CLINICAL LABORATORY MEDICINE

2nd Edition

2002

Lippincott Williams & Wilkins

Philadelphia

530 Walnut Street, Philadelphia, PA 19106 LWW.com

0-683-30751-7

Printed in China

Library of Congress Cataloging-in-Publication Data

Clinical laboratory medicine / editor, Kenneth D. McClatchey.—2nd ed.

p.; cm.

Includes bibliographical references and index.

ISBN 0-683-30751-7

1. Diagnosis, Laboratory. I. McClatchey, Kenneth D.

[DNLM: 1. Laboratory Techniques and Procedures. QY 4 C6414 2001]

RB37.C5897 2001

616.07′56—dc21

2001029271

10 9 8 7 6 5 4 3 2 1

Editor

Kenneth D. McClatchey M.D., D.D.S.

Professor and Chairman
Department of Pathology, Loyola University Medical Center, Loyola University/Foster G. McGaw Hospital, Maywood, Illinois

SECTION CHIEFS

Serhan Alkan M.D.

Associate Professor
Department of Pathology, Loyola University Medical Center, Maywood, Illinois

Emanuel Hackel Ph.D.
Department of Medicine, Michigan State University, East Lansing, Michigan

David F. Keren M.D.

Medical Director
Warde Medical Laboratory; Clinical Professor, Department of Pathology, The University of Michigan, Ann Arbor, Michigan

Elizabeth Lee-Lewandrowski Ph.D., M.P.H.
Clinical Chemistry, Massachusetts General Hospital, Boston, Massachusetts

Kent B. Lewandrowski M.D.
Clinical Chemistry, Massachusetts General Hospital, Boston, Massachusetts

Kenneth D. McClatchey M.D., D.D.S.

Professor and Chairman
Department of Pathology, Loyola University Medical Center, Loyola University/Foster G. McGaw Hospital, Maywood, Illinois

Harold A. Oberman M.D.

Director
Blood Bank and Transfusion Service; Professor, Department of Pathology, University of Michigan Health Center, University of Michigan, Ann Arbor, Michigan

Douglas A. Triplett M.D.

Vice President
Medical Education; Director, Midwest Hemostasis and Thrombosis Laboratories, Ball Memorial Hospital; Assistant Dean and Professor, Department of Pathology, Indiana University School of Medicine, Muncie, Indiana

Daniel L. Van Dyke Ph.D.

Chairman
Medical Genetics, Henry Ford Hospital; Professor, Department of Genetics, Case Western Reserve University, Henry Ford Health Science Center, Detroit, Michigan

Patrick C. J. Ward M.D.

Professor and Head
Department of Pathology and Laboratory Medicine, Duluth School of Medicine, University of Minnesota, Duluth, Minnesota

Gail L. Woods M.D.
Department of Clinical Biology, University of Texas Medical Branch, Galveston, Galveston, Texas

Ruth W. Weinberg

Acquisitions Editor

Raymond E. Reter

Developmental Editor

Colin Warnock

Manufacturing Manager

Thomas Foley

Production Editor

Mark Lerner

Cover Designer

Compositor: Maryland Composition Co., Inc.

CONTRIBUTING AUTHORS

György Ábel M.D., Ph.D.

Director
Clinical Immunology Laboratory, Department of Laboratory Medicine, Lahey Clinic, Burlington, Massachusetts

Serhan Alkan M.D.

Associate Professor
Department of Pathology, Loyola University Medical Center, Maywood, Illinois

Hesham M. Amin M.B., B.Ch., M.Sc.

Hematopathology Fellow
Department of Pathology, Loyola University Medical Center, Maywood, Illinois

F. Philip Anderson Ph.D., D.A.B.C.C., F.A.C.B.

Assistant Professor of Pathology; Co-Director of Clinical Chemistry
Medical College of Virginia Campus, Virginia Commonwealth University, Richmond, Virginia

Mary E. Barkus M.D.
Department of Pathology, Medical College of Virginia, Richmond, Virginia

Patrick G. Beatty M.D.

Professor
Department of Internal Medicine, University of Utah, Salt Lake City, Utah

Miriam G. Blitzer Ph.D.

Professor and Chief
Department of Human Genetics, University of Maryland School of Medicine; and, University of Maryland Medical Center, Department of Pediatrics, Baltimore, Maryland

Bruce F. Bower M.D., F.A.C.E.
Clinical Endocrinology, Hartford Hospital, Hartford, Connecticut

John T. Brandt M.D.

Senior Clinical Research Pathologist
Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana

Eric Brestel M.D.
Department of Medicine, East Carolina University, School of Medicine, Greenville, North Carolina

Arlene M. Buller M.D.
Genetics and Pathology Department, Medical College of Virginia/Virginia Commonwealth University, Richmond, Virginia

Robert W. Burnett M.D.
Department of Pathology, Hartford Hospital, Hartford, Connecticut

Kai Cao M.D.

Scientist I
American Red Cross, National Histocompatibility Laboratory, Baltimore, Maryland

John L. Carey III M.D.
Department of Pathology, Henry Ford Hospital, Detroit, Michigan

Thomas E. Carey Ph.D.

Otolaryngology
Head, Neck and Surgery Department, The University of Michigan, Ann Arbor, Michigan

Aaron Carroll M.D.
Seattle, Washington

Karen C. Carroll M.D.
Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah

Stephen Cordner M.D.
Victorian Institute of Forensic Medicine, Monash University, Melbourne, Australia

Tina M. Cowan Ph.D.

Emanuel Hackel Ph.D.
Department of Medicine, Michigan State University, East Lansing, Michigan

Curtis A. Hanson M.D.

Associate Professor
Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota

Alfred E. Hartmann M.D.

Associate Professor
Department of Laboratory Medicine, University of South Dakota School of Medicine; and, Medical Director, Department of Pathology, Avera McKennan Hospital, Sioux Falls, South Dakota

Gerald A. Hoeltge M.D.

Head
Section of Transfusion Medicine, Department of Clinical Pathology, The Cleveland Clinic Foundation, Cleveland, Ohio

Donald R. Hoffman Ph.D.

Professor
Department of Pathology and Laboratory Medicine, Brody School of Medicine at East Carolina University, Greenville, North Carolina

Earle W. Holmes Ph.D.

Professor
Department of Pathology, Loyola University Stritch School of Medicine; and Director, Special Chemistry, Endocrinology, and Immunoserology Laboratories, Loyola University Medical Center, Maywood, Illinois

James D. Hoyer
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota

Jane C. Huang M.D.

Clinical Assistant Professor of Pathology
Harborview Medical Center, University of Washington, Seattle, WA

Karen James M.D.

Adjunct Associate Professor
Walker College of Business, Appalachian State University, Boone, North Carolina

Stephen G. Jenkins
Department of Pathology, University of Florida Health Science Center, Jacksonville, Florida

Susan T. Johnson M.T. (A.S.C.P.) S.B.B.

Manager
Immunohematology Services, The Blood Center of Southeastern Wisconsin, Milwaukee, Wisconsin

M. John Judd F.I.B.M.S., M.I.Biol.

Professor
Division of Immunohematology, Department of Pathology, University of Michigan Medical Center, Ann Arbor, Michigan

Malek Kamoun M.D., Ph.D.

Director
Clinical Immunology and Histocompatibility Testing Laboratory, Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania; and, Professor, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Jeffrey A. Kant M.D., Ph.D.

Director
Division of Molecular Diagnostics, Department of Pathology, University of Pittsburgh Medical Center; and, Professor, Department of Pathology and Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania

David F. Keren M.D.

Medical Director
Warde Medical Laboratory; and, Clinical Professor of Pathology, University of Michigan, Ann Arbor, Michigan

Rhett P. Ketterling M.D.

Assistant Professor
Department of Laboratory Medicine, Mayo Medical School, Rochester, Minnesota

Franklin P. Koontz Ph.D.

Co-Director of Clinical Microbiology; and Professor
Department of Pathology, University of Iowa Hospitals, Iowa City, Iowa

Alexander Kratz M.D., Ph.D.
Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, Connecticut

Brian D. Kueck M.D.

Assistant Clinical Professor
Department of Pathology, Medical College of Wisconsin, Milwaukee, Wisconsin

Tai C. Kwong Ph.D.

Professor
Department of Pathology and Laboratory Medicine, University of Rochester; and, Director, Toxicology Laboratory, Strong Memorial Hospital, Rochester, New York

Michael Laposata M.D., Ph.D.
Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

Elizabeth Lee-Lewandrowski Ph.D., M.P.H.
Clinical Chemistry, Massachusetts General Hospital, Boston, Massachusetts

Kent B. Lewandrowski M.D.
Clinical Chemistry, Massachusetts General Hospital, Boston, Massachusetts

Brian J. Linehan M.B., Ch.B., Dip., Obst. F.R.C.P.A.
Medlab Hamilton, Hamilton, New Zealand

Naomi L. C. Luban M.D.

Director
Transfusion Medicine, Department of Laboratory Medicine, Children's Hospital; and, Professor, Departments of Pediatrics and Pathology, George Washington University School of Medicine and Health Sciences, Washington, DC

Kenneth D. McClatchey M.D., D.D.S.

Professor and Chairman
Department of Pathology, Loyola University Medical Center and Loyola University/Foster G. McGaw Hospital, Maywood, Illinois

J. Philip McCoy Jr. Ph.D.
Cooper Hospital/University Medical Center, Robert Wood Johnson Medical School at Camden, Camden, New Jersey

Steven Mechanic
Yale University School of Medicine, Blood Bank, New Haven, CT

Jay E. Menitove M.D.

Executive Director and Medical Director
Community Blood Center; Clinical Professor of Medicine, University of Missouri-Kansas City; and, Clinical Professor of Medicine, Kansas University School of Medicine, Kansas City, Kansas

Fredrik Mertens M.D., Ph.D.

Senior Consultant
Department of Clinical Genetics, Lund University Hospital; and, Associate Professor, Department of Laboratory Medicine, Division of Clinical Genetics, Lund University, Lund, Sweden

Derek Middleton Ph.D.

Director
Northern Ireland Histocompatibility and Immunogenetics Laboratory, City Hospital, Belfast, Northern Ireland; and Visiting Professor, School of Biomedical Science, University of Ulster, Coleraine, Northern Ireland

Robert E. Moore Ph.D.

Associate Director of Clinical Chemistry
Department of Pathology and Laboratory Medicine, Hartford Hospitals, Hartford, Connecticut; and Professor, Department of Pathobiology, University of Connecticut, Storrs, Connecticut

S. Breanndan Moore M.D.
Division of Transfusion Medicine, Mayo Clinic, Rochester, Minnesota

Imaad B. Mujeeb M.D.

University of Iowa, College of Medicine, Iowa City, Iowa

Frederick S. Nolte Ph.D.

Laboratory Directory
Clinical Microbiology and Molecular Diagnostics, Clinical Laboratories, Emory University Hospital; and, Associate Professor, Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia

Harold A. Oberman M.D.

Director
Blood Bank and Transfusion Service, Department of Pathology, University of Michigan Health Center; and Professor, Department of Pathology, University of Michigan, Ann Arbor, Michigan

James C. Overall Jr. M.D.

Professor Emeritus
Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, Utah

Catherine G. Palmer Ph.D.

Emeritus Professor
Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana

Darryl E. Palmer-Toy M.D., Ph.D.

Assistant Professor
Department of Pathology, Johns Hopkins Medical Institutions, Baltimore, Maryland

Michael A. Pfaller M.D.

Professor
Department of Pathology, University of Iowa College of Medicine; and Co-Director, Clinical Microbiology Laboratory, Department of Pathology, University of Iowa Health Care, Iowa City, Iowa

Carl L. Pierson Ph.D.

Director
Clinical Microbiology Laboratory, Department of Pathology, University of Michigan Health System; and, Assistant Professor, Department of Pathology, University of Michigan, Ann Arbor, Michigan

Michael Pins M.D.
Department of Pathology, Northwestern University Medical School, Chicago, Illinois

Michele D. Raible M.D., Pharm. D.
Loyola University Medical Center, Maywood, Illinois

Jayashree Ramasethu D.C.H., M.D., M.B.B.S.

Fellow
Department of Pediatrics, Division of Neonatology, Georgetown University Hospital, Washington, DC

James H. Ray Ph.D.

Director
Cytogenetics, Quest Diagnostics, Inc., Teterboro, New Jersey

Barbara S. Reisner Ph.D.
Department of Pathology, University of Texas Medical Branch, Galveston, Texas

Mary J. Reznicek M.D.

Pathologist
Greater Milwaukee Pathologists, Aurora Consolidated Laboratories, West Allis, Wisconsin

Alan Richter Ph.D.
GED Consulting

Roger S. Riley M.D., Ph.D.
Department of Pathology, Clinical Hematology, Medical College of Virginia Hospitals/Virginia Commonwealth University, Richmond, Virginia

Michael G. Rinaldi Ph.D.

Professor
Departments of Pathology and Medicine, University of Texas Health Science Center; and, Director, Veterans Affairs Reference Laboratory, Pathology and Laboratory Medicine, Audie L. Murphy Memorial Veterans Hospital, San Antonio, Texas

Jacquelyn R. Roberson M.D.
Medical Genetics Center, Clara Ford Pavilion, Detroit, Michigan

Marisa Saint Martin M.D.
Loyola University Medical School, Maywood, Illinois

Ramon L. Sandin M.D.

Clinical Pathologist; Medical Director, Clinical Microbiology and Virology Laboratories
H. Lee Moffitt Cancer Center & Research Institute; and, Associate Professor, Department of Pathology, University of South Florida, College of Medicine, Tampa, Florida

Michael A. Saubolle Ph.D., D.(A.B.M.M.)

Director
Clinical Microbiology, Laboratory Sciences of Arizona, Good Samaritan Regional Medical Center, Phoenix, Arizona; and Clinical Associate Professor, Department of Medicine, University of Arizona, Phoenix/Tucson, Arizona

Morton K. Schwartz Ph.D.

Chairman
Clinical Laboratories, Memorial Sloan Kettering Cancer Center; and, Professor, Molecular Pharmacology and Therapeutics, Sloan Kettering Division, Graduate School of Medical Sciences, Cornell University, New York, New York

Stuart Schwartz Ph.D.

Director
Center for Human Genetics Laboratory, University Hospital of Cleveland; and, Professor, Department of Genetics, Case Western Reserve University, Cleveland, Ohio

David L. Sewell Ph.D.

Director
Microbiology Section, Pathology and Laboratory Medicine Service, Veterans' Affairs Medical Center; and, Professor, Department of Pathology, Oregon Health Sciences University, Portland, Oregon

James B. Smart Ph.D.
Department of Pathology, University of Michigan, Ann Arbor, Michigan

Gary Stack M.D., Ph.D.

Assistant Director
Blood Bank, Yale/New Haven Hospital; and, Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut

Peggy J. Stupca M.S.

QA Education Specialist
Laboratory Medicine and Pathology, Mayo Clinic; and Assistant Professor, Department of Laboratory Medicine, Mayo Medical School, Mayo Clinic, Rochester, Minnesota

Deanna A. Sutton M.S., M.T., SM (ASCP), RM, SM (NRM)

Administrative Director
Fungus Testing Laboratory; and, Assistant Professor, Department of Pathology, University of Texas Health Science Center at San Antonio, San Antonio, Texas

Karl S. Theil M.D.

Associate Professor of Clinical Pathology
Department of Pathology, Ohio State University, Columbus, Ohio

John G. Toffaletti Ph.D.

Director of Blood Gas Services
Clinical Laboratories; and, Associate Professor, Pathology and Clinical Laboratories, Duke Medical Center, Durham, North Carolina

Eleanor M. Travers M.D.
Potomac, Maryland

Douglas A. Triplett M.D., F.A.C.P., F.C.A.P.

Vice President
Medical Education, and Director, Midwest Hemostasis and Thrombosis Laboratories, Ball Memorial Hospital; and, Assistant Dean and Professor, Department of Pathology, Indiana University School of Medicine, Muncie, Indiana

Daniel L. Van Dyke Ph.D.

Chairman

Medical Genetics, Henry Ford Hospital; and Professor, Department of Genetics, Case Western Reserve University/Henry Ford Health Science Center, Detroit, Michigan

Patrick C. J. Ward M.D.

Professor and Head
Department of Pathology and Laboratory Medicine, University of Minnesota, Duluth School of Medicine, Duluth, Minnesota

Jeffrey S. Warren M. D.

Warthin-Weller Endowed Professor
Department of Pathology, University of Michigan Medical School; and, Director, Division of Clinical Pathology, University of Michigan Hospitals, Ann Arbor, Michigan

Anne Wiktor B.S.M.T.

Senior Technologist
Cytogenetics, Henry Ford Health System, Detroit, Michigan

David S. Wilkinson M.D., Ph.D.

Professor and Chair
Department of Pathology, Virginia Commonwealth University; and, Chair, Department of Pathology, Medical College of Virginia Hospitals, Richmond, Virginia

Gail L. Woods M.D.
Clinical Microbiology, University of Texas Medical Branch, Galveston, Galveston, Texas

Alan H. B. Wu

Director
Clinical Chemistry, Department of Pathology and Laboratory Medicine, Hartford Hospital, Hartford, Connecticut; and Professor, Laboratory Medicine, University of Connecticut Health Center, Farmington, Connecticut

William A. Wyatt B.A.

Assistant Supervisor
Department of Laboratory Medicine, Mayo Clinic, Rochester, Minnesota

Jane M. Yang M.D.
Department of Chemistry, Massachusetts General Hospital, Boston, Massachusetts

Chester M. Zmijewski Ph.D.

Emeritus Professor
Departments of Pathology and Laboratory Medicine, University of Pennsylvania Medical Center, Philadelphia, Pennsylvania

DEDICATION

To my wife, Martha, and our children: Sean, Suki, Suni, and Stephen. All of whom provide constant support yet remind me of the balance that must be maintained in one's life —no matter what the venture.

PREFACE

It is an honor to be the editor of the second edition of *Clinical Laboratory Medicine*. This textbook should ultimately educate a wide variety of readers—from medical laboratory technicians, medical technologists, medical students, residents in pathology and internal medicine—to clinical laboratory scientists, pathologists, and other physicians with an interest in clinical laboratory medicine.

From beginning to end, this textbook is devoted to providing the reader with the latest in clinical laboratory medicine in an organized, well-illustrated manner. The general outline of every chapter in the text is as follows:

I. Introduction
II. Disease status (including pathophysiology)
III. Diagnostic methods
IV. Interpretation (including sensitivity, specificity, precision, and accuracy issues, where applicable)

Since the late 1950s, there has been a rapid increase in clinical laboratory automation. The concomitant emergence of quality management issues and molecular pathology appplications across multiple disciplines has made it difficult for textbooks in clinical pathology and clinical laboratory medicne to keep up with the advances in this diverse science. There was even difficulty in defining the new technologies. This textbook has been designed to include as many of the “new” technologies as is humanly possible at this time. For example, the “new” technologies in this revision include molecular pathology, cytogenetics, HLA, molecular biology, and even such areas in the general section as quality management, cost accounting, and informatics.

The eleven sections of this textbook start with a beautiful color plate and a short introduction. Chapters are designed to allow the reader to focus on necessary information quickly and efficiently. Finally, the section editors and chapter authors feel that this text provides a relevant, comprehensive, scientific, and artful approach to the ever-widening scope of clinical laboratory medicine. We welcome the readers' input regarding this second edition so that future editions will be even better.

Kenneth D. McClatchey M.D., D.D.S.

ACKNOWLEDGMENTS

It is with deep satisfaction and gratitude that I acknowledge the collaboration of my esteemed colleagues as section editors. Each section editor has performed his or her tasks in a diligent and professional manner. In addition, the chapter authors have completed, in a timely and professional manner, chapters that bring to the reader the latest science from the ever-widening universe of clinical laboratory medicine. They have accomplished their task using a format that allows both students and practicing health professionals to efficiently learn the many disciplines of clinical laboratory medicine. The textbook would not have been designed and written if I had not been influenced during my career by laboratory professionals such as Drs. John G. Batsakis, and Adam J. French. Their inspiration to pursue a career in clinical laboratory medicine was immeasurable.

I would also like to thank Ray Reter and Ruth Weinberg at Lippincott Williams & Wilkins.

Finally, and most importantly, *Clinical Laboratory Medicine* is written to acknowledge the many laboratory professionals who have devoted their lives to better patient care by constantly striving to improve the art and science of clinical laboratory medicine.

Contents

Section 1 General Laboratory

General Laboratory - Introduction

Kenneth D. McClatchey

Section Chief

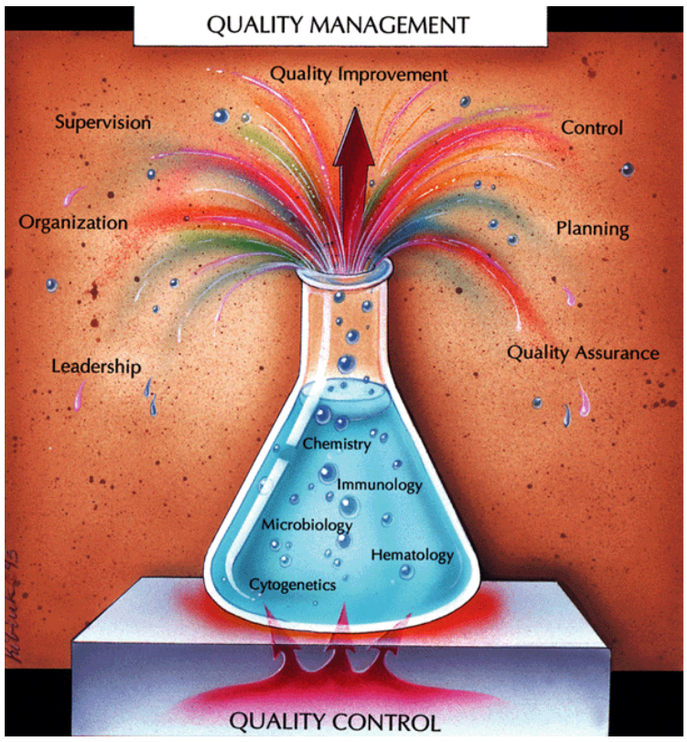


Figure.

1 Basic Laboratory Management

2 Business Management of the Clinical Laboratory

3 Laboratory Safety

4 The Interpretation of Laboratory Tests

5 Laboratory Information Systems

6 Ethics in Laboratory Medicine

Management in the clinical laboratory consists of the many processes of planning and organizing, leading the efforts of staff, and using other laboratory and hospital resources to achieve stated goals. As clinical laboratory medicine continues to grow as a service industry, modern management theory and practice have become more and more a part of our lives. At the same time, patients faced with escalating medical costs, are insisting that they receive value for their health insurance dollar. Such requirements are vividly reflected in the actions of hospital and laboratory accrediting agencies, as well as payors such as insurance companies. It is no longer satisfactory to ignore the business aspects of our practices. Now we must know the language of quality assurance, quality improvement, quality systems utilization management, cost analysis, billing, collecting, etc. In addition, being a responsible manager requires a solid foundation in risk management and laboratory safety. Furthermore, the manager, as a health professional, must provide accurate laboratory data for his or her patients, knowing full well that variability is a part of the preanalytic and analytic test process. This variability must be kept to a minimum, yet it should be clearly explained, when necessary, to colleagues. In the laboratory of today and tomorrow, laboratory information systems are the backbone of success. The ability to communicate accurately in the high-volume environment of a clinical laboratory is a prerequisite for successful management of a clinical laboratory. Knowing the basic requirements of laboratory information systems (as well as having a basic knowledge of contract negotiation for information systems) allows the health professional an opportunity to control his or her destiny in result reporting, order entry, archiving, test menu selection, etc.

This section is intended to acquaint the reader with selected basic management skills that will hold him or her in good stead as quality management becomes more and more a part of the laboratory professional's life.

1

Basic Laboratory Management

Eleanor M. Travers

Kenneth D. McClatchey

GENERAL MANAGEMENT ISSUES

Part of "1 - Basic Laboratory Management"

Laboratory professionals in the 2000s must be versatile and effective as medical scientists, managers of people, clinical consultants, and business managers. Without the integration of these skills and attributes, and the active participation of pathologists, the profession of laboratory medicine is likely to be drawn into increasing conflict with other health professions and health administrative agencies. Patients, clinical laboratories, and clinicians will not be well served by such divisions (1,2). In fact, without flexibility and adaptability, clinical laboratory managers might find themselves quite frustrated in today's "bottom-line" environment.

After World War II, clinical laboratory medicine grew rapidly, but managers primarily focused on only quality control (1,2). Quality control provides assurance that a laboratory functions properly for the benefit of the patient using daily or run-to-run data points plotted over time on charts (3). However, there has been a major paradigm shift to business practice in the philosophy and the role of the laboratory manager. An astute laboratory manager will respond to this paradigm by plotting the laboratory's financial data points as often as quality control and other performance indicator data points.

With the advent of clinical laboratory reimbursement losses and increasing scientific discoveries in medical practice in the 1990s, the control of budgeted dollars became a major problem for managers of clinical laboratory services. By the late 1990s, the spiraling growth of laboratory medicine tests and services was increasingly viewed as a major financial loss by third-party reimbursement agencies and payers.

In reality, what happened was that the health care industry, including laboratory medicine, sped toward a trillion dollar price tag by the turn of the century despite continued efforts to control costs.

At the same time and at the same breakneck pace, reimbursement policies and mechanisms for laboratory medicine were developed to create disincentives for the use of tests and services. During the decade of the 1990s, both federal and private payers severely limited payments to both providers and laboratories, which had negative impacts on laboratory management and practice (1,4). The events leading up to the era of reduced payment (reimbursement), the framework for the punitive policies of the late 1990s, are summarized thusly:

1960s

Medicare: Publicly funded health insurance for all citizens over age 65)

Medicaid: Health insurance for citizens receiving public assistance

1970s

Health maintenance organizations (HMOs) are established)

Professional standard review organizations are established

1980s

Diagnosis-related group (DRG) system of reimbursement is introduced)

Deficit Reduction Act developed to control outpatient fees, including laboratory fees

HMOs begin to flourish, with enrollment of more than 30 million

Preferred provider organizations (physicians and hospitals contracted with agencies for discounted services) develop rapidly

Independent practice associations (loosely structured groups of physicians and hospitals contracted to provide discounted services) become established

Exclusive provider organizations (strict contracts with physicians and hospitals to provide discounted services to a select group of patients) make their appearance

1990s

Resource-based relative value units, a complex reimbursement program for physician services that is based on overall work effort and includes a correlation of multiple disciplines of medicine, are established to replace DRGs)

Clinical Laboratory Improvement Amendments of 1988 (CLIA '88), which applies to laboratories licensed for interstate commerce under the Clinical Laboratory Improvement Act of 1967 as well as laboratories certified to receive reimbursement from Medicare and Medicaid, were implemented through the publication of final regulations in 1993. Major provisions of the final regulations include

Laboratory administration

Proficiency testing

Patient test management

Quality control
Personnel standards
Quality assurance
Laboratory accreditation
Enforcement and sanctions

This, above all other federal regulatory actions, set the standard for reimbursement policy in 2000 and beyond. No laboratory is allowed to receive reimbursement from Medicare/Medicaid unless it is CLIA licensed or has received a waiver from CLIA.

The Balanced Budget Act of 1997 (BBA97) and the Balanced Budget Reconciliation Amendments of 1999 profoundly affected the volume, growth, and reimbursement of clinical laboratories, mostly in a negative manner. Laboratory managers may experience increased workload for preventive tests (e.g., diabetes, cancer) as a result of the BBA97

Basic Management Concepts

In laboratories today, there are four basic economic problems: money, machines, management, and manpower. Economists refer to the same four topics as capital, technology, management, and labor. If one thinks about the problems facing laboratory managers today, the same broad areas are applicable.

Management is probably the world's oldest unheralded profession. In *The History of Management Thought* (5), Claude S. George, Jr., traces the profession of management from the ancient Egyptians, who engaged in planning, organizing, and controlling the activities of their numerous workers, to China, where planning and control systems were established around 1100 BC; to the United States, where Joseph Wharton established college courses in business management at the University of Pennsylvania in 1881. George also wrote that in 400 BC, Socrates noted the universality of the management function in human behavior. In 350 BC, Plato established the principle of specialization for human efficiency. In 1436, the Arsenal of Venice employed more than 1,000 people. They used accounting systems, planning, inventory control, assembly line techniques, interchangeable parts, and a formal system of personnel management to build ships and armament. The first “clinical” manager was probably Robert Owen, an industrialist in Scotland who, in the 1820s, first tackled the problems of productivity and motivation, the relationship of worker to work, worker to enterprise, and worker to management. These basic problems are still applicable in management venues in health care and the clinical laboratory.

The science of management in the United States began in the late 1800s with the efforts of Frederick W. Taylor. Taylor defined his work on scientific management as a systematic or scientific investigation of all the facts and elements of the work being managed. It was the antithesis of “management by tradition,” which was the rule before his work (and incidentally used widely today in medicine in teaching clinical skills). Beginning with the industrial revolution to today's explosion of science and technology, management theory evolved, bringing with it the involvement of the tools of the behavioral sciences. As a discipline, it is young—barely a century old (5,6).

Certainly management is a science, but more important, it is people interacting with one another. As Peter F. Drucker states,

> Every achievement of management is the achievement of a manager. Every failure is a failure of a manager. People manage, rather than forces or facts. The vision, dedication, and integrity of managers determine whether there is management or mismanagement. Management must always be done in organization—that is, with a web of human relations (7).

Interestingly, Drucker further compares a manager to a teacher: “Only a teacher has the same twofold dimension, the dimension of skill and performance and the dimension of personality, example, and integrity” (7).

Managers come in different forms and with different labels: administrator, commander, and executive, for example. They have evolved with the growth of institutions in the industrialized world. Drucker (7) states that in the early 1900s, people asked, “What do you do?” Today they tend to ask, “Whom do you work for?” (7).

The science of management in this country that has been practiced for decades is being challenged. The traditional scientific approach to management is changing; learning to look on people as human resources and opportunities rather than problems, costs, and threats. In such approaches, the art of leadership is critical (8,9,10 and 11).

According to Max DePree (11), leadership is “liberating people to do what is required of them in the most effective and humane way possible.” A leader, according to DePree, has the following attributes:

- Has consistent and dependable integrity
- Cherishes heterogeneity and diversity
- Searches out competence
- Is open to contrary opinion
- Communicates easily at all levels
- Understands the concept of equity and consistently advocates it
- Leads through serving
- Is open to the skills and talents of others
- Is intimate with the organization and its work
- Is able to see the broad picture (beyond his or her own area of focus)
- Is a spokesperson and diplomat
- Can be a tribal storyteller (an important way of transmitting corporate culture)
- Tells why rather than how

DePree goes on to state that the measure of leadership is “not the quality of the head, but the tone of the body” (11). The sign of outstanding leadership appears primarily among the followers. It is fundamental that leaders endorse a concept of persons, thus an understanding of the diversity of people's gifts, talents, and skills.

There are two types of relationships in our industrialized society that influence the management of people: *contractual* and *covenantal*.

Contractual relationships cover the *quid pro quo* of working together. Contractual relationships break down during the demands of conflict and change. Alexander Solzhenitsyn states that

whenever the tissue of life is woven of legalistic relationships, an atmosphere of spiritual mediocrity that paralyzes men's noblest impulses is created (11).

Covenantal relationships induce freedom, not paralysis. A covenantal relationship rests on shared commitment to ideas, issues, values, goals, and management processes. Covenantal relationships are open to influence (11).

A good manager studies to improve his/her management skills and applies them in daily practice. A good manager is also an entrepreneur; one who tends to accept change as part of the business of managing. A good manager is a "high-performance" manager (6). The high-performance manager is:

A strategist. One who looks to the future, makes educated guesses about the major forces and trends he or she can see, and interprets them in terms of opportunities for growth and progress.

A problem solver. One who clearly perceives the differences between the anticipated future and the unfolding present and who decides what must be done with those factors under his or her control to influence the environment or to adapt to it most effectively.

A leader. One who offers those who answer to him or her a clear course of action that will gain their commitment and serve their individual objectives as well as the higher objectives of the organization.

A teacher. One who guides others and helps them to identify and solve problems so that they can perform their tasks effectively and can develop themselves as individuals as well as workers.

It is of utmost importance in the management process to realize that the human resources of an organization make the organization successful. Human resource management is an art form. The basic components of human resource management are motivation, delegation, and supervision. *Motivation* is a basic psychological need of any employee, thus an important task of any manager is to create motivational work environments. *Delegation* is the critical ability to place tasks and responsibilities at the level at which they can best be accomplished. *Supervision* is the combination of technical and human relations skills that ensures that organizational goals are met and policies followed.

The management process has been artfully diagramed by Hardwick (1) to delineate the tasks associated with the basic elements of management: ideas, things, and people (Fig 1.1).

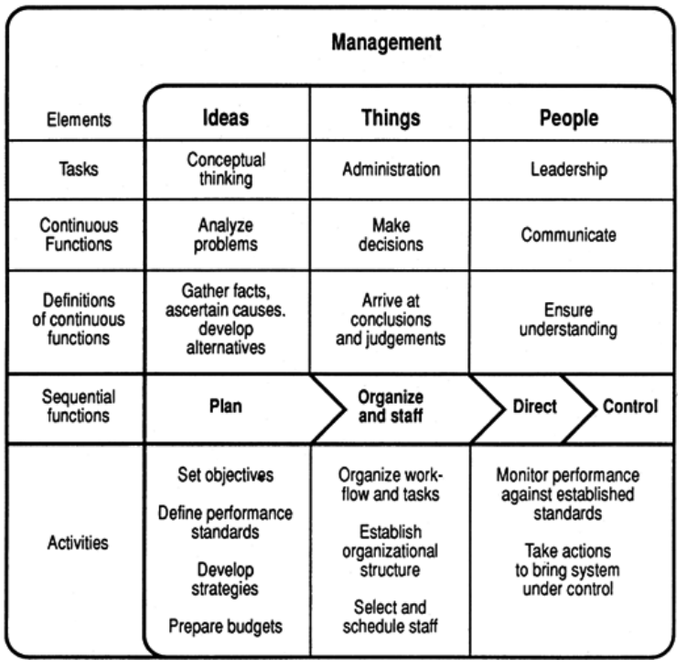


FIGURE 1.1. The management process. (From Hardwick DF. *Directing the clinical laboratory*. New York: Field and Wood, 1990:2, with permission.)

The Evolution of Quality Management Concepts

The evolution of quality concepts begins with the publication of Shewhart's *Economic Control of Quality of Manufactured Product* in 1931 (12). This book put quality on a scientific footing. Shewhart developed the concept of statistical control and describes it as follows:

> A phenomenon will be said to be controlled when through the use of past experience we can predict, at least within limits, how the phenomenon may be expected to vary in the future. Here it is understood that prediction means that we can state, at least approximately, the probability that the observed phenomenon will fall within the given limits (12).

An outgrowth of such a method was the process control chart, well known in every clinical laboratory in this country (Fig. 1.2).

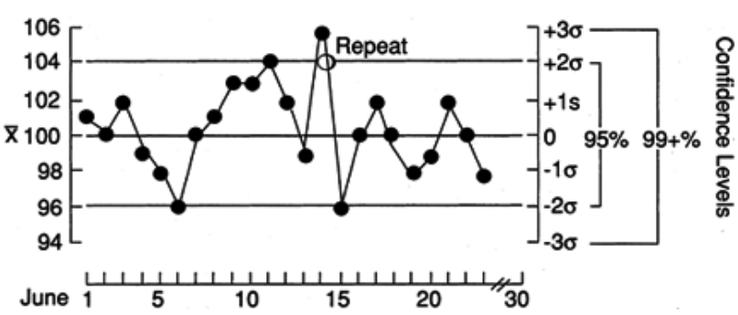


FIGURE 1.2. Levey-Jennings or Shewhart process control chart. (From Weisbrot IM. *Statistics for the clinical laboratory*. Philadelphia: JB Lippincott, 1985:64, with permission.)

Joseph Juran worked with Shewhart at Bell Laboratories and eventually played a prominent role in the evolution of quality

management from the 1960s through the 1980s. In 1951, Juran published a book entitled *Juran's Quality Control Handbook* (13). He subsequently developed a system of management called “managing for quality” (14). His system is based on what he calls the Juran trilogy:

- Quality planning
- Quality control
- Quality improvement

Quality planning includes

- Determining who the customers are
- Determining the needs of the customers
- Developing product features that respond to customer needs
- Developing the processes to produce the product features
- Transferring the plans to the operating forces

Quality control includes

- Evaluating actual product performance
- Comparing actual performance to product goals
- Acting on the difference

Quality improvement includes

- Establishing the infrastructure
- Identifying the improvement projects
- Establishing project teams
- Providing the teams with resources, training, and motivation to:

  diagnose the causes

  stimulate remedies

  establish controls to hold the gains

The Juran trilogy can be diagramed to demonstrate how beneficial change can be used to develop improved levels of performance (Fig. 1.3). There is no question that improved levels of performance place an increased workload on all levels of the management team. An example of workload allocation is demonstrated with the Itoh model. In the Itoh model, time is allocated to (a) development, (b) improvement, and (c) control and maintenance throughout the management structure (Fig. 1.4) (14).

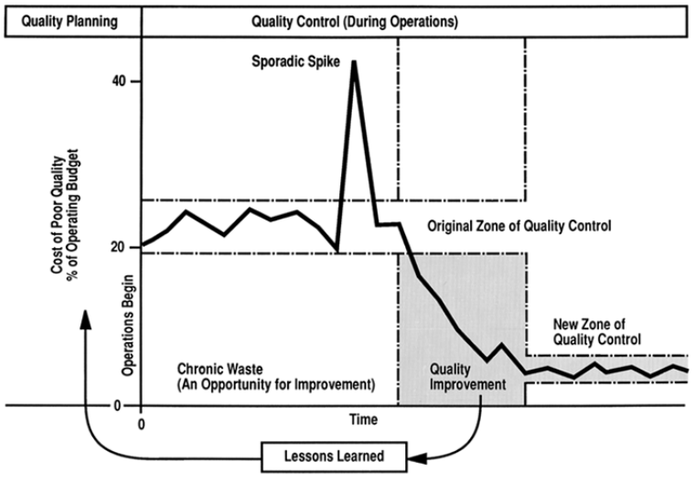


FIGURE 1.3. The Juran trilogy. (From Juran JM. *Juran on planning for quality*. New York: The Free Press, 1988:12, with permission.)

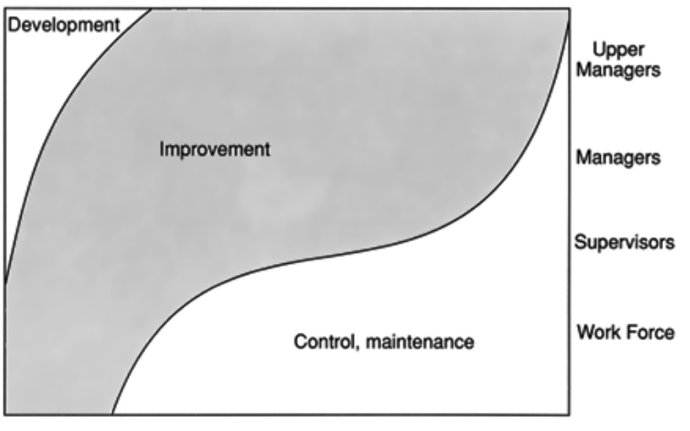


FIGURE 1.4. The Itoh model. (From Juran JM. *Making quality happen: upper management's role*. Wilton, CT: The Juran Institute, 1993, with permission.)

After World War II, W. Edwards Deming, another associate of Dr. Walter Shewhart, also espoused a philosophy of management that calls for organizations to produce products and services that help people live better. Deming believes that real profits are generated by loyal customers. In addition, Deming radically departs from the classic relationships among quality,

costs, productivity, and profit. According to Deming, as quality is increased, cost decreased due to elimination of repeats and waste. Improved quality leads to lower costs and higher productivity. Lower costs and higher productivity lead to lower prices. A competitive edge, it is hoped, is established (15).

"Companies," including hospitals and clinical laboratories, that follow Deming's philosophy have some distinct differences from the "standard" company (Fig. 1.5). Advocates of Deming's quality management philosophy believe that it is workers who pay for management errors—when profits go down, management cuts back on labor costs. They also believe that quality is difficult to define, yet the lack of quality is easy to define. The Deming philosophy ultimately describes quality as "anything that enhances the product from the viewpoint of the customer." In laboratory practice, there are two customers: the physician (the patient's advocate) who orders tests and the hospital/corporation that uses the services provided by laboratories to make revenue and measure contribution (nonprofit) or profit (for profit).

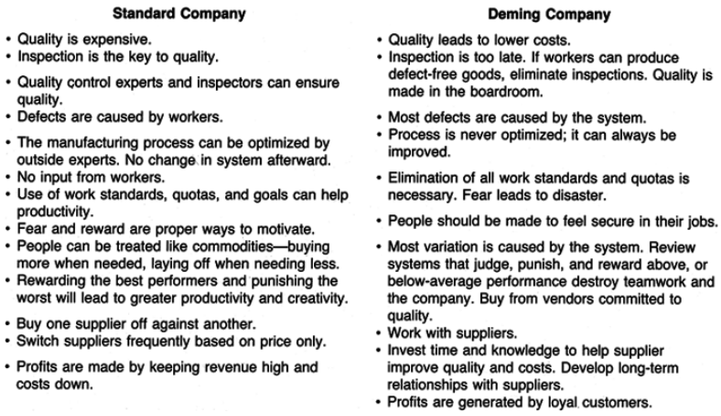


FIGURE 1.5. Comparison of Deming management principles and traditional management principles.

Development of a Service Economy

Integrated with the evolution of management in this country is the transition from an industrial economy to a "service economy." The evolution of a service economy in the United States parallels the increasing importance of quality improvement programs (16,17 and 18). A service economy is also directly related to the purchasing power of the targeted market population. The health industry has grown because both the consumers (patients with health insurance) and the providers who order laboratory tests and services have no information on the true costs of health care. Naisbitt (19) noted that in 1956 for the first time in American history, white-collar workers in technical, managerial, and clerical positions outnumbered blue-collar workers. He called such a change the "era of the information society" (Fig. 1.6). The "information society," throughout its evolution and maturation, requires a service economy. The "classic" management pyramid structure (Fig. 1.7), therefore, is now an inverted pyramid (Fig. 1.8) because the workers are empowered to participate in the management decisions that affect the workplace.

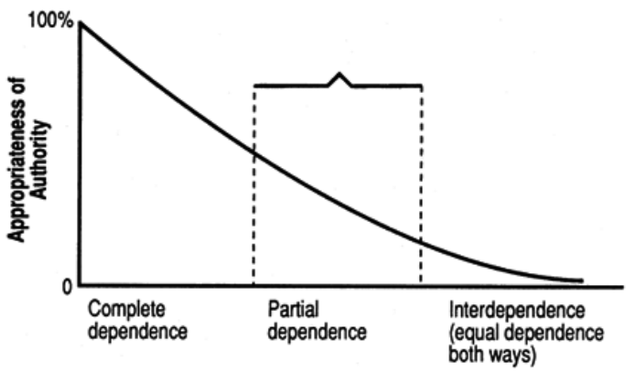


FIGURE 1.6. Dependence of subordinates in U.S. industry today. (From Naisbitt J. *Megatrends: ten new directions transforming our lives*. New York: Warner Books, 1982, with permission.)

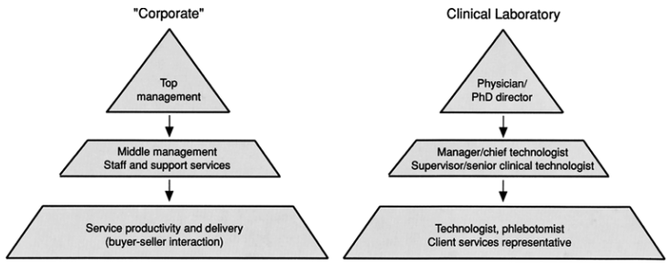


FIGURE 1.7. Classic management structure.

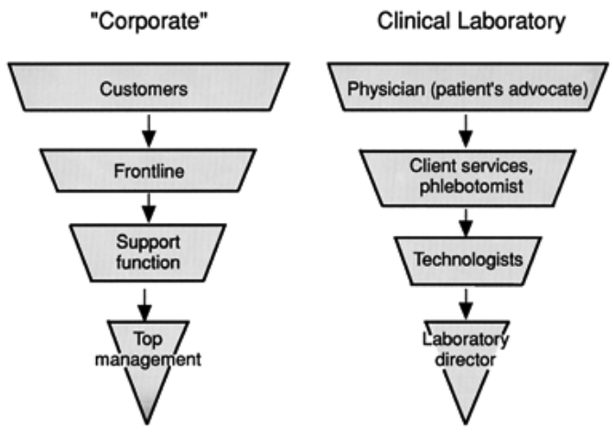


FIGURE 1.8. Service-oriented management structure.

A service economy meets customer expectations in the course of selling and postsales activity as well as providing a series of functions that match or better the competition in a way that provides an incremental profit for the producer of laboratory tests and services. For most services, four basic characteristics can be identified:

1. Services are more or less intangible.
2. Services are activities or a series of activities rather than things.
3. Services are at least to some extent produced and consumed simultaneously.

4. The customer participates in the production process at least to some extent.

In 1990, Christian Gronroos (18) described the “hidden service sector” and the “official service sector” in our economy. Such service sectors have direct relationships with laboratory medicine. Gronroos states that today, service firms have come to realize that competition is now so severe that mere technical solutions offered to customers are not sufficient to create a competitive edge. Further, Gronroos points out that in so-called “industrial sector” of the economy, manufacturing firms must offer customers a variety of services, as an integral part of their total offering, to succeed in business and survive competitive threats.

If one analyzes laboratory medicine today, especially in the outpatient environment, the outcome of the analysis is the same; performing a package of tests at an average cost and selling them at a reasonable price is no longer sufficient to create a competitive edge. An important reason for this is the array of “hidden (indirect) costs” required in the outpatient environment, which include information services (i.e., computerized order entry, line connects, printers, and terminals), couriers, specimen containers, laboratory consultation costs of marketing and advertising, the costs of regulatory compliance, inspection, accreditation, and retraining staff, and many other costs to run the business.

Another way to address the issue is that the clinical laboratory “test” in today's service economy, especially in the ambulatory care setting, is not simply a number or a brief statement; it is a cluster of issues that surround the laboratory test, such as:

- The ethical and moral responsibility to provide the best care for each patient,
- The need to rapidly provide medical information to a provider,
- The seller's (the laboratory or its representative) need to avoid financial losses,
- The organization represented by the laboratory (e.g., private corporation, hospital, university, and clinic) and its need to avoid financial losses,
- The laboratory service's reputation, which needs to be maintained in the eyes of the public it serves,
- The laboratory service's personnel and management's need to retain highly qualified staff,

TABLE 1.1. SYNOPSIS OF EXTERNAL (MACRO) AND INTERNAL (MICRO) LABORATORY MANAGEMENT BENCHMARKING PROGRAMS (2)

Programs	Scope of Program	Management Focus	Primary Users	Management Uses
Global management				
1. Labtrends™	External	Macro	Hospital administrators Laboratory directors/managers Laboratory section supervisors	Operating characteristics and performance Operating characteristics and performance Operating characteristics and performance problems
2. Laboratory Management Index Program	External	Macro	Laboratory directors Laboratory managers Laboratory section supervisors Hospital administrators	Global laboratory productivity, efficiency, utilization, and cost effectiveness
3. Mecon-Peer	External	Macro	Hospital administrators Laboratory directors/managers Laboratory section supervisors	Performance benchmarks Identify improvement opportunities Network operational issues
4. Data Comparison Reporting System (DCRS)	External	Macro	Hospital administrators in voluntary hospitals Laboratory directors/managers	Decision support systems Comparative financial and resource utilization data
5. Canadian Institute for Health Information (CIHI)	External	Macro	Hospital administrators (Canada) Laboratory directors/managers Laboratory section supervisors	Budgeting, productivity, utilization, efficiency
6. Welcan UK System	Internal	Micro	Section supervisors in chemistry, hematology, cytogenetics, histopathology, immunology, microbiology, and automated testing	Work load unit reporting
Labor/equipment management				
1. Equipment Test Labor Productivity index (ETI)	Internal	Micro	Laboratory managers Laboratory section supervisors Lead technologists	• Finds *total productive* minutes of time, based on lab's own unique equipment and methods • Finds total *billable test* productive time
2. CAP Workload Recording Method (Discontinued in 1993)	Internal	Micro	Laboratory managers Laboratory section supervisors Lead technologists	• Finds *average analytical* minutes used for staffing decisions Comparison of equipment and methods Labor cost estimates for equipment/methods

This table provides a synopsis of the currently available laboratory management bench-marking programs in the United States, showing their scope, management focus, primary users, and management uses. Management bench-marking programs are important for cost control in a reimbursement environment demanding lowered utilization and higher productivity.
From Travers EM. *Clinical laboratory management*. Baltimore: Williams & Wilkins, 1996:663, with permission.

- The buyer (the physician as the patient's agent) of laboratory tests and services,
- The organization represented by the buyer,
- The image of both organizations in the marketplace.

It is argued that ultimately the intangibles—the service activities that companies or laboratories offer their customers or patients—may be the most important source of value added in the service economy. It has been noted by Albrecht and Zemke that the marketplace prefers to do business with those who serve and declines involvement with those who only supply materials and goods (17).

The first evidence of interest in service quality began in the 1980s. At the same time, interest in quality assurance in laboratory medicine grew very rapidly. Quality assurance applications for laboratory medicine are currently updated and well defined by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) (20,21) and the federal mandates of the Clinical Laboratory Improvement Act amendments.

At the core of a successful service economy and, coincidentally, at the core of a successful quality improvement program are four commonly shared principles:

Involvement: Recognition by management that a situation exists.

Measurement: Necessary but cannot be used alone as a measurement of success or failure.

Reward: The success of employee-owned companies demonstrates the need for rewards.

Follow-through: Successful service economies or successful quality improvement programs become a way of life.

The overall success of these entities is incumbent on the leadership of management. Service and quality improvement require management's personal, visible involvement.

In discussing service management and marketing, Gronroos (18) observed that “far too often people in an organization view customers as an abstract phenomenon or a mass that is present somewhere. Customers are seen in terms of numbers. When someone stops being a customer there are always new potential customers to take their place. Customers, individuals and organization alike are numbers only.” In reality this is, of course, not true. Every single customer forms a customer relationship with the seller that the firm has to develop and maintain (18). This applies to laboratory practice in its managerial sense, i.e., the laboratory manager must make his or her customers (the physicians) happy with the laboratory's services. If one were to substitute the physician (the patient's advocate) for the customer and the hospital (or corporation) for the organization, this discussion could not be more timely, as health care and clinical laboratory medicine jointly struggle to emphasize the clinical importance of laboratory tests and services, while decreasing their costs.

Gronroos developed a “customer relationship life cycle” (CRLC) that can be modified or adopted to health care and clinical laboratory medicine. In the CRLC, the initial stage is when the customer establishes a need that a firm may be able to satisfy. The customer subsequently makes him- or herself aware of a firm's services. A first purchase follows if all is viewed positively. Subsequent purchases follow if a positive relationship continues. “During this process the customer may observe the firm's ability to take care of his or her problems and provide services which the customer determines to have an acceptable technical and functional quality.” Such a service certainly fits the patient's health care provider “life cycle” (Fig. 1.9).

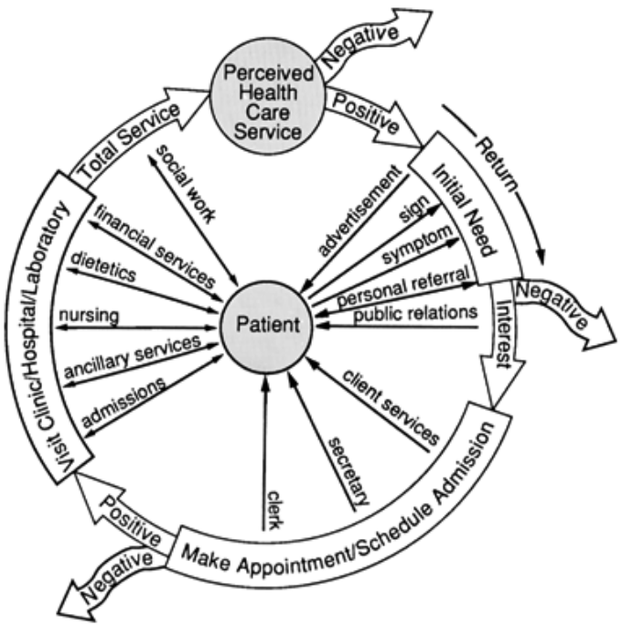


FIGURE 1.9. Patient-health care provider “life cycle.” (Modified from Gronroos C. *Service management and marketing*. Lexington, MA: DC Heath, 1990:130.)

Theory X, Theory Y

In 1960, Douglas McGregor (22) contrasted two types of companies. The first, X, was organized to demonstrate that no one works unless they have to or are made to work. Because of this human characteristic of dislike of work, most people must be coerced, controlled, directed, or threatened with punishment to get them to put forth adequate effort toward the achievement of organizational objectives. Thus, theory X assumes that the average human being prefers to be directed, wishes to avoid responsibility, has relatively little ambition, and wants security above all else.

McGregor's alternative theory, Y, demonstrates that, on the contrary, people love to work and, given the right conditions, will strive to do their very best. Theory Y leads to a preoccupation with the nature of relationships, with the creation of an environment that will encourage commitment to organizational objectives and will provide opportunities for the maximum exercise of initiative, ingenuity, and self-direction in achieving them. McGregor further explains that the appropriateness of authority varies as a function of dependence (Fig 1.6). For example, when the dependence in a relationship is complete, such as between a parent and child, authority can be used almost exclusively. Conversely, when dependence is almost equal, such as two staff physicians working on the same patient, authority is useless as a means of control. In today's clinical laboratory with numerous

highly trained professionals, the predominant relationship is one of partial dependence; thus, persuasion and application of professional guidelines are the ways to achieve goals.

Toughness is the hallmark of theory X. A company or laboratory that relies on the objectives of theory X applies hire-and-fire techniques.

Conversely, theory Y companies or laboratories operate with a velvet glove. Theory Y gives people room to exercise initiative instead of strict supervision. With the development of quality management systems such as continuous quality improvement and total quality, theory Y companies and laboratories are in vogue.

Further evidence for promoting quality management programs is provided by the widely accepted need-hierarchy theory developed by Abraham Maslow (23). Maslow describes an ascending hierarchy of human needs, which each individual strives to achieve in his or her personal and work environments (Fig. 1.10):

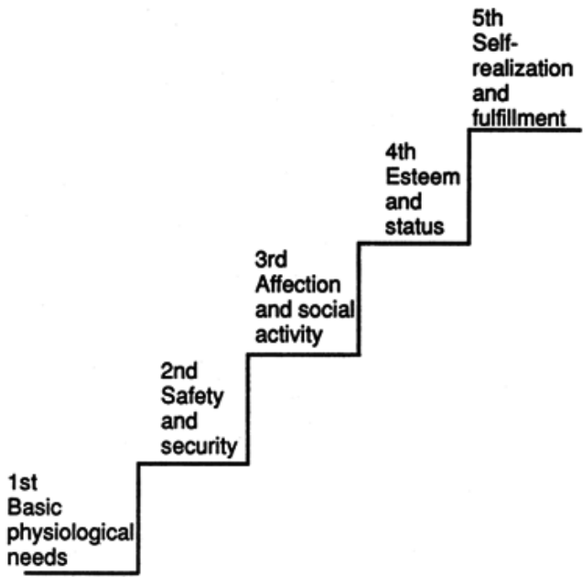


FIGURE 1.10. Hierarchy of human needs. (From Maslow A. *Psychol Rev* 1943;50:370–396, with permission.)

1. Physiologic or survival needs
2. Safety or security needs
3. Social needs
4. Esteem or ego needs
5. Self-actualization or self-fulfillment needs

Modern methods of quality management address these needs.

Continuous Quality Improvement

As quality assurance has matured, there has been a shift to a more consumer-oriented, more positive approach to health care management. Such an approach is called continuous quality improvement (CQI). CQI is a consumer-oriented, proactive response to negative public perceptions of the health care business (Fig. 1.11).

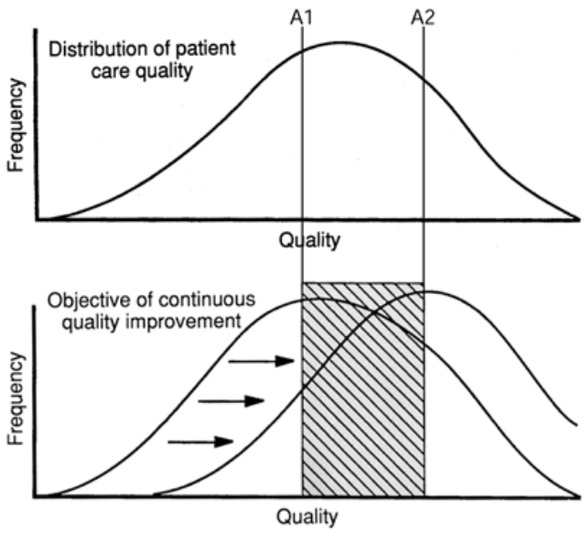


FIGURE 1.11. Shifting the mean: quality assurance to AI. **A**, quality assurance; **B**, continuous quality improvement.

As described by Richard E. Thompson, M.D. (24), there are 11 essential points that differentiate quality assurance from CQI:

1. The goal of CQI is to *improve* the norm of practice and behavior rather than to strive to comply with a standard based on normative behavior.
2. CQI uses supportive, positive approaches to change behavior rather than more investigative techniques.
3. In CQI, individual and corporate attitudes are as critical and as valid and/or useful as quality assurance data.
4. In CQI, performance indicators rather than quality criteria may be easier to define and implement.
5. CQI, using a cooperative effort, avoids the separatist approach to quality assurance, utilization review, and risk management that is often seen in hospital management settings.
6. CQI, because of the overall approach, is open to applications research.
7. CQI methods appear to be less complex than quality assurance methods simply because many quality assurance methods are developed by a committee.
8. CQI changes traditional medical staff functions such as credentialing and peer review to a patient-protective methodology.
9. In contrast to the absolute confidentiality of quality assurance, CQI allows public comparison of physicians and/or hospitals without providing litigious information to anyone.
10. CQI affects medical staff bylaws and organization structure by allowing physicians to feel comfortable about “taking responsibility for yourself.”
11. CQI provides a system that the JCAHO will find *functionally* useful.

Quality System Management

Recently, quality system management has gained attention as a method to provide consistent, high-quality, cost-effective laboratory medicine. Much of a quality system approach to management is based on a comprehensive/coordinated effort to meet quality objectives modeled after the International Organization for Standardization 9000 series quality standards.

A quality system approach to management requires that clinical laboratory personnel have a clear vision of the essentials of what needs to be managed, for example:

- organization
- personnel
- equipment
- purchasing and inventory
- process control
- documents and records
- occurrence management
- internal assessment
- process improvement
- service and satisfaction

Typically the essentials of a quality system can be applied to all aspects of a laboratory operation.

Modeled after the National Committee for Clinical Laboratory Standards approved guideline, *A Quality System Model for Health Care*, a laboratory should apply the essentials of what needs to be managed to a laboratory's path of work flow (25).

The ultimate goal for such a management approach is to provide an environment that the managers of clinical laboratories can apply to all the services in the clinical laboratory environment.

JCAHO—The Agenda for Change

In 1975, the Joint Commission on Accreditation of Hospitals (now called the JCAHO) published in their *Accreditation Manual for Hospitals* a requirement "to demonstrate that the quality of care was consistently optimal by continually evaluating care through reliable and valid measures." Subsequently, in 1980 the Joint Commission adopted the quality assurance standard for hospitals. The 1980 quality assurance standard (a) emphasized the value of a coordinated hospital-wide quality assurance program, (b) allowed greater flexibility in approaches to problem identification assessment and resolution, and (c) emphasized the importance of focusing quality assurance activity on areas where demonstrable problem resolution is possible (20,21,26).

In addition, the Joint Commission launched the Agenda for Change in 1986 to focus HCOs and the general public on quality of patient care. The goals of the Agenda for Change are

1. To use valid and reliable clinical indicators as screening devices to identify potential problems in the organization, provision, or monitoring of care
2. To apply more relevant organizational standards and related indicator measures in evaluating the effectiveness of organizations' governance and management
3. To render accreditation decisions that reflect more accurately the adequacy of an organization's attention to providing high-quality care.

The quality assurance activities of the Agenda for Change must follow a monitoring and evaluation process. The monitoring and evaluation process has 10 basic steps:

1. Assign responsibility
2. Delineate the scope of care (inventory clinical activities)
3. Identify important aspects of care (e.g., high risk, high volume, and/or problem prone)
4. Identify indicators (e.g., variables related to outcome of care)
5. Establish thresholds for evaluation
6. Collect and organize data
7. Evaluate care
8. Take action to solve problems
9. Assess the actions and document improvement
10. Communicate relevant information to the organization-wide quality assurance program

The 10-step process for monitoring and evaluation can also be used for such hospital-wide programs as infection control, utilization review, risk management, and safety management. In many hospitals, such programs are under the governance of the quality assurance administrative team, usually directed by a physician.

Conclusion and Summary

The quality management issues of quality assurance and CQI, including the JCAHO Agenda for Change, have had and will have a profound impact on the bottom line of the practice of laboratory medicine. When these are combined with the key variables of business performance, as described by Garvin (9):

- price
- advertising
- market share
- cost
- productivity
- profitability

The laboratory of the future will be assured of success, both in quality of care and in management. However, unless quality management is practiced assiduously, such entities as quality assurance, CQI, and even the Agenda for Change certainly will lose their impact.

An understanding of the principles of managing laboratory costs is vital to the business success of the laboratory. The section that follows provides the tools needed to manage successfully a cost-effective laboratory. A balance between the principles described in the preceding section and those contained in the next section is the secret to management of a productive, efficient, cost-effective, and profitable (contributory) laboratory (Fig 1.12).

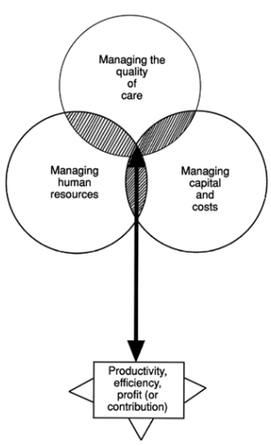


FIGURE 1.12. Key elements for success in laboratory management.

MANAGING LABORATORY COSTS

Part of "1 - Basic Laboratory Management"

Strategies, Problems, and Principal Areas of Laboratory Expenditure

Strategies

Establishing strategies for cost-effective laboratory management begins with an assessment of the laboratory's *external* and *internal* fiscal environments.

In the external environment, the laboratory manager must understand both the organization's and the laboratory's strategic plan. Its market share and product line, its financial goals, its role in supporting the local medical community, its administrative responsibility to its source of operating capital (i.e., a hospital, corporation, multihospital group, HMO, or federal,

state, or municipal government), its capital asset structure (where applicable), and the limitations on current budget, known as operating funds. In the internal environment, the laboratory manager must assess the age and condition of equipment, staffing adequacy and balance, salary structure, and types of reimbursement received and their multiple sources, and must analyze the volume and complexity of the laboratory's input (workload), especially tests created by the explosion of molecular biology and genome-related technologies. If the laboratory is engaged in teaching resident physicians or laboratory professionals or is performing research and development, the additional impact of these resource-consuming programs must also be assessed.

The Fiscal Management Team

Accomplishing this comprehensive fiscal assessment is impossible without the cooperation and interaction of a management team in the laboratory that works closely with hospital, corporate, or institutional management. Although the overall responsibility for cost-effective management rests with the laboratory director, the task ultimately requires an integrated effort between top management, middle management/supervisory personnel, and laboratory technical staff (Fig. 1.13).

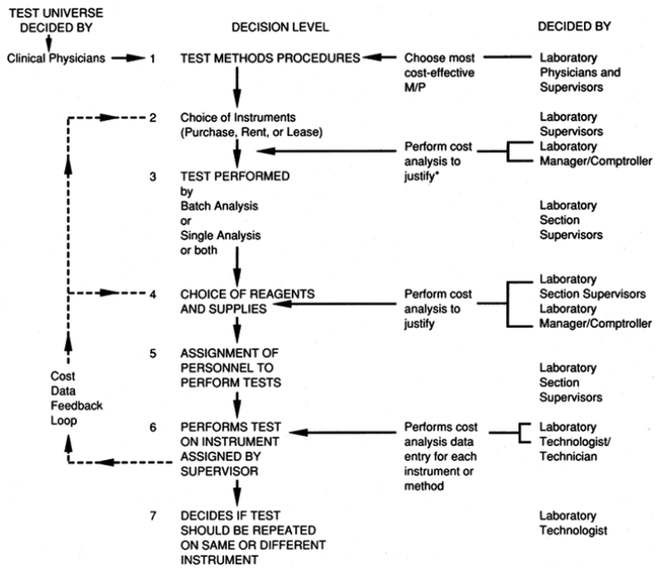


FIGURE 1.13. Laboratory financial decision flow chart. *Cost analysis may indicate that it is more cost effective to send out tests to other laboratories, rather than purchase, rent, or lease an instrument. (From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:5, with permission.)

The Pathologist's Role in Laboratory Fiscal Management

The pathologist's role is critical to success in fiscal management and rests on recognizing medical care priorities, good planning, excellent control, and recognition of the major differences in managing the costs of anatomic and clinical pathology. The value of a good laboratory manager is directly related to his or her ability to foresee future technological and workload trends, anticipate and deal with sudden environmental changes, and make critical decisions for managing limited resources, based on these abilities.

Problems

Principal Fiscal Management Problems

The major fiscal problems facing laboratory managers include (a) managing enormous increases in workload, mandated regulations, compliance and quality management requirements, (b) managing “hidden” and uncontrolled costs, and (c) finding and installing adequate fiscal and management information systems to capture the majority of data input and relieve the technical and administrative staff of the onerous, inefficient task of manual data entry. A subset of this problem (as of this writing) is that there is still little or no compatibility between different laboratory computer vendor's information systems, let alone fiscal and management systems. This makes it difficult if not impossible for a laboratory manager to compare the laboratory's fiscal and utilization data with those of laboratories of similar complexity and workload in the same community. However, several external programs to assist managers in benchmarking operating characteristics, financial, productivity, utilization, efficiency, billable tests, and other management indicators can be purchased (Table 1.1).

Lack of Cost Control Measures

Additional measures that can be used to control costs include (a) controlling automation, primarily by avoiding additional or excessive labor costs associated with the acquisition and operation of unnecessary and excessive automated equipment; (b) avoiding the additional costs for reagents and supplies associated with excessive automation (and workload/quality control demands); (c) providing more efficiency in chemistry and hematology (70% to 80% of total laboratory workload) (27); and finally (d) controlling the total laboratory budget more closely (especially labor costs) by consolidating groups of tests and convincing clinicians to reduce unnecessary routine, daily testing practices.

Principal Areas of Laboratory Expenditure

Direct Costs

Labor costs. Of all costs in laboratories, the most difficult to control are labor costs. Labor costs account for the majority of expenditures, regardless of the complexity and size of the laboratory (70% to 85%, depending on seniority, benefits, and number and size of the staff) (27). A categorization of labor costs appears in Table 1.2.

TABLE 1.2. LABOR SUBACCOUNT CATEGORIES

Operating (or standard cost) labor	
	Labor, quality control at standard cost
	Labor, variance from standard cost
	Supervisors' salaries
	Physician remuneration
	Doctoral level salaries
	Technologist salaries
	Technician salaries
	Secretarial/clerical salaries
	Aide and other operating salaries
Nonoperating labor	
	Executive personnel remuneration
	Maintenance personnel salaries
	Supervisors' salaries
	Driver salaries
	Pilot salaries
	Data processing personnel salaries
	Other nonoperating personnel salaries

Adapted from Gaither JF, Resinger HE. *Cost accounting in the laboratory*. Mundelein, IL: American Pathology Foundation, 1981:V-20.

Reagent Costs. The second major category of laboratory expenditure is for reagents and supplies (consumables). These usually account for 10% to 15% of total laboratory expenditures, depending on how carefully discounts are obtained and how efficiently reagents are used. In fact, the inefficient and excessive use of quality control and standards, especially in chemistry, can accumulate

to 30% to 35% of chemistry costs if not carefully monitored (Table 1.3).

TABLE 1.3. MAJOR SUPPLY REAGENT AND BLOOD SUBACCOUNTS

Materials and supplies, total
Materials, laboratory, total (or at standard cost)
Supplies, quality control, at standard cost
Variance from standard cost, laboratory materials, and quality control supplies
Supplies, other laboratory
Materials, other
Supplies, office
Supplies, teaching
Supplies, data processing
Supplies, other

Adapted from Gaither JF, Resinger HE. *Cost accounting in the laboratory*. Mundelein, IL: American Pathology Foundation, 1981:V-20.

Costs for Blood and Components. Costs for blood and components are far more difficult to control based on the shortage of raw products from donors, the cyclical and unpredictable nature of patients' illnesses, and the need for emergency or operative treatment. One of the most difficult areas of cost to control is the blood bank's transfusion service because the hospital's case mix and clinical factors will have a profound influence on the type and volume of blood and components purchased and utilized. Furthermore, laboratories that do not collect and process their own blood are dependent on outside sources and have limited ability to control their costs and/or optimally set their annual charges for blood and components.

A major factor that has escalated the cost of collecting and processing blood is the cost of testing for transmissible disease, new improvements including leukocyte reduction, nucleic acid amplification testing, and viral inactivation make the transfusion service a costly operation (29). In transfusion-related services, one must not forget that the cost of the testing reagents and labor may not be the major source of cost. The major cost is dependent on the charges for blood products that suppliers of blood impose on hospital laboratories to recover their labor costs to perform mandated safety tests.

Fee-Basis Testing Costs. Fee-basis tests are send-out tests that cannot be performed in-house due to lack of equipment, technical expertise, or personnel or for other reasons. This category of expenditure may be a major source of additional expense for laboratories that do not have a large, highly trained staff with special testing expertise and/or state-of-the-art equipment. Also, fee-basis testing may be necessary if the type and complexity of tests ordered are of a complex or esoteric nature, i.e., those ordered in teaching or tertiary care settings.

Large increases in fee-basis testing expense can be avoided by controlling the test-ordering process at the clinical level and by choosing the lowest cost for the desired tests through competitive bidding and comparison of prices. It is also important for laboratory management to advise clinicians of test costs and turnaround time using published lists of tests with this information.

Capital Equipment Acquisition Costs. There are several ways to acquire equipment, each of which has its advantages and disadvantages (28). The principal methods are purchase, lease, rental, and special rental contracts (e.g., cost per test, reagent rental, and other customized contracts). Without consideration of the cash flow requirements of each method, appropriate planning is necessary to ensure that the laboratory's budget can accommodate the acquisition because additional equipment may escalate operational costs.

The criterion for choosing among the different alternatives is the minimization of the cost of acquiring the equipment and the yearly net cost outflow associated with each alternative. These areas are usually best handled by an accountant who understands laboratory operations and who will perform the required analysis to compare purchase with a time-sale contract, financial lease arrangement, rental, or other contractual method.

However, the current difficulties with availability of money for investment in capital equipment make it imperative that hospital laboratory managers present a complete and convincing presentation to hospital or corporate management when requesting *purchase* of new or replacement equipment. The process itself is segmented into five distinct categories:

1. Development of depreciable costs
2. Estimation of appropriate depreciation life and timing
3. Isolation of related cash flow variables
4. Forecast of production volume and revenue
5. Measurement and comparison with hospital criteria

The laboratory manager should realize that other equally important factors must be considered in the capital acquisition process, namely, turnaround time, rapid response, precision, accuracy, reproducibility, availability of service, and safety (29).

A combination of the two processes discussed above should result in a smooth final acquisition determination. A detailed description of the elements of this process is beyond the scope of this chapter but can be reviewed in an article by Oszustowicz (30) and in textbooks by Travers (27,31), which provide a format for a marketing opportunity package for capital equipment (see Methods for Acquiring Equipment section).

Indirect Costs

Indirect costs are the costs of specimen collection, processing, testing and reporting that cannot be directly traced to test production. A partial list of basic items used in estimating indirect costs for tests is included in Fig. 1.14. A detailed explanation of

indirect costs using formulas and examples can be found in *Cost Accounting in the Clinical Laboratory*, published by the National Committee for Clinical Laboratory Standards (32), and textbooks by Travers (27,31).

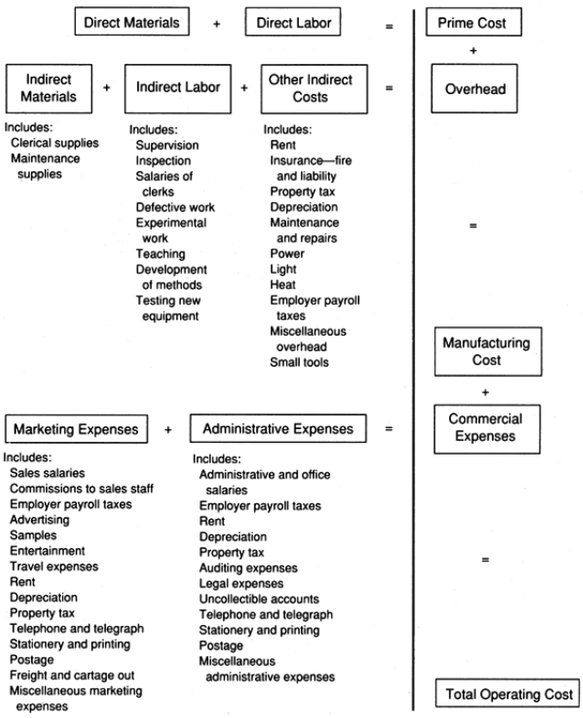


FIGURE 1.14. Analysis of total operating cost. (Adapted from Matz A, Usry MF. *Cost accounting: planning and control*, 7th ed. Cincinnati, OH: Southwestern Publishing, 1980:45.)

Indirect costs are assigned to total cost based on a multitude of factors that are part of hospital or corporate operations. They are added to the cost of tests as a percentage rather than as actual amounts for each item.

Total indirect costs to operate an organization or department must be recovered by reducing operating expenses, by charging for research, epidemiologic, or environmental tests, or through the pricing of tests and services. The total cost of performing a test or procedure is the sum of the direct cost plus a proportionate share of indirect costs.

The indirect cost percentage will vary in each laboratory based on the total number of factors that the organizations and the laboratory uses to calculate indirect costs. These are considered as highly significant expenditures because in some private sector settings, indirect costs may be as high as 148% of total cost (27,31).

The Laboratory Fiscal System

Every organization—regardless of its size or complexity, profit or nonprofit status—must maintain a system for financial data collection and coding for recording in journals and ledgers. The prerequisite for accomplishing these tasks is a properly established account structure. This fiscal system is known as the chart of accounts.

The Chart of Accounts

The organization's chart of accounts is the fundamental means for budgetary and control accounting. It provides control accounts for the recognized elements of cost, and it segregates and details all expenses not included in prime cost (the total cost of direct materials and labor in test production). The chart of accounts is the foundation of the financial management system for the laboratory. It includes information on the historic financial database and current financial information about the laboratory and is a future financial planning aid. It also allows consistency of financial data collection for planning and control. The comprehensive chart of accounts (Fig. 1.15) provides management data breakdown in categories needed by management, namely, assets, liabilities/equity, revenues, and costs (expenses).

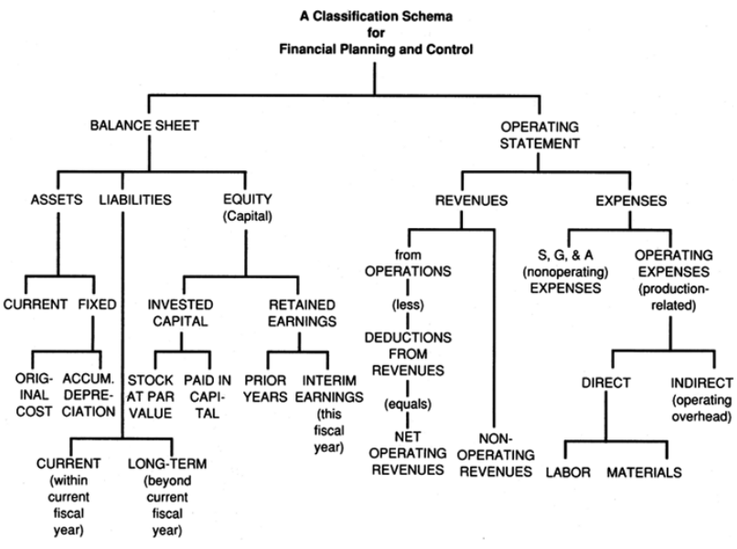


FIGURE 1.15. The chart of accounts for laboratories. (From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:36, with permission.)

The chart of accounts has two major sections: (a) the balance sheet and (b) the operating statement (income statement). The balance sheet is divided into three major sections: assets, liabilities, and equity. The operating statement is divided into revenues and expenses. Because laboratories vary in their activities and complexity, the chart of accounts will also vary, according to the information that the managers believe they need to accomplish the task of sound fiscal management. Not even a large, independent laboratory will use more than a small fraction of the

available accounts, but the accounts they need are available, because the fine structure of the chart of accounts meets virtually every conceivable need for detail (27).

The Balance Sheet

The purpose of the balance sheet is to report the financial position of the laboratory at a particular time. Financial position refers to the amount of resources (i.e., assets) and liabilities of the laboratory on a specific date. The balance sheet provides the laboratory manager with a recorded, systematic statement of each transaction involving assets, liabilities, or equity. The accounting process provides the raw data for each subaccount, whether it is handwritten, mechanized, or computerized. Each transaction is recorded in one of the major sections of the balance sheet and is ultimately used to construct the laboratory's financial statements. The basic formula for the fundamental accounting model is:

Assets = Liabilities + Equity

This formula leads to an understanding of the basic and broad perspective of accounting as it relates to overall fiscal management. The basic objective of cost-effective laboratory management is to balance this equation. In situations in which improved profit is desired, the need to minimize and reduce liabilities is even greater, while maximizing assets and equity. In not-for-profit situations, in which assets or profit are not desired, minimizing liabilities becomes even more important to preserve operating capital (27,32). Major categories of liabilities are listed in Fig. 1.15.

The Operating (Income) Statement

This section of the chart of accounts includes a format for determining the laboratory's revenues, expenses, and profit (or losses).*a* Revenues are obtained from operations or from nonoperating sources. Operating revenues are characterized by type of payment agreement and type of guarantor. Nonoperating revenues are obtained from categories such as technical school operations, grants, investments, rent, and miscellaneous revenue-generating activities other than laboratory operations.

In a hospital, the laboratory is one of many cost centers; however, it is also a revenue center, i.e., it charges and/or is compensated for its services. It is necessary, then, for the laboratory and other hospital revenue centers to produce revenues in excess of expenses to offset losses incurred by other nonrevenue-producing centers. In laboratories, the operating statement is used to report the profit performance of the laboratory for a specific time period (e.g., year, quarter, month). Profit (or net income) represents the difference between revenues and expenses, i.e., the results of operations. Even not-for-profit laboratories need to make a profit (e.g., efficiency or savings) to survive, expand, and make a “contribution” to the nonrevenue-producing departments.

Therefore, the income statement reflects one of the fundamental principles of accounting, i.e., Revenue - Expenses = Net income or (Net operating revenues + Nonoperating revenues - Expenses) = Net income (Profit or Contribution).

Revenues include sales of laboratory tests, whereas expenses include the general categories of costs of test specimen collection, processing, performance and reporting, marketing and selling expenses, general and administrative expenses, financial expenses, and income tax expenses.*b* More detailed information and a listing of income statement categories can be found in Gaither and Resinger (33) and Travers (27,31).

DETERMINING PROFIT OR LOSS

Part of "1 - Basic Laboratory Management"

One of the techniques used to establish profit or loss is known as cost-volume-profit analysis, also known as marginal analysis in general industry. It is used to provide relevant information for selecting product lines, pricing individual products, and developing market strategies. It is a valuable method for analyzing profitability in many firms. Cost-volume-profit analysis permits the study of the relationships between prices, variable costs, fixed costs, and workload volume to aid in management decision making. In laboratory use, the key relationship is the difference between test price and variable costs per test, otherwise known as the *contribution margin*.

Distinction between Contribution and Profit Margin

There is an important distinction between profit and contribution. As noted previously, profit (or net income) represents the difference between revenues and expenses. However, contribution is computed by deducting the costs for which the laboratory is responsible (both direct and indirect) from the revenues gained for the tests generated by the laboratory. Contribution can also be applied to laboratory sections, e.g., contribution for hematology section is calculated by taking all fixed and variable costs that are the responsibility of the hematology section and deducting them from the revenue generated in the hematology section for the period in which the cost was incurred.

Impact of Contribution on the Rate-Setting Process

The impact of contribution on the rate-setting process by hospital or corporate management is significant because it affords hospital management the opportunity to calculate the approximate level of patient service revenue required to cover budgeted economic costs. The portion of the rate charge that directly affects the operating margin is the contribution factor.

There are both internal and external reasons for placing minimum and maximum limitations on departmental rate increases

and decreases. Among these factors are competition, governing board policy, limitations imposed by rate-setting authorities, and contractual arrangements. Each of these factors is reviewed as applicable, and upper and lower limits are established within which rates may be charged. For example, suppose hospital administration has determined that the room rate is well below that of neighboring hospitals and a room rate increase of 5% to 25% is desirable. However, emergency department charges are slightly higher than acceptable. A survey has shown that laboratory rates could be increased by as much as 20%, but the hospital's governing board has indicated that a 10% reduction would be desirable. Finally, the radiologists' contract provides that there be no reduction in departmental rates while their contract is in effect. Based on this information, the hospital establishes the following rate change limitations:

	Lower Limit	Upper Limit
Routine care	+5%	+25%
Emergency department	-5%	0%
Laboratory	-10%	+20%
Radiology	0%	+10%

By applying rate increases to the departments with the highest contribution factors and rate decreases to those with the lowest contribution factors until the target profit contribution factor is reached, the aggregate rate change will be minimized (33).

Management Responsibilities and Contribution Margin

During times of unrestrained resource available, profit-maximizing decision makers will elect to produce all the products they can as long as their marginal revenue (MR) (the price or charge) exceeds marginal cost. However, in times of limited resource availability, profit-maximizing decision makers should (and will only) select those tests for market expansion that will have the greatest effect on producing a high contribution margin.

The next major management topic concerns the manager's responsibility to plan the future use of operating funds, known as budgeting.

BUDGETING

Part of "1 - Basic Laboratory Management"

Principles

A budget is generally defined as a comprehensive plan or guide for some future period of time, usually, by convention, a year, a quarter, or a month. Budget plans can be expressed in a variety of ways, e.g., in man-hours, dollars, outputs. A budget, however, is primarily a financial plan and is therefore expressed in dollars. Thus, the term budget might be defined more explicitly as a comprehensive financial plan, based on anticipated outputs and predetermined hospital goals and policies for future operations, that is expressed in dollars of expense and corresponding dollars of revenue.

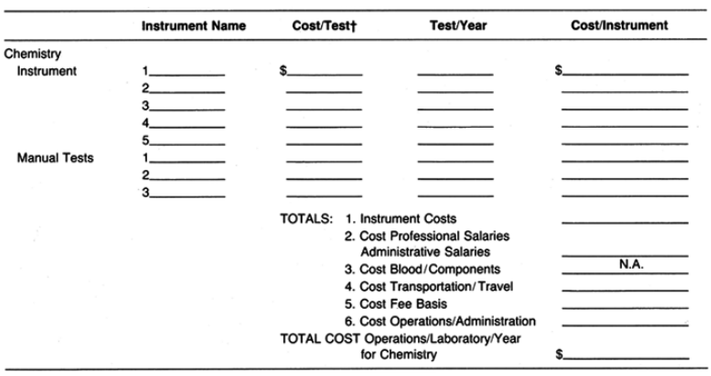
For purposes of understanding and discussion, the budget translation process can be viewed as consisting of two primary steps: (a) converting the hospital's or corporation's strategy into an operating plan for the laboratory and (b) expressing the operating plan as a budget for the laboratory. The first step involves defining the specific workloads, projects, and other operating activities that must be undertaken to transform the corporate strategy into reality. The second step focuses on expressing the operating plan in financial terms, for example, dollars of expense, revenues, capital purchases. Therefore, there are four basic types of budgets: statistical (e.g., workload), expense, revenue, and capital budgets.

Before constructing the budget, an additional element is necessary: a budget preparation procedure. This includes the technical, mechanical aspect as well as the management of the entire process. The technical part focuses on (34)

- projecting workload (volume)
- converting volume estimates into resource requirements and revenue estimates
- converting resource requirements into direct cost estimates
- calculating indirect costs
- adjusting revenues and total costs to obtain the necessary equality

Sectional Budget Analysis

Once the basic concepts of budgeting are understood, the actual construction of budgets for laboratory-specific elements can begin. A general sectional analysis format (Fig. 1.16) is adaptable for all laboratory sections and includes entries for all major areas of laboratory instrumentation and manual methods. Direct costs are established first, using a microcost analysis method (see Appendix A), then these data are entered into the appropriate section for instrument based or manual methods. The cost is then multiplied by the raw test count (the workload) for each category for the period in which the budget is constructed. The final product for each instrument or method is added to produce the total direct cost for test production. Added to this figure are other major sectional costs, namely, professional salaries and fees, blood and component costs, transportation, fee-basis (send-out) costs, costs for operations, administration, travel, depending on the laboratory's organization structure and fiscal control points. This basic worksheet combines the principles of (a) where resources are consumed, (b) how rapidly they are consumed, and (c) assigning a site to link the cause and the site to where the cost was generated.



	Instrument Name	Cost/Test†	Test/Year	Cost/Instrument
Chemistry				
Instrument	1_____	$_____	_____	$_____
	2_____	_____	_____	_____
	3_____	_____	_____	_____
	4_____	_____	_____	_____
	5_____	_____	_____	_____
Manual Tests	1_____	_____	_____	_____
	2_____	_____	_____	_____
	3_____	_____	_____	_____
		TOTALS: 1. Instrument Costs		_____
		2. Cost Professional Salaries Administrative Salaries		_____
		3. Cost Blood/Components		N.A.
		4. Cost Transportation/Travel		_____
		5. Cost Fee Basis		_____
		6. Cost Operations/Administration		_____
		TOTAL COST Operations/Laboratory/Year for Chemistry		$_____

FIGURE 1.16. Worksheet for sectional analysis of chemistry costs based on equipment configuration manual procedures. *This format can be utilized for all laboratory sections. Chemistry is used here as an example. For instance, if used for the blood bank section, item 3 in TOTALS would be applicable. Cost/test includes direct costs (labor [technical/clerical/phlebotomy], reagents and supplies) and indirect costs (amortization, depreciation, maintenance, site preparation, and all other costs). (From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:130, with permission.)

These are the principles of activity-based costing, a newer method than product costing for only one test/service.

Departmental Budget Analysis and Allocating Costs across Sections

Although direct costs noted in the previous section comprise the major costs per section, it is important to allocate costs to a budget format that also include the costs for the time of general employees who are not specifically assigned to a laboratory section but whose work helps to produce or interpret tests (e.g., secretaries, couriers, computer personnel, maintenance personnel). These are known as indirect costs. Some laboratories use formulas

for allocation; however, there is no uniform set of formulas that will apply to every laboratory. An experienced manager should be able to create accurate formulas for each production area and the department. The only requirement is that the formulas must realistically reflect the section's or department's use of that function. Examples of allocation within a laboratory section budget are given in Fig. 1.17 and within a departmental budget in Fig. 1.18. Figure 1.19 further illustrates budgeted expenses for a total laboratory, by breaking major cost categories into their cost accounting subcomponents.

		A (Demand)	B (Astra)	C (ACA)	G (ABL Blood Gas)	H (Centrifichem)	
Example 1:	General Function That Varies per Instrument						Total Chemistry Tests
	Total Billable* Procedures	39,232	65,068	23,922	5,052	6,890	140,164
	Percent of Total Chemistry Tests	27.9	46.4	17.1	3.6	4.9	
		x	x	x	x	x	
	Phlebotomy Salaries† for Chemistry	$100,128	$100,128	$100,128	$100,128	$100,128	
		=	=	=	=	=	
	Cost of Phlebotomy per Instrument	$ 27,936	$ 46,459	$ 17,122	$ 3,605	$ 4,906	
Example 2:	General Function That Applies Equally to All Instruments/Sections						Total Professional Salaries for Chemistry¶
	Pathologist's Salaries (Professional Fees)‡	12,500	12,500	12,500	12,500	12,500	62,500

*Test result actually billed to patient. Example: a chemistry profile with 18 determinations counts as one billable procedure, not 18 tests.
†Determined by the following ratio:
1. $\frac{\text{Total Billable Procedures/Chem}}{\text{Total Billable Procedures/Lab}} = \frac{175,999}{394,000} = .447 = 44.7\%$ 2. 44.7% × Total Phlebotomy Salaries/Lab = .447 × $224,000 = $100,128
‡Pathologist's time in this lab is allocated as follows:
50% Histology/Surgical Pathology
10% Chemistry
10% Microbiology
10% Hematology
10% Immunology
10% Blood Bank
¶For Chemistry:
a. 10% × Total Prof. Fees/Lab = 10% × $625,000 = $62,500/Chem
b. 62,500/5 (instruments) = $12,500 Prof. Fees/Instrument

FIGURE 1.17. Allocating costs for general functions among chemistry workstations. (Adapted from Sharp JW. A cost accounting system targeted to DRGs. *Med Lab Observer* 1985;17:34–41.)

	Chemistry	Serology	Histology	Microbiology	Total Laboratory
Billable procedures	176,000	17,000	21,000	42,000	415,000
Technical FTEs	21	3	8	9	64
Revenue	$ 2,640,000	$ 230,000	$ 1,050,000	$ 756,000	$ 6,716,000
Expenses					
Salaries (technical)	$ 595,000	$ 75,000	$ 200,000	$ 225,000	$ 1,600,000
Supplies	600,000	75,000	40,000	155,000	1,035,000
Allocated expenses					
Phlebotomy[1]	$ 100,061	$ 9,655	$ 0	$ 23,870	$ 223,992
Path. secretaries[2]	7,500	1,875	60,000	1,075	74,200
Professional fees[2]	62,500	62,500	312,500	62,500	625,000
Lab receptionists[1]	67,005	6,472	0	15,990	150,000
Reference lab fees[3]	227,500	32,500	16,250	16,250	325,000
Cost of blood[3]	0	0	0	0	310,000
Equip. maintenance[3]	65,000	2,500	5,000	15,000	128,500
Depreciation[4]	75,063	3,146	11,438	21,161	169,995
Travel[5]	3,333	3,333	3,333	3,333	19,998
Other[3]	19,167	19,167	19,167	19,167	115,002
Direct expense	$ 1,752,129	$ 291,158	$ 667,688	$ 558,346	$ 4,776,687
Indirect expense[6]	$ 506,530	$ 48,926	$ 60,438	$ 120,877	$ 1,194,375
Total expense	$ 2,258,659	$ 340,084	$ 728,126	$ 679,223	$ 5,971,062
Net income(loss)	$ 381,341	($ 110,084)	$ 321,874	$ 76,777	$ 744,938

FIGURE 1.18. Example of a department laboratory budget. [1] Allocation based on section's percentage of all nonhistology billable procedures (394,000). [2] Based on estimated work done for each section. [3] Based on utilization by each section. [4] Depreciation over seven years: zero salvage value. [5] Total expense shared equally by all six sections. [6] Based on section's percentage of 415,000 procedures. (Adapted from Sharp JW. A cost accounting system targeted to DRGs. *Med Lab Observer* 1985;17:34–41)

Traceability	Cost Behavior: Variable		Fixed		Semifixed		Total
Direct	Salaries (part-time)	$4,000	Salaries (full-time)	$1,000	Salaries (other)	$10,000	
	Supplies	5,000					
		$9,000					$20,000
Indirect	Employee Benefits	$150	Depreciation	$160	Maintenance	$250	
	Housekeeping	100	Administration	500	Laundry	100	
			Housekeeping	100			
		$250		$760		$350	$1,360
	Total	$9,250		$1,760		10,350	$21,360

FIGURE 1.19. Laboratory cost example. (Modified from Cleverley WO. *Handbook of health care accounting and finance, volume I*. Rockville, MD. Aspen Publishers, 1982:146.)

COST ACCOUNTING

Part of "1 - Basic Laboratory Management"

Definitions

Accounting focuses on measuring and reporting the input and output of resources in an organization and the resources controlled by the organization. It provides an information base that allows decision makers to assess the potential financial implications and potentials of various alternatives being considered. Berman (35) and Horngren (36) point out that "the essential strength of the managerial cost accounting process is that it links 'promises' made during the budget process back to the responsibility center" (36).

The subject of accounting can be divided into two distinct major specialties: financial accounting or external financial reporting to shareholders, creditors, the IRS, and regulatory and other external bodies and managerial accounting, which deal with internal reporting. Within managerial accounting (also known as responsibility accounting), the subject of cost accounting defines, measures, reports, and analyzes the elements of costs associated with providing a unit of output—in the specific case of laboratories, a test or service, also known as a procedure in reimbursement coding manuals.

In the past, cost accounting referred to the methods for accumulating and assigning historical costs to units of products and departments, primarily for purposes of inventory valuation and income determination. Cost accounting is now indistinguishable from management accounting because it serves multiple purposes. Fundamentally, cost accounting now refers to the gathering and providing of information for decision needs (which range from large-scale needs such as management of recurring operations to the making of nonrecurring strategic management decisions) and for the formulation of major organizational policies. The smallest unit able to be costed is a product. The cost accounting method needed to find the costs of a laboratory's product(s), tests, and services is known as job-order costing.

Cost Categories Used to Determine Operational Expenses

Figure 1.19 lists the categories of costs that encompass both the traceability of costs (direct or indirect) and the behavior of costs (fixed or variable). Figure 1.19 also illustrates how laboratory costs fall into both traceable and behavior categories and breaks down total costs for each major laboratory expense into direct or indirect, fixed or variable costs. Direct costs are test-specific costs that can be directly traced to the production of tests and services. These include all costs for labor and supplies needed to produce a test. Indirect costs (common or joint costs) cannot be traced to a particular test; however, they contribute to the provision of an adequate work environment. Examples are section-specific and general supervisory salaries, general quality control, miscellaneous supplies, education, travel, and global costs such as administrative costs, computers, maintenance, and security (Fig. 1.14). Ultimately all indirect costs must be allocated to each laboratory section and to each test. Variable costs are costs related to changes in the volume of tests produced. Examples are reagents, supplies, and labor that can be controlled by laboratory supervisors by obtaining the best discounts, purchasing only when necessary, and using the instrument and method that uses the least costly labor and material. Fixed costs are costs that remain constant regardless of the volume of tests produced. They buy the capacity to do business—space, equipment, and administrative and other general personnel (e.g., preventive maintenance contracts, equipment depreciation, insurance policies, automobiles for courier service).

Cost Accounting for Laboratory Tests and Services

Accounting for the cost of a laboratory test or service considers the prime cost aspects of analysis: (a) labor costs and (b) material (reagent, supply) costs, and (c) indirect costs, also known as overhead. The interrelationship of these three areas is illustrated in Fig. 1.14, which also shows the business elements of test cost that must be added if the test is produced in a corporate setting, e.g., marketing and administrative expenses.

In 1989, Travers (27) published a detailed generic cost accounting method (Appendix A) that can be used in any section of the laboratory to derive a total cost for any test produced, using any procedure, instrument or test method. Known as the instrument cost accounting technique (ICAT), the method utilizes generally accepted business accounting techniques to construct actual total test costs. If charges or prices need to be established, ICAT provides the basis for identifying actual costs for labor, reagents, supplies, and instrument-related (and other indirect) costs incurred with the performance of the test. Using the total actual cost as a base amount, laboratory managers can realistically project test charges and prices by adding desired profit and contribution margin elements (see Setting Test Price and Charges section). Standard laboratory cost management texts and guidelines provide worksheets that can be used to develop test costs (27,31).

Cost Accounting for Disease Categories

In the future, laboratory medicine will become even more dependent on accurate and realistic total reimbursement to survive financially. Therefore, the value of the ICAT becomes more important because it can be used to calculate the actual laboratory costs for all tests and services performed for a patient's hospitalization by the patient's disease category. Once the cost per test for each type of test (e.g., sodium, potassium, glucose) is determined for the instrument or method used, the cost of a particular pattern of disease-related tests ordered can be established (Fig. 1.20). This technique will become more important as cost accounting software is developed that will rapidly associate workload and activities as tests are produced with the costs of each test or service.

			Cost Per Test					
			Direct			Indirect		
Test	Value of Test(s)	Equipment	Reagents/ Supplies	Instrument	Labor GS-9 GS-7	Indirect	Frequency	Total Cost
BUN	Highly Valuable	Demand	$ .13					
Potassium	Highly Valuable	"	.22					
CO_2	Highly Valuable	"	.21					
Chloride	Highly Valuable	"	.22					
Creatinine	Highly Valuable	"	.09	$ .04	$ 7.01	$ .79	1	$ 9.45
Sodium	Highly Valuable	"	.13					
Glucose	Highly Valuable	"	.13					
24-Hour Creatinine	Highly Valuable	"	.09					
Protein	Highly Valuable	"	.15					
Cholesterol	Moderate Value	"	.24					
Urine Microscopic	Highly Valuable	Manual	.15	.03	1.55	.19	1	1.92
Hematocrit	Moderate Value	Coulter S + IV	.27	.36	2.09	.30	1	3.02
Hemoglobin	Moderate Value							
TOTAL COST/LDRG/NEW ADMISSION			$2.03	$ .43	$10.65	$1.28		$14.39

FIGURE 1.20. Example of a cost analysis worksheet. Major diagnostic related group: 294.MDC 10M. Diabetes ≥ 36. Principle diagnosis: 250.40 adult-onset type (diabetes mellitus with renal manifestations/noninsulin dependent). Patient status: new admission, stable. Frequency of test(s) ordered: order once on admission only. LDRG, laboratory diagnostic related group; MDC. (From Travers EM. *Laboratory diagnostic related groups*. Department of Academic Affairs, Continuing Education Workshop, Washington, DC, May 1984, with permission.)

Activity-Based Costing

Activity-based costing links fiscal resources consumed in the cause of the activity (workload generated) and the site where the cost was generated. Activity-based costs consider the direct cost of the test/service produced, multiplied by the volume of the test/service produced. Indirect costs are separately allocated by

tracing costs associated with the production process and adding them to the total direct cost of production. Known as allocation, the process is described in Fig. 1.17.

Contrasted with actual (job order, product) costing, activity-based costing is less accurate. However, activity-based costing is more helpful to managers responsible for budget formulation and make-or-buy decisions.

The Role of Cost Accounting in Variance Analysis

Variance analysis is the comparison of the deviation of actual costs from budgeted or expected standard costs. Cost accounting provides the tools to develop actual costs, and over time, the averaging of actual costs provides the expected or standard cost figure. In times of shrinking economic resources, variance analysis provides a tool to monitor the financial performance of an operating department. The significance of variance analysis is its ability to break the total difference between standard and actual costs into elements such that the causes of the difference are revealed. This allows the manager to decide what action should be taken, by constructing an explanation worksheet to use as a basis for mutual resolution of variance problems (Fig. 1.21).

Expense Item	Amount	Explanation of Variance	Current Corrective Action
Direct Materials Material Usage	(380)	Poor reagent handling-outdating.	Variance should be eliminated with new plan for handling.
Direct Labor Labor Efficiency	(8855)	Lost three experienced workers. New replacements need training on equipment and upset routine of entire department.	New replacements not envisioned in setting standards last year. Expect training to be completed soon. Will be included next year. No standard revision believed necessary now.
Labor Rate	(150)	New union contract rate for workers at higher rates than used in planning standards.	No standard revision believed necessary. Will set up separate action for rate variance.
Overhead Efficiency	(4679)	Caused by 1519 excess hours of labor.	See Labor Efficiency above.
Overhead Spending Supervisor	300	Budgeted as fixed expense with provision for salary increase at midyear.	Not required, since variance will disappear by year end.
Indirect Labor	(752)	Cost of working every Saturday during quarter for extra volume.	Short-term contract completed.
Employee Benefits	200	Budgeted as fixed expense with provision for rate increase at midyear.	Not required since variance should disappear by year end.
Supplies	50	Actual supplies used less than anticipated.	Not required. Year-to-date excess use has been reduced.
Maintenance & Repair	(884)	Repair of equipment causing down-time.	Temporary situation, now corrected.
Power & Light	(30)	Additional power requirements based on Saturday work.	Not required.
Taxes & Insurance	(20)	Not controllable.	Not required.

FIGURE 1.21. Example of a variance explanation worksheet. (Adapted from Gaither J, Resinger H. *Cost accounting in the laboratory*. Half Day, IL: American Pathology Foundation, 1981:10–14.)

For example, if the department chairperson claims that the purchase of a certain analyzer will result in labor savings, but no labor savings are experienced, then this shortfall in performance needs to be either (a) explained by exogenous circumstances beyond the manager's control or (b) utilized annually to discount the judgment and the budget of the manager in question (36). More detailed explanations of variance analysis can be found in standard laboratory cost management references (27,32).

Setting Test Prices and Charges: Make-or-Buy Decisions

One of the ultimate goals of a for-profit institution is to provide a profit or contribution for the organization's shareholders through the production of a product. The preceding section developed a rationale and methods for deciding whether a test (product) is profitable to perform, using cost accounting templates and variance analysis. Both the complex, active tertiary care laboratory and the less active laboratory must manage critically short resources by establishing actual cost and determining whether their prices or charges are adequate to cover costs and provide a profit (or contribution) margin to satisfy hospital or corporate management. Similarly, decisions must be made to determine whether certain tests should be discontinued in-house and sent to other laboratories on a fee basis. The latter analysis is known as make-or-buy analysis.

Pricing Strategies

There are several ways to arrive at a price or charge for a test, which range in complexity from using the price charged by neighboring hospitals or commercial laboratories to elaborate computerized systems. The most widely used method is comparison pricing, using rates from competing HCOs in the community. Price or rate setting is not often considered until hospital or corporate management asks for additional revenue from the laboratory. Frequently, increasing the price or charge still may not generate sufficient revenue; therefore, hospital or corporate fiscal managers may take the initiative to raise prices or charges on certain high-volume tests higher than the competitor's price or charge, regardless of how inaccurately the test prices

or charges reflect the laboratory's actual cost. Very often, the laboratory manager and director are never advised of this action. All too often, there is so little communication by corporate decision makers that laboratory managers do not ever know what is charged for their tests and services.

The comparison pricing approach is not recommended but is widely used. However, the only truly accurate comparison pricing method to set prices or charges is to measure actual direct and indirect costs incurred in test performance, using the microcosting approach discussed in the Cost Accounting section and in Appendix A. However, pricing tests does not simply consist of the mechanical task of adding up costs and tacking on a percentage that represents the desired profit margin. Price setting is one of the most difficult fields of business because the manager's intuition, ability to understand the marketplace, reimbursement regulations, legal and government regulations (such as the Robinson-Patman Act), and awareness of the competition's prices (in the case of commercial laboratories) are all parts of the process.

It is interesting to note that many laboratories do not have a problem with pricing tests. This is especially true for commercial laboratories. If a market price exists, customers will usually not pay more than this price, and there is a little reason for the test to be sold at a lower price. This is especially true for small laboratories that must compete in an industry in which one or a few large companies exercise price leadership. In this case, the laboratory makes no pricing calculations; it simply charges the market price.

There are situations, however, in which ethics must prevail. In some laboratories that offer the market price or lower than the market price, there is always the chance that management may decide to cut corners on the quality of service to obtain more business. Examples of “corner cutting” include failure to spend adequate time screening cytology slides (laboratory costs are the highest costs for any test) or similar activities that might reduce the quality of a test result. Making the decision to charge more for a test is difficult, especially if the costs for labor are higher, but the ethical laboratory manager should be prepared to explain to customers why the added cost will make a difference in the quality of patient care. This subject is not to be taken lightly. Severe federal penalties are assessed by the Department of Justice auditors when instances of price setting to destroy the laboratory's competitors are discovered and corroborated.

The Economic Basis for Setting Prices or Charges

The description of pricing in standard economics textbooks is stated in the law of supply and demand, i.e., the demand for a test is determined by how many customers want to buy it. It is only reasonable to assume that the lower the price, the more customers will be attracted to use that laboratory's tests and services. The supply part of the law refers to the test's cost. Similarly, the supply curve reflects the fact that unit costs decrease as test volume increases. This is because the fixed costs of a laboratory are distributed over more tests as test volume increases (examples of fixed costs are given in Fig. 1.19.) Figure 1.22 demonstrates the general rule that the most commonly used way to arrive at the best selling price is to estimate revenue and costs at several possible

selling prices and select the price at which the difference between revenue and costs is the largest (37).

Unit Selling Price	Estimated Quantity Sold	Total Revenue	Fixed Cost	Variable Cost (at $100 per unit)	Total Cost	Profit
$250	500	$125,000	$50,000	$ 50,000	$100,000	$25,000
200	1,000	200,000	50,000	100,000	150,000	50,000*
150	1,500	225,000	50,000	150,000	200,000	25,000
125 *	2,000	250,000	50,000	200,000	250,000	0

FIGURE 1.22. Pricing analysis example. *Preferred alternative. (From Anthony RN, Welsch GA. *Fundamentals of management accounting*. Homewood, IL: Richard D. Irwin, 1974:134–140, with permission.)

The Selling Price

The selling price (or charge or rate) for a test should be high enough to (a) recover direct costs, (b) recover a fair share of all applicable indirect costs, and (c) yield satisfactory profit (for the for-profit corporation) or provide a contribution back to the hospital. Stated another way, profit is for the for-profit organization and efficiency is provided for the not-for-profit organization. This must be understood as a statement of general tendency, rather than a prescription for setting the selling price for each test. There are some situations, however, in which the process works in reverse, i.e., the selling price that must be charged to meet competition is taken as a given; the problem then is to determine how much cost the laboratory can afford to incur if it is to earn a satisfactory profit at the given price. The laboratory manager must then design equipment configurations and methods to “fit” the desired price. If the test cannot be produced at the desired price, then it may have to be discontinued or the calculations should be redone, taking the next higher price point as a given (37).

The Billable Test

Billable tests are defined as those that, when charged to a patient, physician, or another entity, will generate revenues for the laboratory. Repeats performed to achieve a billable test must not be counted, provided that the original test was counted. Typically, billable test records can be obtained through many sources; the source chosen will vary based on the accuracy of the data in a given institution. Potential data sources include (a) actual manual test counts, (b) instrument meter readings, (c) computerized workload summaries, (d) laboratory accounting system summaries, health insurance information, insurance registers, and (e) hospital billing system summaries derived from encounter forms.

Whichever method is used, an audit should be performed before or during the data collection to ensure the accuracy of the information being compiled. Since 1998, severe penalties have been imposed by the federal government for false claims for Medicare reimbursement, even though fraud could not be proved.

Methods for Computation of Prices or Charges

Price-setting procedures are difficult and there are many methods for price setting, which vary in complexity. The manager should always consult a qualified attorney who is familiar with laboratory regulations and billing requirements before setting prices for tests. The techniques noted below, which are beyond the scope of this chapter, are discussed in other textbooks (27,33,35,37). The method recommended in this chapter is full cost pricing because it includes all resources expended for a test and is traceable back to the actual base cost of production of a specific test or service, in the event of an audit. Other methods include

1. *Profit maximization*, which relates total revenue to total cost; the objective of this method is to obtain a price that contributes the largest amount to profit.
2. Pricing based on *return of capital employed* is based on a percentage markup on cost.
3. *Conversion cost* pricing attempts to direct the manager's attention to the amount of labor and indirect costs required to produce tests.
4. The *contribution margin* approach to pricing indicates a laboratory's contribution to the recovery of fixed costs and to profit; this method sets the price below full cost (see method 7) but above the variable cost, where the difference between price and variable cost per test is contribution margin.
5. *Standard cost* pricing represents the cost that should be attained in an efficiently operated laboratory at normal capacity.
6. *Direct cost* pricing sets prices at a certain percentage above the direct cost incurred in producing a test; it is used by those who believe that the results of using indirect costs are not sufficiently valid to be useful; they prefer to base pricing decisions on direct costs because these costs can be measured with a high degree of accuracy, using microcosting methods.
7. *Full cost* pricing (total cost method) includes all the resources expended for a test; it is nearly identical to the ICAT discussed above. Appendix A provides worksheets for this technique.
8. *Relative value* pricing assigns a weighting factor to each test, based on the labor effort required to perform the test or procedure; it is lengthy and time-consuming and requires that detailed, up-to-date information is included for all techniques, instruments, and methods; its weakness is the use of average values for each procedure because it is known that some laboratories produce the same test faster and more efficiently than others.

Make-or-Buy Decisions

Laboratory managers have the responsibility to decide where it is most expedient to perform tests—in-house or in another laboratory.

This problem arises particularly in connection with the use of idle equipment, idle space, or idle labor or with not enough equipment, space, or labor to perform tests. Faced with a make-or-buy decision, the manager should (a) consider the quality of the tests offered, (b) compare the cost of buying the test with performing it in-house, (c) determine the medically necessary turnaround time, (d) evaluate consultation and clinician review capabilities, and (e) estimate the problems with the logistics of sample transport.

If the decision is made to send tests to another laboratory, the laboratory's financial manager should prepare a statement that compares the laboratory's cost of performing needed tests with the potential vendor's price. The statement should present the differential costs of the tests as well as a share of existing fixed expenses and a profit figure that compares the costs on a comparable basis. The laboratory's budget should also be revised to indicate the effect of the change once the transfer of tests is made. A worksheet for calculating in-house test costs and pricing to aid make-or-buy decisions can be found in Appendix B.1.

The Break-Even Point

The break-even point is the point at which there is no profit or loss, or, in other words, the point at which total revenues and total expenses are equal and profit is zero. The most useful purpose of break-even analysis in the laboratory is for making decisions on pricing tests. It is calculated by dividing the laboratory's total fixed costs plus desired profit by the difference between revenue per test and variable cost per test.

The general formula for the break-even point is as follows:

$$\text{Break-even point} = \frac{\text{Fixed expenses} + \text{Desired profit}}{\text{Revenues} - \text{Variable expenses}}$$

Using the graphic technique, the break-even point can be plotted with dollars of revenue on the *y* axis and test volume on the *x* axis. The graph is a composite of (a) the revenue in dollars plotted against test volume, (b) total variable expense plotted against test volume, and (c) total fixed expense plotted against test volume (Fig. 1.23). More detailed examples and techniques for determining the break-even point are presented in Gaither and Resinger (33) and Travers (27). Figure 1.23 shows not only the break-even point but, as revenues increase beyond it, it also shows the development of net income (profit). The break-even volume is useful for determining whether a procedure should be performed in-house or sent to a reference facility. If the calculated break-even volume is higher than current volume, it is not economically feasible to perform the test in-house (27,31,35,37).

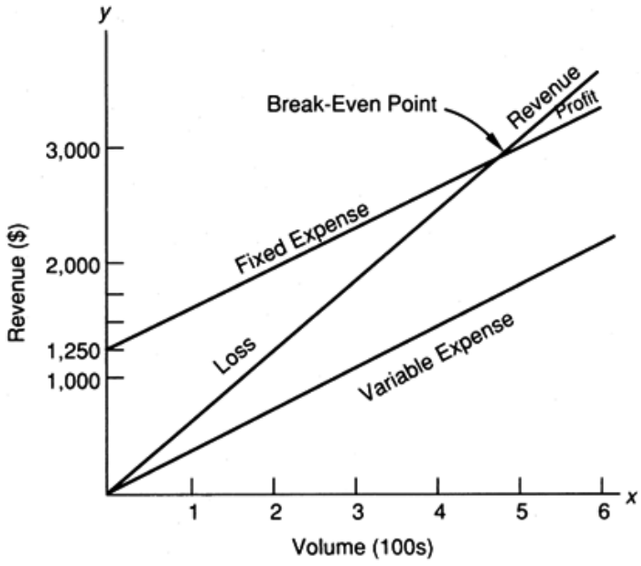


FIGURE 1.23. Break-even chart. (Adapted from Gaither J, Resinger H. *Cost accounting in the laboratory*. Half Day, IL: American Pathology Foundation, 1981:10–14.)

Cost Control and Cost Avoidance

Causes

When all problems and sources of difficulty in managing laboratories have been considered, there are several general conclusions that can be drawn and easily recognized by experienced managers who have long struggled with the dynamics of workload and budget. These basic assumptions are (27)

- The supply (output) of tests is directly proportional to clinical demand
- Costs are directly proportional to laboratory test output until the maximum point of efficiency is exceeded and compliance with regulatory actions and quality control requires more procedures
- Continuous test-order input, without reduction in output, will escalate costs

As previously noted (in the Strategies, Problems, and Principal Areas of Laboratory Expenditure section), there are three basic areas that need correction before any improvement can be made in reducing or avoiding costs. There are (a) lack of information about clinical user needs, (b) lack of physician/provider education in the use of diagnostic procedures, and (c) increasing workload, quality control, inspection, compliance, and accreditation requirements. Problems with uncontrolled costs in laboratories arises first from the inefficient use of automated equipment, primarily from the acquisition and operation of too many high cost analyzers in the automated sections of the laboratory. It is critical to perform a workflow and configuration analysis to determine the actual need for each analyzer, based on the input (workload demand) from clinical services (38). The problems with operating too much automated equipment are (a) it generates too many labor costs, (b) optimum discounts are not always available unless test volume per instrument is very large, (c) it causes a requirement for additional quality control, standards and maintenance, and (d) it is usually not used to full capacity in noncommercial settings (39). Furthermore, if excessive or additional equipment is operated when staffing is less than optimal, the staff will be overstressed, will exceed its point of maximum efficiency, may produce errors, and will increase expenditures by using more reagents and supplies to keep up with the unchecked demand (Fig. 1.24).

SOLUTIONS

Finding solutions to the combined clinical-laboratory problems noted above is not an easy task. In the preceding section, only a few of the most significant problems were outlined. In fact, there is little evidence that clinicians have made any attempt to alter their test-ordering behavior. Even restrictive reimbursement policies have not succeeded in reducing the annual volume of laboratory tests. The only way to begin to reduce unnecessary demand is for the laboratory physician and clinical scientists to take the lead in promoting test use appropriateness in the health care environment.

Test Appropriateness Education

In 1990, the JCAHO encouraged test appropriateness education when it published a standard that provided laboratory physicians with a strong policy to support education of clinical physicians in test appropriateness (40). The standard (PA 1.2.7), part of the pathology and medical laboratory standards in JCAHO's annually published *Accreditation Manual for Hospitals* (40), stated: "Within the hospital's overall assurance program, the director of pathology and medical laboratory services assures that the department/service participates in the monitoring and evaluation of the quality and appropriateness of the services provided."

Furthermore, this standard (now part of a matrix of factors) is a key factor in the decision-making process for accreditation. The standard allows the laboratory director and managers to set up education guidelines, working in conjunction with clinical physicians and other providers for test appropriateness training.

The Pareto Principle (80/20 Rule)

The task of deciding appropriateness for all tests in the laboratory's armamentarium is impossible to accomplish unless a method is used to select the tests and services that are most commonly employed. Therefore, the Pareto principle (the 80/20 rule) can be applied to measure utilization of the most commonly performed tests. The most commonly performed tests account for the major portion of the laboratory's resource expenditure. The application of this concept is discussed in *Cost Accounting Guidelines* published by the National Committee on Clinical Laboratory Standards (32), which describes a method for selecting tests.

Management Control through Interpretive Reporting

Proceeding hand-in-hand with test appropriateness education for clinicians, the laboratory director and managers must always take the lead by providing consultations (where allowed by reimbursement regulations) and by providing interpretive reporting of laboratory test results, especially for abnormal values. There is a great need for proper use guidelines for newer, esoteric tests now ordered in large numbers because of new discoveries in molecular biology and the human genome. Also, written educational information, improved cumulative laboratory reports with narrative statements, and electronic information about proper test use should be available to providers at their workstations. Some areas to begin this exercise are therapeutic drug monitoring for antibiotics, cardiac, pulmonary, and other commonly used therapeutics (40), immunologic tests, DNA probes, thyroid function tests, and other tests as they emerge on the market.

The philosophy is to convince the clinician that when a patient is seen in inpatient practice, the hospital is only prospectively reimbursed at a fixed rate for DRGs, which does not include an equitable rate for laboratory tests. If the tests ordered are excessive or inappropriate, the laboratory spends its valuable resources that could be used for other tests. As of this writing, the outpatient reimbursement for the top 25 laboratory tests has been cut back to 74% of actual cost by the Health Care Financing Administration. The need for interpretive reporting is obvious, to conserve scarce resources and help our clinical colleagues who no longer have the time or learning resources to educate themselves on proper test use.

Management Control Using Optimized Test Groupings

This method allows laboratory management to establish optimized sets of diagnostic criteria that are "customized" for each practice setting. Using the Delphi technique, minimum diagnostic criteria are designed by the hospital's top clinical specialists, agreed on by the physician staff and medical executive board. The laboratory's role is to provide the test selection forms, educate the clinical staff, prove to clinicians that smaller, optimized test groupings for diseases and conditions provide greater quality of care because they avoid "reflex" testing, iatrogenic anemia from multiple phlebotomy episodes, and faster turnaround time. An additional benefit accrues when clinicians are advised which tests are not eligible for reimbursement.

Optimizing Diagnostic Patterns of Care

This method was established in 1985 to advise clinicians on how to set up customized, unique test patterns, based on the standard of medical practice in their own hospital setting (41,42). A similar

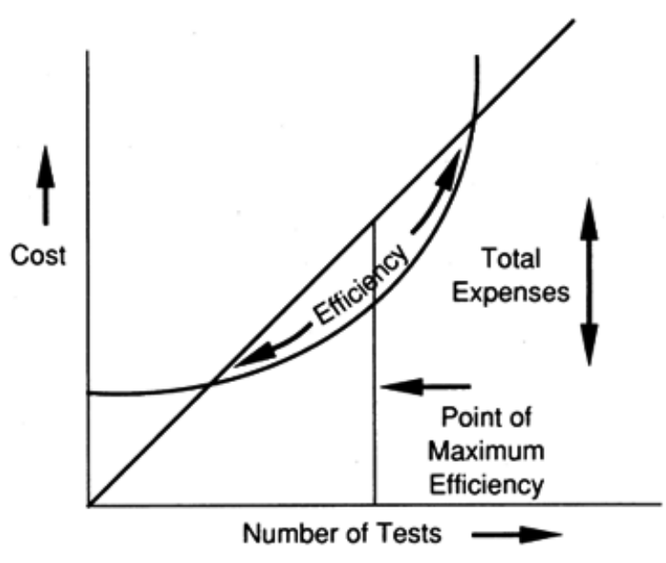


FIGURE 1.24. Effect of exceeding the point of maximum efficiency in the clinical laboratory. (From Craig TM. Dupont Insight. Wilmington, DE: Clinical Diagnostics Division, H.I. Dupont De Nemours Co, Inc., 1983, with permission.)

method, known as “preferred practice patterns” was introduced in 1985 in three Boston hospitals where a group of diseases and operations was studied to identify the lowest cost, highest quality preferred practice patterns. Their approach was to identify the “least resource usage for the same quality output” (43).

Management Control by Clinical Service Leaders

In 1980, Martin et al. (44) emphasized that clinical service leaders have the greatest amount of control and influence over the test-ordering behavior of their house staff in teaching hospitals. This factor helps to reduce unnecessary testing when the laboratory computer system interfaces with clinical service providers to provide them with computerized information about the test menu and frequency and volume of testing by each clinical service and subspecialty. Again, the importance of interpersonal relationship, communication, and networking cannot be emphasized enough in this setting. It is a management control system of great importance for controlling test costs in daily hospital practice.

Deemphasizing the Use of Multitest Profiles

The wasteful practice of multitest profiling, a vestige of past “brainwashing” from the early days of laboratory automation in the 1960s, created a “shotgun” approach to testing because most clinicians believed that it was faster and cheaper to perform large groups of tests simultaneously. It is wasteful because it perpetuates reflex (repeat) testing profiles with larger numbers of tests to verify a previously abnormal test result, causes workload and cost increases, promotes iatrogenic anemia, and gives the false impression that screening for other diseases provides better patient care. In fact, multitest profiling has failed to demonstrate an improved health yield because there is no longer a continuing channel for long-term patient follow-up (44). It is laboratory management's responsibility to make these facts known to clinical providers. Reimbursement for multitest profiles was no longer allowed by Health Care Financing Administration in 1999 for tests listed on the Clinical Laboratory Fee Schedule.

Providing More Efficiency in Automated Areas of the Laboratory

As previously emphasized, the injudicious selection of automated equipment and its inefficient operation causes costs to escalate, especially labor costs, which account for 75% to 85% of all costs. Workflow and optimum configuration studies must be performed before additional equipment is added. Personnel must be utilized more efficiently to perform highly complex, manually oriented tasks, whereas instruments should be chosen to perform high-volume tests at the lowest possible cost per test for tests that are not labor intensive.

Monitoring Laboratory Budgets More Closely

Bulk discounts must be obtained, based on high-volume test needs. Unnecessary testing must be discouraged, and actual and/or activity-based cost analysis must be performed on the high-volume tests that account for 80% of the laboratory's expenditures (see Cost Accounting section). If a test cost is too high and there is insufficient test volume ordered in-house to break even or make a profit, then the test should be sent to another laboratory. Labor scheduling should be seriously looked at, and wasted time must be identified and eliminated. In situations in which part-time staff can be utilized, this option should be exercised to reduce salary costs and provide optimum staffing for the peak periods of workload.

Justifying Capital Equipment Acquisition

The difficulties with availability of capital dollars for investment in laboratory equipment has made it imperative that laboratory managers present a complete, convincing presentation to hospital or corporate managers when requesting additional or replacement equipment.

There are basically four ways to acquire equipment: purchase, rental, lease, or special contractual arrangements with vendors for a cost-per-test fee, in which the equipment is owned by the vendor, who charges a set fee for each test performed.

Regardless of the type of acquisition, the laboratory manager's task is made more complex because of the need to justify the acquisition based on the presentation of cost data to hospital or corporate management, the promise of improved service, and the presentation of data on revenue impact (45). Failure of existing equipment or failure of management strategies to expand the laboratory's market share can further complicate the matter.

WORKFLOW IMPROVEMENT

Part of "1 - Basic Laboratory Management"

It was noted earlier that improving laboratory workflow is the most important goal in acquiring a new analyzer. An instrument should not be selected because of its technology unless the technology has a direct positive impact on improving workflow and use of personnel. The laboratory's workload and test menu must justify additional equipment and the new equipment's throughput must improve turnaround time (27). An analyzer that is forced into an old workflow pattern rarely makes use of all its capabilities. Broader test menus on an instrument are also important because they reduce the total number of pieces of equipment in the laboratory by consolidating workstations, eliminating or reducing batching, replacing existing equipment, and, if possible, reducing staff (10).

COST ANALYSIS

Part of "1 - Basic Laboratory Management"

In addition, before acquisition, a cost analysis should be performed to compare the (i) cost of performing a test (or tests) on the existing equipment with (ii) the cost on the new equipment (see Cost Accounting section). Appendix C provides a worksheet for comparing cost alternatives for the acquisition of laboratory equipment.

METHODS FOR ACQUIRING EQUIPMENT

Part of "1 - Basic Laboratory Management"

Purchase

Purchase of major equipment also requires the performance of complicated analytical techniques to provide hospital or corporate management with information to assess several aspects of financial burden for the hospital or corporation. These include (a) development of depreciable costs, (b) estimation of appropriate depreciation life and timing, (c) isolation of related cash flow variables, (d) forecast of production volume and revenue, and (e) measurement and comparison with hospital corporate criteria. Appendix D provides a guideline for the recommended steps in a capital acquisition process for laboratory equipment. A format for deriving labor management indicators for instrument selection can be found in standard laboratory and business cost management texts (27,30,31,46,47).

If purchase is the desired method of acquisition, then the following key factors should be considered:

- The merits of each capital expenditure
- Future net increases in cash inflows or net savings in cash outflows
- Required total or lifetime investment

Economic or cost-benefit ratios are not the only quantitative considerations to be used in evaluating capital expenditures. The following are also very important:

- The number of patients who benefit
- The degree to which the patients will benefit
- Demographic characteristics of the patients who will benefit
- The benefit to the parent organization for marketing new tests and services

The following factors are also key elements to analyze cost benefits:

- Anticipated time span (years) of project
- Original capital investment
- Annual operating expenses
- Present value of all expenditures
- Average annual savings or net earnings
- Number of patients served
- Age of patients served
- Average cost per patient served
- Intangible short- and long-term health care benefits

Leasing/Rental Arrangements

Many HCOs have entered into leasing or rental arrangements out of necessity, rather than choice. Lease financing has turned into an important source of financing because of the scarcity of capital and the financial squeeze in which many HCOs have found themselves. Leasing, if used properly, can offer advantages. However, financial losses can also occur as a result of not fully understanding the factors involved in a lease arrangement and not properly evaluating the cost of a lease versus other alternatives. The principal advantage of leasing is that it serves as a protection against technological obsolescence and gives the laboratory greater flexibility in the replacement of equipment. With the purchase of equipment, private-sector hospitals will ordinarily be reimbursed only for depreciation and operating costs. Depreciation is calculated on the basis of the original cost of the equipment and may be estimated according to an unrealistic depreciation schedule.

With a lease, the private sector HCO avoids tying up capital. It can use the equipment and (in the private sector) still be reimbursed for the full cost of ownership and the net worth of the equipment is not affected by acquiring the services of the equipment. Federal and state HCOs also avoid the expenditure of capital but will also incur the additional cost for reagents, supplies, and incidentals required for test performance. The vendor also profits by deducting the lease (or rental) costs over time from corporate taxes.

The disadvantages associated with leasing result from the fact that the HCO does not own the asset. Therefore, the residual value at the end of the lease period belongs to the lessor. Second, the failure to correctly estimate the correct technological life of the equipment might result in significant costs.

Cost-per-Test Contracting (Reagent Rental)

This type of acquisition is beneficial in that the vendor provides all reagents, supplies, disposables, automatic data processing, training, installation, maintenance, and all other needs except labor. The HCO or laboratory incurs the burden of labor expense (which encourages the optimal use of the equipment, as well as the more efficient use of labor). The vendor charges a fixed cost per test and provides quality controls, standards, and calibrators. Therefore, it behooves the laboratory to utilize the analyzer judiciously and to perform tests only in the efficient way, avoiding unnecessary tests, repeats, duplicates, wastage, and all other tests that ordinarily would be charged (because they are billed for all reportable results). This arrangement is commonly found in health care networks and multihospital corporations. In 1990, it was established for federal and military hospitals by a contract negotiated with laboratory equipment vendors through the General Services Administration (27).

The advantages of this type of contracting include vendor ownership of equipment, thus avoiding depreciation, maintenance costs, shipping costs, and training costs. Reagents and supplies are provided at the lowest possible cost for the duration of the contract between the HCO and the vendor. An additional advantage for federal and military hospital laboratories is the fixed reagent and supply costs that are negotiated, for instance, for a 5-year contract period.

CONCLUSION

Part of "1 - Basic Laboratory Management"

In summary, before prospective payment and overwhelming reductions in reimbursement of tests and services, the purchase of an analyzer could be justified on the basis of its technology alone. Currently, a good cost-benefit analysis must be performed as well as sensitivity analysis, computations of net present value break-even analysis, internal rate of return, and potential reduction in

variable costs of production. These tools are applied to find the true acquisition costs, as well as lifetime operational costs (45).

Measuring Laboratory Productivity

Managing laboratory costs is impossible without managing laboratory productivity. As the cost of hospital care has increased, the pressure for higher levels of productivity in the provision of services has also increased. One of the interesting aspects of laboratory operations is the potential for substantial improvements in productivity. An increase in productivity will have occurred if (a) the quality of tests provided increases for a given level of technical effort, (b) the quantity of technical effort used to produce a given quality of tests decreases, or (c) given amounts of tests and technical effort remain constant but the quality of tests increases (27,47).

An assessment of laboratory productivity implies that many factors must be included in an analytical process that determines the relative productivity of one laboratory compared with another or to determine total laboratory productivity (Table 1.4). For the purpose of this discussion, the concepts of work units and standard test times, usually thought of as “productivity” standard for laboratories, are considered only as one of several factors required for the total assessment of laboratory productivity. The shortcomings and limitations of using work units or standard times for tests as measures of planning and controlling labor activity are discussed in standard laboratory management references (27,31,32,47). More important, however, is the fact that relative value work units and standard test times by themselves lack the basic ingredients necessary to enhance the improvement of productivity in a managed-care era clinical laboratory (48,49).

TABLE 1.4. TYPES OF PRODUCTIVITY

Major Type	Measures	Subtype	Components	Methods of Measurement	Measurement Tools
Operational	Timeliness of services provided	Labor	Technical staff, clerical staff	Industrial engineering time standards	Task list, list of test times per instrument or method times, analysis sheets, personnel utilization report
	Quality and timeliness of services provided	Automation	Instruments, computers	Maintenance standards, temperature standards	Maintenance checklists, manufacturer's tolerance limits
	Quality and timeliness of agency	Clinical	Test turnaround time, precision, accuracy, reproducibility	Quality control standards	Quality control reagents, reference standards
	Quality of services provided	Workload	Tests, quality control standards	Raw test count, relative value per test	Manual recording, computerized recording
Fiscal	Economic efficiency	Departmental, sectional, inpatient, outpatient	Laboratory work stations, laboratory sections	Cost accounting, cost allocation	Standard costs, operating expenses = actual costs, operating revenues, ratios

From Travers EM. *Managing costs in clinical Laboratories*. New York: McGraw-Hill Information Systems, 1989:204, with permission.

Instead, it is more important as a whole that total laboratory productivity, rather than the workload output of individual technical and clerical workers, must be the center of management's concern.

Types of Laboratory Productivity

There are four types of operational productivity: labor, automation, clinical, and workload productivity, and four types of fiscal productivity: departmental, sectional, inpatient, and outpatient productivity (Table 1.4). Although all are important and must be integrated to achieve successful laboratory operations and financial performance, the most important of all eight types is labor productivity because all other categories of productivity and successful financial outcome depend on the efficient utilization of labor resources.

Labor Productivity

The growing importance of the role of management in improving labor productivity in clinical laboratories is especially pertinent in times of reduced reimbursement and economic constraint. The basic concepts underlying productivity and productivity measurement in businesses are quite clear and simple. However, in a clinical laboratory environment, measuring output in terms of a vast test universe of more than 900 tests and services is not always simple. Laboratory tests are subject to technological changes in methodology, and their performance times vary greatly, depending on the equipment and/or method used to perform the test. Furthermore, the laboratory manager is not always equipped with the knowledge or the time to apply workload measurement criteria precisely. Consequently, a more global approach

is required to assess the laboratory's labor productivity. Appendix E provides a basic checklist/score sheet and a list of solutions for managers to assess their own causes of low labor productivity, as well as solutions for low labor productivity.

The Productivity of Automation

This subtype of productivity is difficult to separate from labor productivity, but it does have distinctly different characteristics that can be improved to enhance the quality and timeliness of services provided and provide more cost-effective utilization of equipment. Unfortunately, automation is often purchased and utilized in this era based on empirical needs, and is not always matched to workload volume or test complexity, which causes under- or over-utilization. Table 1.5 provides suggestions for improving the productivity of automation.

TABLE 1.5. STEPS FOR IMPROVING THE PRODUCTIVITY OF AUTOMATION

Consolidate priority tests.
Multichannel vs. discrete test costs should be determined.
Optimize test groupings (laboratory disease-related groups).
Interact with clinicians to prioritize/consolidate test/disease profiles.
Utilize cost accounting techniques for instruments, sections, and departments.
Utilize automatic data processing to decrease turnaround time- decrease telephone/clerical/billing time.
Improve capital productivity.
Use cost accounting. Determine most/least expensive procedures. Evaluate cost of lease/purchase or rent vs. capital purchase.

From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw Hill Information Systems, 1989:2135 with permission.

Clinical and Workload Productivity

Clinical productivity is synonymous with quality. The importance of producing a product (test) of the highest quality is especially important in view of increasing regulations from private sector accrediting bodies (JCAHO, College of American Pathologists) and federal regulatory action (CLIA '88). A laboratory manager's greatest concern is balancing quality with economic productivity and financial outcome, and there is no room for a trade-off of one or the other. However, there is some concern in today's highly computerized laboratory that there is too much quality control, especially when analyzers have been programmed to recognize errors instantaneously. This can be an additional area of unnecessary expense and, if carefully reviewed, might result in considerable savings without reducing the quality of test results.

The other aspect of clinical productivity is the need to perform more cogent, specific test patterns, as discussed in the Test Appropriateness Education section. Table 1.6 provides recommendations for improving clinical productivity. These recommendations, if implemented, can improve workload productivity by reducing unnecessary testing.

TABLE 1.6. RECOMMENDATIONS FOR IMPROVING CLINICAL PRODUCTIVITY

Develop laboratory disease-related test groups.
Interact with clinical physicians to prioritize and consolidate tests/disease.
Integrate test groups into laboratory's quality assurance plan.
Interact with clinical staff to determine their optimum test choices.
Provide interpretive reporting for clinicians.
Interact with hospital management to demonstrate cost avoidance and encourage administrative decisions that discourage un-necessary testing.

From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:213, with permission.

Improving Fiscal Productivity

Improving the economic efficiency of a laboratory can be applied to an entire department, a laboratory section or workstation(s), or to a hospital component to which services are directed, namely, inpatient or outpatient services. The causes of inefficient fiscal operations are most commonly related to salaries that are too high and/or being paid to too many laboratory staff members. Other nonlabor expenses can be inordinately high (Table 1.7)

but should be considered for review after labor costs have been streamlined because labor costs account for 75% to 85% (or more) of total costs.

TABLE 1.7. CAUSES OF HIGH NONLABOR EXPENSES

Waste
Too many dollars spent on research, development, and nonpatient care
Excessive controls, standards, and repeat tests
Nonlaboratory-related expenses charged against laboratory by the hospital or corporation
Too little spent on capital equipment, too much on reagent rental
Poor purchasing practices, poorly negotiated deals
An inconsistent definition of capital equipment (capital vs. noncapital expenses threshold is high)
Excessive inventory and/or poor inventory control
Excessive send-outs or too few send-outs, with no emphasis on true cost analysis
Excessive preventive maintenance
Poor preventive maintenance practices resulting in high equipment failure
Excessive travel and education expenses
Poor monitoring of long-distance telephone costs
Courier services are not cost controlled
Inappropriate use of substitution products (such as in blood bank, the use of more expensive, but clinically equivalent blood products)
Capital equipment not properly sized for the laboratory workload flow
Stat equipment has high reagent cost, yet fast turnaround time is used for routine testing
Excessive accrediting agency fees
Excessive proficiency testing fees
Excessive dues to organizations
Excessive, unnecessary subscriptions
Excessive, unnecessary consulting fees
An unusual laboratory test mix, based on the hospital patient population
Lack of instrument standardization, resulting in excessive repeats
Need for backup instrumentation
Computer paper wastage

Adapted from Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:217, with permission.

REFERENCES

1. Hardwick DF, Morrison JI. *Directing the clinical laboratory*. New York: Field and Wood, 1990:1–45.

2. Snyder JR, Senhauser DA. *Administration and supervision in laboratory medicine*. Philadelphia: JB Lippincott, 1989:1–20.

3. Westgard JO, Klee GG. Quality assurance. In: Tietz NW, ed. *Textbook of clinical chemistry*. Philadelphia: WB Saunders, 1986:424–458.

4. Gray BH, Field MJ. *Controlling costs and changing patient care, the role of utilization management*. Washington, DC: National Academy Press, 1989:27–58.

5. George CS Jr. *The history of management thought*. Englewood Cliffs, NJ: Prentice-Hall, 1965:13–17.

6. Albrecht K. *Successful management by objectives*. New York: Prentice-Hall Press, 1988:1–40.

7. Drucker PF. *Managing for results*. New York: Harper & Row, 1964:1–50.

8. deleted in proofs.

9. Garvin DA. *Managing quality*. New York: The Free Press, 1988:1–20.

10. Couch JB. *Health care quality, management for the 21st century*. Tampa, FL: The American College of Physician Executives, 1991:1–22.

11. DePree M. *Leadership is an art*. New York: Bantam/Doubleday/Dell Publishing Group, 1989.

12. Shewhart WA. *Economic control of quality of manufactured product*. New York: D. Van Nostrand, 1931.

13. Juran JM. *Juran's quality control handbook*, 4th ed. New York: McGraw-Hill, 1988.

14. Juran JM. *Juran on leadership for quality: an executive handbook*. New York: Collier Macmillan, 1989:1–80.

15. DeAguayo R. *Deming: the American who taught the Japanese about quality*. New York: Simon and Schuster, 1990:1–66.

16. Davidow WH, Uttal B. *Total customer service, the ultimate weapon*. New York: Harper & Row, 1989.

17. Albrecht K, Zemke R. *Service America*. Homewood, IL: Warner Books, 1985:1–191.

18. Gronroos C. *Service management and marketing*. Lexington, MA: Lexington Books, 1990:1–257.

19. Naisbitt J. *Megatrends: ten new directions transforming our lives*. New York: Warner Books, 1982.

20. Clark GB. Quality assurance, an administrative means to a managerial end, Part I. A historical overview. *Clin Lab Manag Rev* 1990;17:7–15.

21. Clark GB. Quality assurance, an administrative means to a managerial end, part II. The JCAHO ten step process selecting indicators of quality. *Clin Lab Manag Rev* 1990;17:224–257.

22. McGregor D. *The human side of enterprise*. New York: McGraw-Hill, 1960:1–58.

23. Maslow A. *Motivation and personality*. New York: Harper and Bros., 1954.

24. Thompson RE. From quality assurance to continuous quality improvement. *Physician Exec* 1991;17:3–8.

25. National Committee for Clinical Laboratory Standards. *A quality system model for health care: approved guidelines*, NCCLS Document GP26-A. Wayne, PA: NCCLS, 1999.

26. Fromberg R. *The joint commission guide to quality assurance*. Chicago: Joint Commission on Accreditation of Healthcare Organizations, 1988.

27. Travers EM. *Clinical laboratory management*, 1st ed., chapter 12. Baltimore: Williams & Wilkins, 1997:141–158

28. Hoerger TJ, Eggleston JL, Lindrooth RC, et al. *Background report on the clinical laboratory industry*, Research Triangle Institute Report RTI6864FR, Center for Economics Research. Research Triangle Park, NC, 1997;21:33–58.

29. Wilkinson SL. How are we, as a nation, going to pay for safety? AABB News 1999:2–3.

30. Oszustowicz RJ. A capital equipment acquisition process. *J Healthcare Finan Manag Assoc* 1982;30 (April).

31. Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989.

32. National Committee for Clinical Laboratory Standards. *Basic Cost Accounting for Clinical Services* (GP11-A), vol. 18, no. 14. Wayne, PA: National Committee for Clinical Laboratory Standards, 1998.

33. Gaither J, Resinger H. *Cost accounting in the laboratory*. Mundelein, IL: The American Pathology Foundation, 1981:56.

34. Cleverley WO. *Handbook of healthcare accounting and finance. Volume I*. Rockville, MD: Aspen Systems, 1982:295.

35. Berman HJ, Weeks LE, Kukla SF. *The financial management of hospitals*, 6th ed. Ann Arbor, MI: Health Administration Press, 1986:457–458.

36. Horngren C. *Cost accounting: a managerial emphasis*, 6th ed. Englewood Cliffs, NJ: Prentice-Hall, 1982.

37. Eastaugh R. *Financing health care: economic efficiency and equity*. Dover, MA: Auburn House, 1987:111.

38. Anthony RN, Welsch GA. *Fundamentals of management accounting*. Homewood, IL: Richard D. Irvin, 1974:134–140.

39. Lifshitz MS, DeCresce RP. Clinical laboratory instrument selection. *Lab Med* 1990;21:367–370.

40. Joint Commission on Accreditation of Healthcare Organizations. *Accreditation Manual for Hospitals*. Chicago: Joint Commission on Accreditation of Healthcare Organizations, 1990:137–138.

41. Travers EM. Misuse of therapeutic drug monitoring. *Clin Lab Med* 1987;7:453–472.

42. Travers EM. *Optimizing diagnostic patterns of care*. Washington, DC: Department of Veterans Affairs, Veterans Health Services and Research Administration, Publication IB 11-82, 1989.

43. McNeil B. Hospital response to DRG-based prospective payment. *Med Decision-Making* 1985;5:15–21.

44. Martin AR, Wolf MA, Thibodeau LA, et al. A trial of two strategies to modify the test-ordering behavior of medical residents. *N Engl J Med* 1980;303:130–236.

45. Home HF. Application of AMHT in clinical medicine. *JAMA* 1972;219:885.

46. Benezra N. Re-equipping the laboratory: Part II: how to win approval for acquisitions. *Med Lab Observer* 1989;21:33–35.

47. Travers EM. *Workbook for managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:117–124.

48. Hadley J. Research on health manpower productivity: a general overview. In: Rafferty J, ed. *Health manpower and productivity*. Lexington, MA: D.C. Heath, 1974.

49. Ozan TM. *Productivity measures for the improvement of productivity in clinical laboratories. Examination of case studies of productivity improvements in clinical laboratories*. Chicago: Center for Hospital Management Engineering of the American Hospital Association, 1974:243–261.

a Profit is defined as relative efficiencies of scale (or break-even + adjustment for inflation) for government-operated laboratories and other not-for-profit laboratories; relative efficiencies of scale include cost avoidance and cost savings produced as a result of efficient management of personnel (most costly aspect of operational costs) and operational resources; these "saved" resources are usually recorded by the not-for-profit hospital's finance department and are redistributed by hospital management during the current fiscal year in which they "saved."

b Publicly funded laboratories will not use these categories unless they generate revenue for the hospital by performing tests for other laboratories using "sharing" agreements that specify that a fee will be collected for a service shared between hospitals. Ordinarily, the revenue collected will only cover costs incurred for tests performed for the duration of the sharing agreement.

APPENDICES

APPENDIX A: INSTRUMENT COST ACCOUNTING TECHNIQUE

If one component of laboratory operations can be identified as the controlling influence, it is the instrument configuration of the laboratory. It determines the major portion of costs (labor) for any laboratory and determines actual costs for the next largest component—reagents and supplies. The information and formats in this section provide laboratory managers with a microcosting "bottom-up" method for determining cost per test per instrument. This is the basic starting point in any laboratory cost accounting technique (Fig. 1.A1).

I. TEST INSTRUMENT

Name: GC/MS — Life Expectancy in Years: 5
Model Number: 59708 — Annual Maintenance Cost: 9442
Manufacturer: H-P — Site-Preparation Cost: 2000
Purchase Price: 90000 — Evaluation Period in Days: 1
Starting Date: 9/15/89
Completion Date: 9/15/89

II. DIRECT TEST MATERIALS

Name of Test/Profile	Total No. Tests/Profiles	Per Test/Profile in Evaluation Period: $ Reagents	$ Test-Related Supplies and Parts	$ Equipment-Related Disposables
A	B	C	D	E
1. Benzoylecgonine	30	$27.57	$12.56 (supplies)	$1.89
2.			$9.00 (parts)	
3.				
4. Controls	10	$190.00	included above	.63
5. calibrators	10	$190.00	" "	.63
6.				
7.				
8.				
9.				
10.				
Totals	30 tests	$407.57	$21.56	$3.15

III. TEST LABOR

Time Segment per Test/Profile	Total FTE Minutes/ Procedure	Annual Salary + Benefits
Preanalytical	21 min.	$33880 + 4404 = $38284
Analytical	180 min.	38284
Postanalytical	120 min.	38284
Totals	321 min. ÷ 15 PRD FACTOR = 3692	

FIGURE 1.A1. Completed ICAT worksheet showing cost accumulation for gas chromatography/mass spectrometry procedure. See Fig. 1A.2 for instructions. (Adapted from Travers EM. *Workbook for managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:72–73.)

IN THE FIRST PART (I):

1. Enter the trade name of the instrument.
2. Enter the instrument model number, if applicable.
3. Enter the manufacturer's name.
4. Enter the purchase price of the equipment.
5. Enter the anticipated life expectancy of the equipment in years.
6. Enter the annual maintenance cost of the equipment.
7. Enter the site-preparation cost.
8. Enter the evaluation period (in whole days) and the starting date of the evaluation period.

IN THE GRID PORTIONS (II AND III):

9. Enter the name or designation for each test or panel that will be performed on the instrument. If the instrument performs both discrete and batch testing, do this analysis for both modes of operation and duplicate this form if necessary (Part A).
10. Enter the total number of tests or profiles performed on this instrument during the evaluation period (Part B).
11. Based on reagent costs incurred during the evaluation period, enter the reagent cost per test profile (Part C).
12. Based on cost of disposables used during the evaluation period, estimate the test-associated disposable cost per test/profile (Part D).
13. Enter the cost of any other equipment-related disposables per test/profile used during the evaluation period (Part E).
14. Continue with the labor analysis section and enter:
15. *Preanalytical time*—Include *time* to collect specimens from patients and laboratory central receiving. This includes all steps up to the actual testing procedure, such as work-list gathering, start-up, sample cup preparation, and preparation of daily quality controls and standards for this instrument. Phlebotomy or centrifugation is included.
16. *Analytical time*—Include labor to analyze the specimen(s) and to perform all routine procedures up to reporting of results. This includes calculation(s) and checking *but does not include* repeats.
17. *Postanalytical time*—Include labor to manually report results or to enter results into a computer system. This includes sorting, filing, and telephone calls related to the final report(s). Routine, daily maintenance normally performed and shut-down time must also be included.

When all blocks have been completed, provide totals where indicated and proceed to Figure A.9 (Formulas for ICAT).

FIGURE 1.A2. Instructions for completing the ICAT worksheet. (From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:112, with permission.)

No standard cost accounting format has been developed for use in laboratories, and standard accounting texts are not suitable for most medical applications. The examples in the texts are not pertinent to laboratories and many of the problems are irrelevant. Therefore, the ICAT was developed as a method and in a form that managers can use in any laboratory for any instrument or combination of instruments to calculate the cost of tests or test profiles.

ICAT is based on standard cost accounting techniques and is derived from standard literature and generally accepted accounting principles in the field of accounting. These cover direct labor, materials, and instrument-related costs. Indirect expenses are

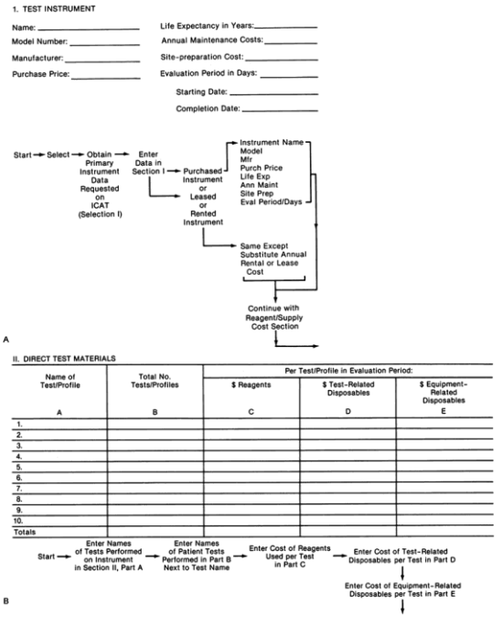


FIGURE 1.A3. Section II. Data entry. **A**, indirect instrument costs; **B**, direct material costs. (Adapted from Travers EM. *Workbook for managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:52 and Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:126.)

usually estimated in each individual setting and will vary considerably from laboratory to laboratory, based on hospital or laboratory reimbursement. Indirect expenses are discussed in a later section because this important component of total cost is not considered in the overall determination of direct test production costs.

ICAT Format

The ICAT method considers the three prime cost aspects of analysis: (a) instrument-related costs, (b) direct materials cost, and (c) labor costs. No two laboratories are alike in their instrumentation configuration; therefore, ICAT was designed to be generic so that it could be used in each section of any laboratory to profile costs associated with every instrument or test method. Direct material expenses on the ICAT form include expenditures for supplies and reagents required to produce tests. Direct labor costs comprise total compensation required for preanalytic, analytic, and postanalytic personnel necessary to perform tests. By dividing all test production activities into the three major work effort segments, ICAT allows managers to estimate which segment of the test production process has the greatest impact on manpower, and, subsequently, on compensation costs.

Sample ICAT Worksheet

The reader should know that the ICAT method is universal and generic and can be used for any test performed in the laboratory. Regardless of whether an instrument is used, the method will be the same. If no instrument is used, then the section on indirect instrument costs (Section 1 of the ICAT) is left out and the remaining sections are completed. The ICAT worksheet is used to accumulate costs (Fig. 1.A1); instructions for completing each major section of the ICAT appear in Fig. 1.A2.

Examples and algorithms for actual cost analysis for gas chromatography/mass spectrometry are presented in Fig. 1.A3,Fig. 1.A4,Fig. 1.A5,Fig. 1.A6,Fig. 1.A7,Fig. 1.A8,Fig. 1.A9,Fig. 1.A10,Fig. 1.A11,Fig. 1.A12,Fig. 1.A13 and Fig. 1.A14

1. Depreciation costs[a] $= \frac{\$90000}{5\ \text{yrs}} \times \frac{1\ \text{year}}{365\ \text{days/yr}} \times \frac{1\ \text{d.}}{30\ \text{tests/pd.}} = \frac{90000}{54750} = \$1.64/\text{test}$

2. Maintenance costs[b] $= \frac{\$47210}{5\ \text{yrs}} \times \frac{1\ \text{year}}{365} \times \frac{1\ \text{d.}}{30\ \text{tests/pd}} = \frac{47210}{54750} = \$0.86/\text{test}$

3. Site-preparation costs[c] $= \frac{\$2000}{5\ \text{yrs}} \times \frac{1\ \text{year}}{365} \times \frac{1\ \text{d.}}{30\ \text{tests/pd.}} = \frac{2000}{54750} = \$0.04/\text{test}$

	ANNUAL COST
EQUIPMENT	
Depreciation[a]	$90000/5 yr; 18000/yr
Lease Fee	—
Rental Fee	—
Maintenance[b]	9000 svc contract; $442 tech labor (2 hr/mo) = $9442/yr
Service Parts Kit[b]	—
Service Contract[b]	incl. in maintenance cost
Interface	—
Accessory Equipment[b]	—
Other	—
Subtotal	$27442/yr
INITIAL & RELOCATION SETUP COSTS	
Heating/Air Conditioning	$—
Water System[c]	500
Drain[c]	500
Power[c]	500
Remodeling[c]	500
Correlation Studies	—
Start-Up Kit[c]	—
Other	$—
Subtotal	$2000/5 yr; $400/yr
LAB INDIRECT COST TOTAL	$27,842

FIGURE 1.A4. Equipment indirect cost worksheet summary. (Adapted from Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:112–115 and Travers EM. *Workbook for managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:56.)

CONSUMABLES	QUANTITY PER 30 TESTS	COST PER RUN
FLUIDS		
Reagents	1	$27.57
Diluents	—	—
Calibration material	1	190
Reference fluids	—	—
Linearity standards	—	—
Control fluids	1	190
Rinse/wash solutions	—	—
Cleaning solutions	—	—
Deionized water	—	—
Analyzer specific solutions	—	—
Other		
Subtotal		$407.57/run
SUPPLIES*		
Chemicals	1	$10.46
Gasses	—	—
Paper	1	$50/yr.; ($0.02/test × 30) = $.60/run
Labels	—	—
Ribbons	1	$60/yr.; ($0.03/test × 30) = $.90/run
Water	—	—
Cartridges, filters, etc	1	$56/yr.; ($0.03/test × 30) = $.90/run
Other	—	—
Subtotal		$12.56/run
PARTS*		
Syringes	1	$600/yr.; ($0.29/test × 30) = $8.70/run
Tubing	—	—
Membranes	1	20.80/yr.; ($0.01/test × 30) = $.30/run
Valves	—	—
Reagent Caps	—	—
Seals	—	—
Reagent vessels	—	—
Cuvettes	—	—
Vials	—	—
Lamps	—	—
Needles	—	—
Probes	—	—
Electrodes	—	—
Filters	—	—
Detectors	—	—
Reaction Chambers	—	—
Other	—	—
Subtotal		$9.00/run
DISPOSABLES*		
Sample cups (Test Tubes)	50	$80/yr.; ($0.04/test × 30) = $1.20/run
Sample caps	50	$30/yr.; ($0.01/test × 30) = $0.30/run
Pipette tips	50	$30/yr.; ($0.01/test × 30) = $0.30/run
Glass tubes	50	$75/yr.; ($0.04/test × 30) = $1.20/run
Bar code labels		
Micro inserts/cups		
Cuvettes		
Other		$10/yr.; ($0.005/test × 30) = $0.15/run
Subtotal		$3.15/run
TOTAL		$432.28/run

FIGURE 1.A5. Laboratory material cost worksheet summary. *Cost estimates for supplies and parts based on 2080 tests/yr. (Adapted from Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:237–238.)

Section I: Indirect Costs of Test Production for Test Instrument

Indirect costs are all costs of test production not traceable to a test (unit of output) or to a segment of the equipment or operational requirements for test production. Indirect costs include those of materials, indirect labor, and all other expenses of test production that cannot conveniently be charged directly to laboratory sections, jobs, or products (Fig. 1.A2). They are assigned to the total cost based on a multitude of factors that are part of the hospital or corporate operations. These are commonly added into the cost of tests as a percentage rather than as actual

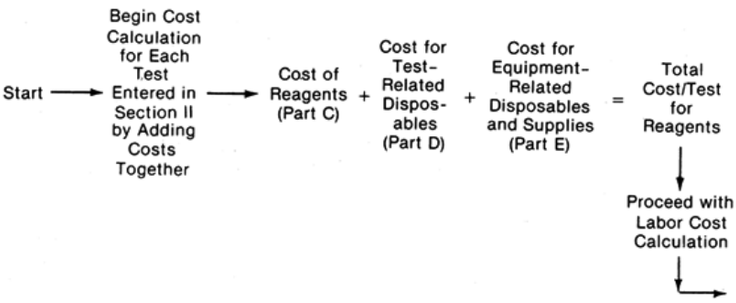


FIGURE 1.A6. Section II: calculation of direct individual costs for reagents and supplies and equipment-related disposables. Part 1. Calculation of reagent and supply costs. (Adapted from Travers EM. *Managing costs in clinical laboratories*. New York, McGraw-Hill Information Services, 1989:128.)

III. TEST LABOR

Time Segment per Test/Profile	Total FTE Minutes/ Procedure	Annual Salary + Benefits
Preanalytical		
Analytical		
Postanalytical		

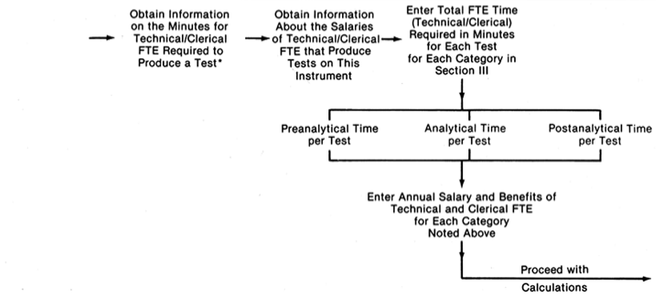


FIGURE 1.A7. Section III data entry (continued): direct labor costs. *Involves all phases of production, starting with specimen collection and ending with delivery of test results to user. (Adapted from Travers EM. Workbook for managing costs in clinical laboratories. New York: McGraw-Hill Information Services, 1989:52.)

LABOR COST ACCUMULATION WORK SHEET

WORK MEASUREMENT PROJECT NONREPETITIVE TIME STUDY — 1. Study Number /

2. Operation *Clinical Laboratory* — 3. Organization *University Hospital*

4. Test Name and Type *Gas Chromatography/Mass Spectrometry; Benzoylecgonine*

5. Location *Toxicology section*

6. Operator's Name *Nancy Broderick*

7. Date *9/15/89*

8. Start Time *8:00 a.m.* — 9. Stop Time *2:09 p.m.* — 10. Elapsed Time *6 hrs, 9 min.*

	11. Element Description	Readings	Time	Level Factor	Normal Time	Occr. Factor	Base Time
	A	B	C	D	E	F	G	H
Pre-Analytical Time (21 min)	1	Make work list (computer)		5 min.				
	2	Get controls		2				
	3	Label tubes		4				
	4	Tune MS		10				
Analytical Time (150 min)	5	Pipette samples		20				
	6	Add internal std to buffer, mix soln.		2				
	7	Shake 20 min		2				
	8	Aspirate top layer		20				
	9	Transfer bottom layer		20				
	10	Evaporate		2				
	11	Add derivatizing reagent		10				
	12	Vortex; incubate		20				
	13	Wash		30				
	14	Centrifuge		4				
	15	Aspirate		20				
	16	Transfer upper layer to another tube		20				
	17	Evaporate		2				
	18	Inject into GC/MS		8				
Post Analytical Time (120 min)	19	Evaluate results		80				
	20	Enter results into computer and controls into log book		30				
	21	Copy data file on to Tape		10				

12. Remarks: Run of 30, including 10 standards, 10 controls

13. Total Base Time: 321

14. PF&D* Allowance: 15 % → 48.15

15. Standard Time for One Test: 369.2 / 30 = 12.3 min. — Min. 369.2 for 30 tests; His. 6.15 for 30 tests

16. Approved Signature

17. Date: 9/15/89

FIGURE 1.A8. Labor time accumulation worksheet. PF&D, personnel fatigue and delay. (From Travers EM. Workbook for managing costs in clinical laboratories. New York: McGraw-Hill Information Services, 1989:30, with permission.)

1. Materials Costs/(Test or Profile) $= \frac{\text{Cost of Reagents + Cost of Test-Associated Supplies \& Parts + Cost of Equipment-Related Disposables}}{\text{\# of Tests or Profiles}}$

$= \frac{407.57 + 21.56 + 3.15}{30 \text{ tests}} = \frac{\$432.28 \text{ run}}{30 \text{ tests}} = \boxed{\$14.40/\text{test}}$

2. Labor Costs/(Test or Profile) $= \frac{\text{Salary Cost*}}{\text{year}} \times \frac{1 \text{ year}}{2080 \text{ hours}} \times \frac{1 \text{ hr}}{60 \text{ min}} \times \text{\# min. per test or run}$

$= \frac{\$38284 \times 1 \text{ yr} \times 1 \text{ hr}}{1 \text{ yr} \times 2080 \text{ hrs} \times 60 \text{ min}} \times 369.2 \text{ min per run} = \$0.31 \times 369.2 = \boxed{\$114.45/\text{run}}$

or $\div 30$ tests $= \boxed{\$3.82/\text{test}}$

3. Instrument Depreciation Costs/(Test or Profile) $= \frac{\text{Instrument Purchase Price}}{\text{Anticipated Instrument Life Expectancy (in years)}} \times \frac{1 \text{ year}}{365 \text{ days}} \times \frac{\text{\# Days in Evaluation Period}}{\text{\# Tests in Evaluation Period}}$

$= \frac{\$90000 \times 1 \text{ yr} \times 1 \text{ day}}{5 \text{ yrs} \times 365 \text{ days } 30 \text{ tests/pd.}} = \frac{\$90000}{54750} = \boxed{\$1.64/\text{test}}$

4. Maintenance Costs/(Test or Profile) $= \frac{\text{Total Cost for Maintenance, Service, and Repairs}}{\text{Anticipated Instrument Life Expectancy (in years)}} \times \frac{1 \text{ year}}{365 \text{ days}} \times \frac{\text{\# Days in Evaluation Period}}{\text{\# Tests in Evaluation Period}}$

$= \frac{\$47210 \times 1 \text{ yr} \times 1 \text{ day}}{5 \text{ yrs} \times 365 \text{ days } 30 \text{ tests/pd.}} = \frac{\$47210}{54750} = \boxed{\$0.86/\text{test}}$

5. Site-Preparation Costs/(Test or Profile) $= \frac{\text{Costs to Install Instrument (Elec., Plumb., A/C, Constr., etc.)}}{\text{Anticipated Instrument Life Expectancy (in years)}} \times \frac{1 \text{ year}}{365 \text{ days}} \times \frac{\text{\# Days in Evaluation Period}}{\text{\# Tests in Evaluation Period}}$

$= \frac{\$2000 \times 1 \text{ yr} \times 1 \text{ day}}{5 \text{ yrs} \times 365 \text{ days } 30 \text{ test/pd.}} = \frac{\$2000}{54750} = \boxed{\$0.04/\text{test}}$

6. Indirect Costs = (Materials + Labor + Instrument Costs) $\times$ 60%†

$= (\$14.40 + 3.82 + 1.64 + .86 + .04) \times .60 = \boxed{\$12.46}$

FIGURE 1.A9. Formulas used for test cost determination for laboratory tests. * Benefits at 13% rate included in salary cost (see Fig. 6); if benefits are not included in salary cost, then a factor must be added to this equation to multiply the results by 1.13 if the benefits are 13%. Percentage will vary in each laboratory/hospital; see Fiscal Services for rates. (Adapted from Travers EM. Managing costs in clinical laboratories. New York: McGraw-Hill, 1989:227-228.)

amounts for each item (see Formula 6, in Fig. 1.A9 and Fig. 1.A10.)

Depreciation, Maintenance, and Site-Preparation Costs Ten entries in part I of the ICAT form (Fig. 1.A1 and Fig. 1.A3) provided the format for indirect instrument-related cost calculations. They include the instrument's name, model number, manufacturer's name, purchase price, life expectancy in years, annual maintenance cost, cost of site preparation, length of the evaluation period in days, and date that the evaluation was started and completed. A cost accumulation worksheet is provided in Fig. 1.A4, which is coded to the formulas that follow. From this information, three different equations are used to calculate each instrument's deprecation, maintenance expenses, and site-preparation costs per test or profile. They are as follows:

$$\text{Depreciation costs} = \frac{\text{purchase \$}}{\text{years life}} \times \frac{\text{1 year}}{365} \times \frac{\text{\# evaluation days}}{\text{\# total tests}}$$

$$\text{Maintenance costs} = \frac{\text{Total maintenance \$}}{\text{years life}} \times \frac{\text{1 year}}{365} \times \frac{\text{\# evaluation days}}{\text{\# total tests}}$$

$$\text{Site-preparation costs} = \frac{\text{site \$}}{\text{years life}} \times \frac{\text{1 year}}{365} \times \frac{\text{\# evaluation days}}{\text{\# total tests}}$$

Section II: Test-Performance Analysis: Direct Materials Costs

It is important to analyze the cost of direct materials used in test production for each test or panel to be performed. If an instrument performs both discrete and batch testing, the analysis should be done on separate ICAT forms as required for both modes of testing. After the name of every test or panel, direct test-performance costs are divided into four segments (Fig. 1.A1, section II; Figure 1A3.B provides a data entry algorithm for direct materials costs):

a. The total number of patient tests or profiles performed during the evaluation period (Fig. 1.A3.B, part IIB)
b. The cost of reagents per test or profile during the evaluation period (Fig. 1.A3.B, part IIC) (use Fig. 1.A5 to accumulate materials costs)
c. The estimated cost of test-associated disposables used per test or profile based on the total cost of disposables used during the evaluation period (Fig. 1.A3B, part IID) (Fig. 1.A5)
d. The cost of any other equipment-related disposables per test or profile used during the evaluation period (Fig. 1.A3B, part IIE) (Fig. 1.A5)

Laboratory activity and consumables to be considered in test-performance analysis include some or all of the following components, depending on whether kits, containerized reagents, or batch or discrete (single) tests are performed: (i) the price per kit or total reagent costs, (ii) the number of tests per kit or total reagent volume, (iii) the number of controls run per test or profile, (iv) the number of standards run per test or profile, (v) single or replicate analysis, as done during each test or profile performance, (vi) the price for expendable items used to perform the test per batch or specimen, (vii) the expense of specimen collection and processing, and (viii) the total number of tests or profiles

DIRECT COSTS		
1. Materials (Reagents/Supplies) Cost	$14.40	Formula 1, Table 9
2. Labor Costs		
a. Preanalytical		
b. Analytical	$ 3.82	Formula 2, Table 9
c. Postanalytical		
INDIRECT INSTRUMENT COSTS		
1. Depreciation Costs	$ 1.64	Formula 3, Table 9
2. Maintenance Costs	.86	Formula 4, Table 9
3. Site-Preparation Costs	.04	Formula 5, Table 9
SUBTOTAL	$	
×(11%) Indirect Costs*	$12.46	Formula 6, Table 9
= Total cost per test	$33.22	

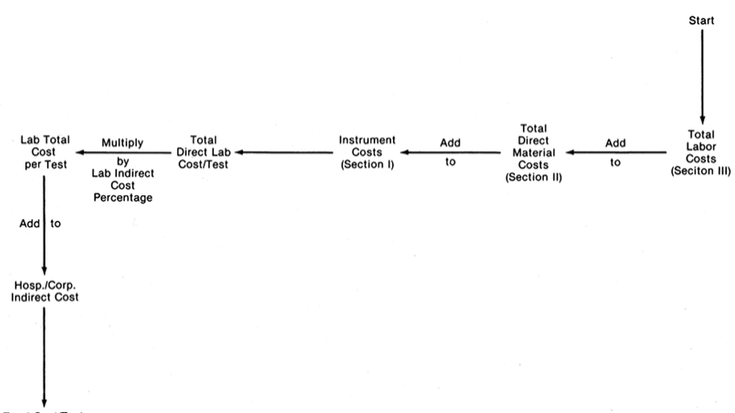


FIGURE 1.A10. Summary sheet-test cost determination. * Exclude all other indirect costs other than instrument-related indirect costs, e.g., rent, insurance, power, heat, lights, taxes, etc.; 11% is only an example since indirect costs are different for each hospital or corporate laboratory; private sector indirect costs may be as high as 146%. (Adapted from Travers EM. *Workbook for managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:57.)

FIGURE 1.A11. Section III, part 2: Calculation of direct labor cost and summary sheet. (Adapted from Travers EM. Managing costs in clinical laboratories. New York: McGraw-Hill Information Services, 1989:128.)

performed during the evaluation period. A flow diagram accumulation of test materials cost is shown in Fig. 1.A6.

To calculate direct materials test costs per test or profile, use the following formula:

$$\text{Material costs} = \frac{\begin{array}{c}\$\ \text{reagents} \\ +\ \$\ \text{test-associated disposables} \\ +\ \text{Equipent related disposables}\end{array}}{\#\ \text{of tests or profiles}}$$

Section III: Labor Costs

Labor costs are divided into three labor-intensive segments of test performance: preanalytic, analytic, and postanalytic time. The total full-time equivalent required to produce a test must be measured in minutes using a stopwatch, and the total compensation (annual salary plus cost of benefits) in dollars of the staff person or persons performing the test are entered for each segment. Figure 1.A7 provides a flow diagram for cumulating labor costs. Figure 1.A8 is a labor-time accumulation worksheet that is used to outline specific labor tasks that are divided into the following times:

1. *Preanalytic time* includes the time to gather specimens from laboratory central receiving and time for all the steps up to the actual testing procedure, such as work-list gathering, startup, sample-cup preparation, preparation of daily quality controls, and standard for this instrument. Preanalytic time includes phlebotomy and centrifugation.

	INSTRUMENT NAME	COST/TEST	TEST/YEAR	COST/INSTRUMENT
Chemistry Instrument	1	$		$
	2			
	3			
	4			
	5			
Manual Tests	1			
	2			
	3			
TOTALS:	1. Instrument Costs			
	2. Cost Professional Salaries/ Administrative Salaries			
	3. Cost Blood/Components			N.A.
	4. Cost Transportation/Travel			
	5. Cost Fee Basis			
	6. Cost Operations/Administration			
TOTAL COST Operations/Laboratory/Year for Chemistry				$

FIGURE 1.A12. Worksheet for sectional analysis of chemistry costs based on equipment configuration/manual procedures. This format can be used for all laboratory sections. (Adapted from Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Services, 1989:130.)

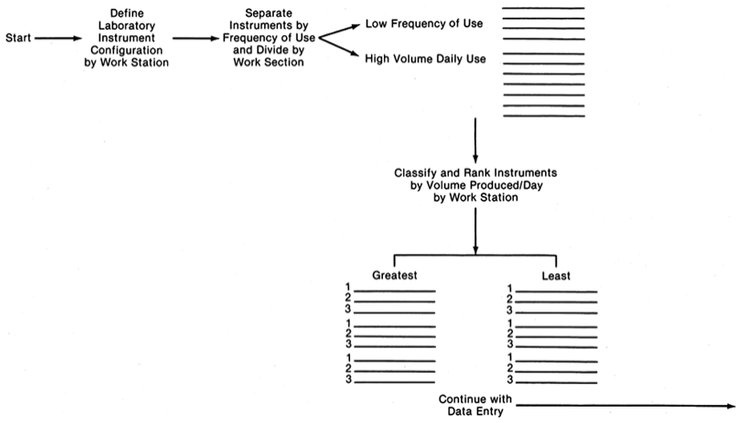


FIGURE 1.A13. Classification of laboratory's configuration by volume. Algorithm for identification of high-volume instruments. (Adapted from Travers EM. Managing costs in clinical laboratories. New York: McGraw-Hill Information Services, 1989:125.)

2. *Analytic time* includes time for the labor effort required to analyze the specimen or specimens and to perform all routine procedures up to reporting of results. It includes calculation(s) and checking but does not include test repeats. Detailed instructions on analytic time-study performance are given.
3. *Postanalytic time* includes time for the labor effort required to report results manually or to enter results into a computer system as well as the sorting, filing, and telephoning related to final reports. Routine, daily maintenance time, and shutdown times are also postanalytic efforts.

After the minutes of time for each task are entered in column D of Fig. 1.A8, they are divided into the categories noted above. The cumulative time for each category (preanalytic, analytic, and postanalytic) is then entered in section III of the ICAT Cost Accumulation Worksheet (Fig. 1.A1).

Formulas Used for Test Cost Determination There are six basic formulas that must be used to calculate material, labor, and indirect costs for test production (Fig. 1.A9). This information accumulated on the ICAT worksheet (see Fig 1.A1) is transferred to each of the formulas that appear in Fig. 1.A9.

When all the formulas have been solved, their totals are entered on the test cost determination summary sheet (Fig. 1.A10) and added to a subtotal. The subtotal is multiplied by the indirect cost percentage for the hospital or corporation to provide a cost per test total (Fig. 1.A11).

Section IV: Synthesis of Test Costs for Laboratory Instruments by Laboratory Section

When the total costs for each instrument-selected analysis in each laboratory section has been calculated, the manager can synthesize costs for all instruments in the section. The format appears in Fig. 1.A12. Laboratory managers should determine which instruments

account for the major workload (Fig. 1.A13) and then proceed to the next phase to perform sectional cost analysis.

The algorithm presented in Fig. 1.A13 allows the manager to apply the 80/20 rule and to select instruments and methods that account for the highest volume (and highest costs) of test production.

Phase V: Sectional Cost Analysis

The generic sectional analysis format is adaptable for all laboratories and includes entries for all major areas of laboratory instrumentation and manual methods. It is assumed that the user will first establish the cost for an instrument's operation using the ICAT and then enter these data into the appropriate category in Fig. 1.A12. A category is also included for entry of manually performed test costs for smaller laboratories that rely on the latter for a significant number of tests.

Once the cost per test for each instrument is established using the ICAT, the laboratory manager must then determine the raw count of tests and profiles per year that have been (or will be) performed on each instrument and multiply this by the cost per test. The final product for each instrument in each section is added together to produce the cost for the equipment configuration of the laboratory.

INSTRUMENT	A	B	C	H	
Example 1: General Function That Varies per Instrument					**Total Chemistry Tests**
Total Billable* Procedures	39,232	65,068	5,052	6,890	116,242
Percent of Total Chemistry Tests	27.9 ×	46.4 ×	3.6 ×	4.9 ×	
Phlebotomy Salaries† for Chemistry	$100,128 =	$100,128 =	$100,128 =	$100,128 =	
Cost of Phlebotomy per Instrument	$ 27,936	$ 46,459	$ 3,605	$ 4,906	
Example 2: General Function That Applies Equally to All Instruments/Sections					**Total Professional Salaries for Chemistry¶**
Pathologist's Salaries (Professional Fees)‡	12,500	12,500	12,500	12,500	50,000

* Test result actually billed to patient. Example: a chemistry profile with 18 determinations counts as one billable procedure, not 18 tests.

† Determined by the following ratio:

1. $\frac{\text{Total Billable Procedures/Chem}}{\text{Total Billable Procedures/Lab}} = \frac{175{,}999}{394{,}000} = .447 = 44.7\%$
2. 44.7% × Total Phlebotomy Salaries/Lab = .447 × $224,000 = $100,128

‡ Pathologist's time in this lab is allocated as follows:

50% Histology/Surgical Pathology	10% Hematology
10% Chemistry	10% Immunology
10% Microbiology	10% Blood Bank

¶ For Chemistry:

a. 10% × Total Prof. Fees/Lab = 10% × $625,000 = $62,500/Chem

b. 62,500/5 (instruments) = $12,500 Prof. Fees/Instrument

FIGURE 1.A14. Allocating costs for general functions among chemistry workstations. (Modified from Sharp JW. A cost accounting system targeted to DRGs. *Med Lab Observer* 1985;17:34–41.)

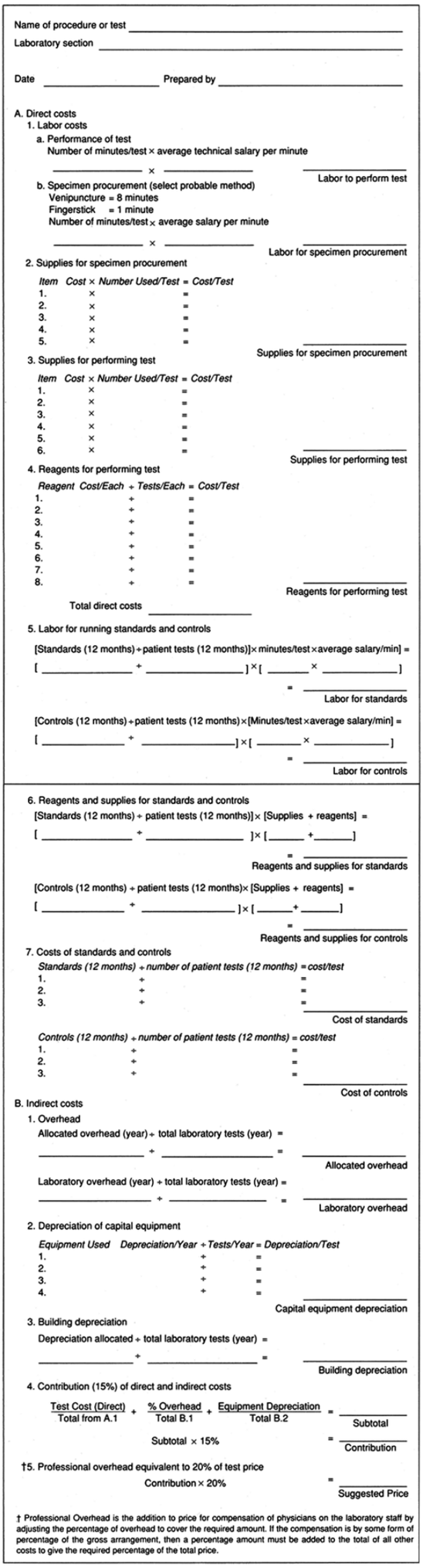
The format can also be used to cumulate other sectional major costs, e.g., professional salaries (professional fees), blood and component costs, transportation costs, fee-bases (send-out) costs, costs for operations, administration, travel, depending on the organizational structure and fiscal control points of the laboratory.

Many cost accounting techniques are concerned only with direct costs for technical and clerical labor and materials for test production, and others use an estimation cost allocation technique for each laboratory section to include nontechnical and professional functions. These functions are not usually assigned to a particular workstation, or even to a section, and they are part of the cost of test production, interpretation, and reporting. This approach is used to distribute the total costs of the laboratory section equally among the total number of workstations. When all costs have been allocated, the total cost of each type of test in the section can be calculated.

Allocating Costs Across Laboratory Sections for All Costs

If physicians, clinical scientists, and technical managerial and general employees such as secretaries and other clerical workers are not specifically assigned to a laboratory section, yet their work helps to produce or interpret tests, then an allocation schedule that spans all departments should be developed. For instance, Sharp uses allocation schedule formulas to assign the costs of nontechnical personnel, profession functions, and other miscellaneous functions to each production center. The formulas may differ for each group or item allocated, and there is no uniform set of allocation formulas that will apply to every laboratory. A good manager should be able to create accurate formulas for each production area and the department based on his or her familiarity with laboratory operations. The only requirement is that the formulas realistically reflect the section's or department's use of that function. For example, allocation of expenses for phlebotomy and for clerical, reception, professional, and other miscellaneous functions allocated to each workstation in chemistry is presented in Fig. 1.A14.

APPENDIX B: PRICE SETTING FORMAT



Name of procedure or test _____

Laboratory section _____

Date _____ Prepared by _____

A. Direct costs

1. Labor costs

a. Performance of test

Number of minutes/test × average technical salary per minute

_____ × _____ = _____ Labor to perform test

b. Specimen procurement (select probable method)

Venipuncture = 8 minutes

Fingerstick = 1 minute

Number of minutes/test × average salary per minute

_____ × _____ = _____ Labor for specimen procurement

2. Supplies for specimen procurement

Item	Cost	×	Number Used/Test	=	Cost/Test
1.		×		=	
2.		×		=	
3.		×		=	
4.		×		=	
5.		×		=	

_____ Supplies for specimen procurement

3. Supplies for performing test

Item	Cost	×	Number Used/Test	=	Cost/Test
1.		×		=	
2.		×		=	
3.		×		=	
4.		×		=	
5.		×		=	
6.		×		=	

_____ Supplies for performing test

4. Reagents for performing test

Reagent	Cost/Each	÷	Tests/Each	=	Cost/Test
1.		÷		=	
2.		÷		=	
3.		÷		=	
4.		÷		=	
5.		÷		=	
6.		÷		=	
7.		÷		=	
8.		÷		=	

_____ Reagents for performing test

Total direct costs _____

5. Labor for running standards and controls

[Standards (12 months) ÷ patient tests (12 months)] × minutes/test × average salary/min] =

[_____ ÷ _____] × [_____ × _____]

= _____ Labor for standards

[Controls (12 months) ÷ patient tests (12 months) × [Minutes/test × average salary/min] =

[_____ ÷ _____] × [_____ × _____]

= _____ Labor for controls

6. Reagents and supplies for standards and controls

[Standards (12 months) ÷ patient tests (12 months)] × [Supplies + reagents] =

[_____ ÷ _____] × [_____ + _____]

= _____ Reagents and supplies for standards

[Controls (12 months) ÷ patient tests (12 months) × [Supplies + reagents] =

[_____ ÷ _____] × [_____ + _____]

= _____ Reagents and supplies for controls

7. Costs of standards and controls

Standards (12 months)	÷	number of patient tests (12 months)	=	cost/test
1.	÷		=	
2.	÷		=	
3.	÷		=	

_____ Cost of standards

Controls (12 months)	÷	number of patient tests (12 months)	=	cost/test
1.	÷		=	
2.	÷		=	
3.	÷		=	

_____ Cost of controls

B. Indirect costs

1. Overhead

Allocated overhead (year) ÷ total laboratory tests (year) =

_____ ÷ _____ = _____ Allocated overhead

Laboratory overhead (year) ÷ total laboratory tests (year) =

_____ ÷ _____ = _____ Laboratory overhead

2. Depreciation of capital equipment

Equipment Used	Depreciation/Year	÷	Tests/Year	=	Depreciation/Test
1.		÷		=	
2.		÷		=	
3.		÷		=	
4.		÷		=	

_____ Capital equipment depreciation

3. Building depreciation

Depreciation allocated ÷ total laboratory tests (year) =

_____ ÷ _____ = _____ Building depreciation

4. Contribution (15%) of direct and indirect costs

$$\frac{\text{Test Cost (Direct)}}{\text{Total from A.1}} + \frac{\%\ \text{Overhead}}{\text{Total B.1}} + \frac{\text{Equipment Depreciation}}{\text{Total B.2}} = \text{Subtotal}$$

Subtotal × 15% = _____ Contribution

†5. Professional overhead equivalent to 20% of test price

Contribution × 20% = _____ Suggested Price

† Professional Overhead is the addition to price for compensation of physicians on the laboratory staff by adjusting the percentage of overhead to cover the required amount. If the compensation is by some form of percentage of the gross arrangement, then a percentage amount must be added to the total of all other costs to give the required percentage of the total price.

FIGURE 1.B1. An example of a format used for single test price setting using the full costing (total cost) method. (Modified from Perryman M. Cost accounting of test procedures. In: Karni KR, Viskochil KR, Amos PA, eds. *Clinical laboratory management*. Boston: Little, Brown, 1982:348–349.)

APPENDIX C: COMPARATIVE LABORATORY EQUIPMENT COST WORKSHEET

Billable Tests _____	Alternative A		Alternative B		Alternative C		Alternative D	
Consumables	Quantity per._____	Annual Cost	Quantity per._____	Annual Cost	Quantity per._____	Annual Cost	Quantity per._____	Annual Cost
FLUIDS								
Reagents	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Diluents	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Calibration Material	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Reference Fluids	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Linearity Standards	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Control Fluids	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Rinse/Wash Solutions	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Cleaning Solutions	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Deionized Water	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Analyzer Specific Solutions	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Other	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Subtotal	_____	$_____	_____	$_____	_____	$_____	_____	$_____
SUPPLIES								
Chemicals	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Gases	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Paper	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Labels	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Ribbons	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Water	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Cartridges, Filters, etc.	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Other	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Subtotal	_____	$_____	_____	$_____	_____	$_____	_____	$_____
PARTS								
Syringes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Tubing	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Membranes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Valves	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Reagent Caps	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Seals	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Reagent Vessels	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Cuvettes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Vials	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Lamps	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Needles	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Probes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Electrodes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Filters	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Detectors	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Reaction Chambers	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Other	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Subtotal	_____	$_____	_____	$_____	_____	$_____	_____	$_____
DISPOSABLES								
Sample Cups	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Sample Caps	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Pipette Tips	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Glass Tubes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Bar Code Labels	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Micro Inserts/Cups	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Cuvettes	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Other	_____	$_____	_____	$_____	_____	$_____	_____	$_____
Subtotals	_____	$_____	_____	$_____	_____	$_____	_____	$_____
TOTALS								
1. Material Cost per Billable Test	_____	$_____	_____	$_____	_____	$_____	_____	$_____
2. Labor Costs (see Appendix I)	N.A.	$_____	N.A.	$_____	N.A.	$_____	N.A.	$_____
3. Indirect Costs (see Appendix I)	N.A.	$_____	N.A.	$_____	N.A.	$_____	N.A.	$_____
TOTAL COST (Sum of 1, 2 & 3)		$_____		$_____		$_____		$_____

FIGURE 1.C1. Comparative laboratory equipment cost worksheet. (Adapted from Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:237–238.)

APPENDIX D: CAPITAL ACQUISITION PROCESS FOR LABORATORY EQUIPMENT

Depreciation Costs

1. Provide generic and trade name of equipment; provide operating description
2. Provide total capital acquisition cost
3. Provide cost of first-year service contract
4. Provide cost of service arrangement included in the total cost, which extends beyond the first year
5. Provide installation cost
6. Provide shipping costs
7. Provide training cost
8. Calculate supply and other costs included in the price of the equipment; which items will be consumed less than the equipment's depreciable life and/or less than the financing term
9. Subtotal items 3–8
10. Calculate value of cash payment discount
11. Subtotal items 9 and 10
12. Calculate the pure capital equipment value (item 11 minus line 2)

Depreciation Life and Timing

13. Assign the equipment's proposed depreciable life
14. Project equipment arrival date
15. Forecast date of equipment acceptance
16. Forecast date of first use of equipment for testing

Related Cash Flow Variables

17. Provide other costs from outside vendors/contractors not included in equipment price
    a. Remodeling costs (outside contractor)
    b. Moving costs (outside contractor)
    c. Other costs (detailed list)
18. Calculate undepreciated value of equipment being replaced on date of first testing use for patient care
19. Provide trade-in value, if any, on the equipment being replaced
20. Calculate the value of departmental depreciation expense recorded for the last time during the year prior to the year of new equipment acquisition
21. Provide manufacturer's payment terms
    a. Amounts to be paid
       1. At signing of purchase order
       2. At time of equipment production
    b. Payments to be financed through internal and/or external sources
       1. At time of delivery
       2. At time of acceptance
    c. Total to be internally/externally financed

Production Volume and Revenues

22. Calculate production forecast data
    a. Test procedures—expected annual average
    b. Relative Value (minutes/test)
    c. Total Relative Values (a × b)
23. Calculate average test market price
24. Forecast gross patient revenue (item 22a × item 23)
25. Determine cost payers
    a. Medicare
    b. Medicaid
    c. Other third-party payers
26. Bad debts and charity expenses
27. Convert production forecast into man-hours

$$\frac{(22b \times 22a)}{60}$$

28. Annualize gross patient revenue per full-time equivalent employee forecast

$$\frac{\text{Actual full-time employee man-hours}}{\text{Annualized FTE man-hours}} = \frac{1500 \text{ man-hours}}{2080 \text{ man-hours}}$$

Raw tests for full-time equivalent employee

$$\frac{\text{Raw tests produced}}{\text{Actual FTE}} = \frac{30{,}000 \text{ tests}}{.72} = 41{,}667 \text{ tests/FTE}$$

29. Calculate hospital planned rate of return (%)
30. Calculate department planned rate of return (%)
31. Calculate net present value
32. Calculate internal rate of return
33. Determine payback period

Financial Profile Summary

Equipment cost	$82,194
Depreciable life	5 years
Date of return	15%
Tests/year	30,000
Total relative values (weighted workload units)	90,000
Average price/test	$10
Estimated cost payer adjustments	
Medicare	25%
Medicaid	5%
Bad debts/charity care	5%
Forecast net income	$59,319
Forecast depreciation/year	$16,439
Annual cash flow (income−depreciation)	$75,758
Net present value	$157,923
Internal rate of return	74%
Payback period	460 days

FIGURE 1.D1. Recommended steps in a capital acquisition process for laboratory equipment. [Modified from Oszustowicz RJ. A capital equipment acquisition process. *J Healthcare Finan Manag Assoc* 1982;30:(April).]

APPENDIX E: CAUSES OF SOLUTIONS FOR LOW LABOR PRODUCTIVITY

	Condition Present In This Laboratory YES	NO
1. There is inadequate automation.	____	____
2. There is poor skill level.	____	____
3. Consistency of reporting CAP WLUs needs to be assessed.	____	____
4. There is poor work flow caused by poor facility design.	____	____
5. There is poor information flow; computerization is lacking or is a problem.	____	____
6. Infrequently ordered tests are performed in-house	____	____
7. There is inappropriate batching of tests (tests are run seven times per week when three times per week are adequate.	____	____
8. Tests are performed in lab when they should be sent out.	____	____
9. There is inappropriate service level demands by physician staff.	____	____
10. There is a high percentage of stat testing.	____	____
11. There are inappropriate test turnaround time standards.	____	____
12. There are inappropriate quality control programs, resulting in test repeats.	____	____
13. There is excessive machine breakdown, resulting in repeat testing.	____	____
14. There is poor distribution of work between shifts, departments, and benches (could more tests be done during night shift?).	____	____
15. There is excessive paid time off due to staff longevity.	____	____
16. There are dollars spent on medical/technical schools or training programs that are not properly assigned to that function.	____	____
17. There is excessive orientation due to staff turnover or poor hiring practices.	____	____
18. The specialization and generalization of technical staff is not properly balanced.	____	____
19. There is an incorrect assessment of workload units.	____	____
20. There is unnecessary quality conytrolling.	____	____
Note: A predominance of "Yes" answers implies low labor productivity may exist in this laboratory; recommended solutions appear in Table 9–4.	TOTAL SCORE	

FIGURE 1.E1. Self-analysis checklist for causes of low labor productivity. (From American Hospital Association. *Assessing laboratory operations*. Chicago: Clinical Services Division, 1987, with permission.)

TABLE E.1. SOLUTIONS FOR LOW LABOR PRODUCTIVITY

1. Evaluate lab consolidation opportunities.
2. Increase work volume and batch sizes by marketing services.
3. Shift routine work to other shifts.
4. Make sure there is adequate space and design so that work flows smoothly.
5. Investigate flexible vs. fixed staffing
6. Increase automation using low cost/test, nonlabor intensive equipment.
7. Computerize laboratory.
8. Increase personnel expectations and performance with consistent hiring, training, and reward practices.
9. Improve relationships with medical staff, i.e., strengthen pathologist's role in developing appropriate service levels and control inappropriate stat testing orders.
10. Investigate increasing test batching, relax turnaround time requirements.
11. Combine sections and/or consolidate workstations for maximum efficiency.
12. Reduce any duplication of services.
13. Cross-train staff–generalists vs. specialists.
14. Evaluate use of satellite laboratories.
15. Improve preventive maintenance and troubleshooting skills of all staff.
16. Send out infrequently performed tests when staff efficiency remains low.
17. Investigate staff reduction.
18. Motivate staff to give 100% when at work.
19. Evaluate organizational chart and restructure if appropriate, i.e., flatten staffing if too "top heavy" in management personnel.
20. Evaluate appropriateness of research and development time.
21. Improve accuracy of workload.
22. Determine point of maximum efficiency for major automated equipment items.
23. Remove labor-intensive equipment.
24. Avoid highly specialized staff and use cross-trained staff.
25. Determine cost of manual vs. automated tests.
26. Determine "make vs. buy" cost.
27. Avoid too many high-salaried staff.

From Travers EM. *Managing costs in clinical laboratories*. New York: McGraw-Hill Information Systems, 1989:212, with permission.

2

Business Management of the Clinical Laboratory

Eleanor M. Travers

- THE VALUE OF MEDICAL DIAGNOSTIC SPECIALITIES TO PATIENT CARE
- THE COLLISION OF MEDICAL CULTURE CHANGES AND MDT INNOVATION
- THE IMPACT OF MDT
- PRESSURE TO PERFORM MORE SPECIFIC/SENSITIVE MEDICAL DIAGNOSTIC TESTS WITHOUT RECOVERY OF OPERATING COSTS
- PARADOX OF ASTOUNDING TECHNOLOGICAL EXPERTISE (HIGH COST) IN THE ERA OF SYSTEMATIC PRICE CONTROLS BY THIRD-PARTY PAYERS AND MANAGED CARE ORGANIZATIONS
- FLEXIBILITY, ADAPTABILITY, AND ALTERNATIVE OPTIONS
- MERGER AND ACQUISITION SHOCK
- MANAGEMENT COPING SKILLS
- LEGISLATIVE FACTORS
- GOVERNMENT'S PUNITIVE INDIFFERENCE
- THE COST OF COMPLIANCE
- BALANCING CARE MANAGEMENT WITH BUSINESS MANAGEMENT
- PROMOTING THE VALUE OF THE LABORATORY
- PATIENT OUTCOME MANAGEMENT
- PUBLISHING EVIDENCE-BASED DIAGNOSTIC REASONING PATTERNS FOR PROVIDERS
- DIAGNOSTIC PATTERN RECOGNITION OF SIGNIFICANT ABNORMALS AND NORMALS
- PUBLISHING EVIDENCE-BASED SENSITIVITY AND SPECIFICITY DATA FOR PROVIDERS
- DISEASE MANAGEMENT
- DEFINITION OF “PROVIDER”
- WHAT IS NECESSARY TO PERFORM DISEASE MANAGEMENT PROPERLY?
- LEADERSHIP AND BUSINESS RELATIONSHIPS
- INFORMATION MANAGEMENT
- TOTAL COST MANAGEMENT
- MANAGEMENT RESPONSIBILITIES AND CONTRIBUTION/PROFIT MARGIN
- UTILIZATION MANAGEMENT
- PRODUCTIVITY MANAGEMENT
- REVENUE MANAGEMENT
- MARGIN
- PRICE (CHARGE, RATE, FEE)
- RISK MANAGEMENT
- REIMBURSEMENT MANAGEMENT
- SUMMARY AND CONCLUSION

THE VALUE OF MEDICAL DIAGNOSTIC SPECIALITIES TO PATIENT CARE

Part of "2 - Business Management of the Clinical Laboratory"

Overview

The American health care system may have forgotten, never knew, or never appreciated the value of the clinical diagnostic laboratory to medical diagnosis and treatment. A staggering array of medical diagnostic information is within easy reach of any clinical provider in an era in which the disappearance of the physical examination and other subjective medical decision-making tools signals the importance of quantitation and measurement (1). Medical diagnostic technologies (MDT) are now the method of choice for providers of medical care because they locate and define problems for which a physical examination and history can only provide clues. If an easily available medical diagnostic technology is an alternative, in this generation and the future, it will be used instead of the time-honored but time-consuming, imprecise hands-on traditional bedside diagnostic method. There is now an entire generation of medical care providers who may not be very comfortable with the “bedside” part (2).

Specifically, MDT as objective, factual tools are used more often by providers than subjective, judgmental medical decision-making tools to confirm clinical impressions and establish definitive diagnoses (Fig. 2.1). Combined with the positive attributes of accuracy and speed in providing diagnostic information, their unchecked, repetitive use is known to increase health care costs through complex and interrelated technical, social, ethical, and legal mechanisms (3,4).

In addition, several interacting cultural changes in the practice of medicine in the United States have propelled the medical diagnostic laboratory to a prominent, cutting-edge role never dreamed of before automation was introduced and forever changed the way that physicians make medical diagnostic decisions. These interacting forces are ambulatory and preventive care, primary care, and the provider's reliance on objective data versus subjective impressions to make medical decisions more rapidly. The relationship of each factor to the emergence of the clinical diagnostic laboratory and the provider of medical care is discussed in the sections that follow.

Ambulatory and Preventive Care Causes Greater Need for Tests

In the late 1990s, providers of medical care in American health care organizations (HCOs), highly dependent on medical diagnostic tests, found it necessary to shift the focus of their practices from centralized inpatient to decentralized ambulatory care (outpatient) operations. This was done primarily to satisfy the need of third-party payers to reduce high insurance risks caused by costly inpatient care. Payers also created a dilemma for managers of medical diagnostic support specialties, forced to follow providers and decentralize test performance and/or provide test results and services for providers in near-patient settings to improve the efficiency of the delivery of patient care.

Providing diagnostic services for smaller, sicker inpatient populations in large, centralized HCOs with fixed, complex, and costly medical diagnostic laboratories caused higher operating costs and logistical nightmares. More technical staff was needed to perform tests at the site of patient care. Workload doubled in centralized sites because of the need to analyze both inpatient and outpatient specimens, and new federal legislation introduced in 1997 (Balanced Budget Act of 1997) created a mandate for increased preventive testing for common chronic diseases of the elderly.

Primary Care Emphasis

In addition, the deemphasis of medical specialist and subspecialist intervention in patient care requiring primary care physicians to approve and supervise a patient's care added new meaning to the use of medical diagnostic tests. Primary care physicians in outpatient care settings needed more tests on a repetitive basis to assure good care for outpatients who were not authorized to receive care from a specialist.

The value of the clinical diagnostic laboratory also increased because the Balanced Budget Act of 1997 mandated new preventive tests for the most common diseases found in the elderly population such as diabetes, prostate cancer, and breast cancer.

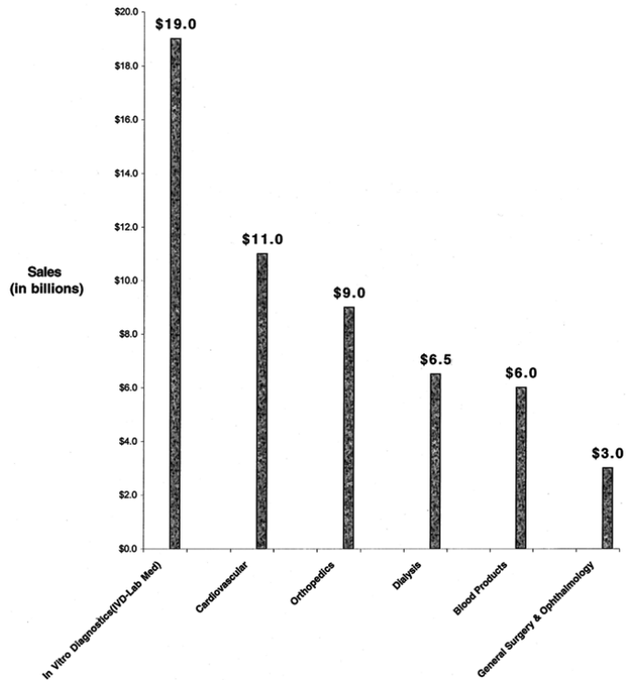


FIGURE 2.1. Medical technology sales by segment: 1998. Laboratory Medicine and other clinical segments that use In-Vitro Diagnostics (IVD) constituted the largest volume of medical technology sales in 1998. Data source: U.S. Bancorp Piper Jaffray Research, Minneapolis, Minnesota.

Provider's Reliance on Objective Data (Technologies) Versus Clinical Findings

It is expected that providers seeing patients in any health care setting must increase their efficiency and productivity. The average time allowed for managed-care providers to evaluate, diagnose, and write up prescriptions and medical records is 20 minutes or less. In many cases, it is not wise or feasible for providers to wait for medical diagnostic test results to be delivered from distant sites, and medical decisions have to be made promptly with simple but reliable point-of-care testing techniques.

It is this reliance on objective and accurate data that makes a provider opt for faster and more complete diagnostic test menus that deliver results promptly to the point of care.

THE COLLISION OF MEDICAL CULTURE CHANGES AND MDT INNOVATION

Part of "2 - Business Management of the Clinical Laboratory"

The interaction of these cultural shifts in medical care delivery increase the value of the medical diagnostic laboratory in the American health care system at the same time that the introduction of new and dynamic diagnostic technologies through research and development offers great hope for better patient care. However, the simultaneous occurrence and entry speed of both these remarkable circumstances into medical care causes considerable stress for managers of medical diagnostic production facilities faced with the responsibility of adding value and reducing costs for his/her organization.

THE IMPACT OF MDT

Part of "2 - Business Management of the Clinical Laboratory"

Despite the need for and demonstrated value of medical diagnostic tests and services to health improvement (5,6), the altruism of providers who assume that ordering a series of medical diagnostic tests will benefit the patient and the federal government's zeal to provide preventive care for the elderly, the steep operating costs of providing medical diagnostic tests as a necessary step in the medical decision-making process is not always considered by those who order them.

The resultant impact of these forces ends squarely at the point of production of medical diagnostic tests causing increased costs for labor and materials to produce, control quality and increase the speed and availability of medical information to providers. Further, the value of testing to providers and investors within the diagnostic testing industry has grown to phenomenal proportions.

Marketing Dynamics

Within the vast universe of medical technology and services, MDT

- constitute the single, largest market segment within the medical technology industry (7,8) (Fig. 2.1),
- comprise nearly 10% of hospital costs (9) (Table 2.1),
- influence the ordering of other more expensive MDT (e.g., the average cost of a diagnostic laboratory test is $0.03 to $0.05, but the use of its result by providers influences $0.70 to $0.80 of subsequent health care spending) (10,11) (Fig. 2.2),

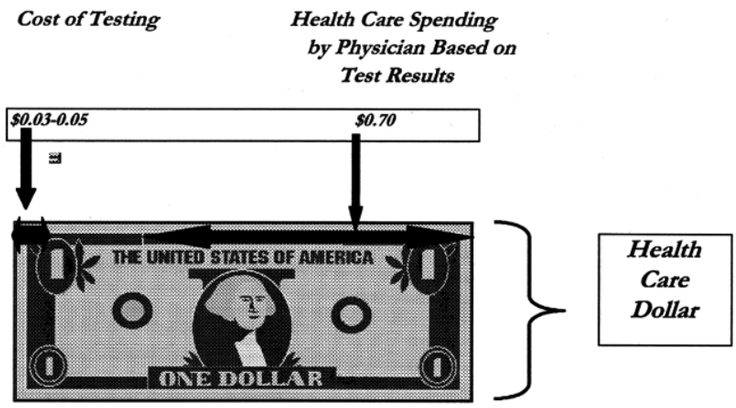


FIGURE 2.2. The impact of a diagnostic laboratory test. Source of Data: Paxton A. Gazing at Tomorrow: CEO Questions, *CAP Today*, 1997:11:25-26.

- are responsible for 80% of patient care data models, based on information developed from interpretation of *in vitro* diagnostics, also known as medical laboratory tests (10).

PRESSURE TO PERFORM MORE SPECIFIC/SENSITIVE MEDICAL DIAGNOSTIC TESTS WITHOUT RECOVERY OF OPERATING COSTS

Part of "2 - Business Management of the Clinical Laboratory"

Urged on by good medical practice and ethics, providers continue to order both existing and newly US Food and Drug Administration (FDA)–approved medical diagnostic tests with greater specificity and sensitivity than existing tests. However, the producer of the test must incur the cost of its production without assurance that costs will be recovered from third-party payers.

Because, in the era of managed care, the directors of insurance and other companies authorized to make payments to HCOs and other providers value greater patient turnaround time and reduced financial risk, the ability of providers to rapidly diagnose and treat patients has become more important in the United States than completeness and continuity of care. To meet this challenge, the research arm of the medical diagnostic industry and other research entities are producing a vast array of new, more specific, and more sensitive products, a costly process when there is no assurance that the research and development costs will be recovered through future sales or reimbursement by third-party payers.

Delayed Reimbursement for New Tests

The reason for this concern about financial losses for tests performed without assured reimbursement is that new medical diagnostic testing products must first be approved by the FDA before they can be used for patient care, a 1- to 2-year process. Other already FDA-approved products must still wait for recognition by the American Medical Association's (AMA) Current Procedural Code (CPT-4) committees, another 1- to 2-year process, before a medical diagnostic laboratory can legally file a claim and bill Medicare or another third-party payer for reimbursement.

The dilemma for managers of medical diagnostic testing laboratories is their inability to inform providers that reimbursement is not authorized for certain tests that do not have approved codes or for tests that have been deleted from the AMA's approved CPT-4 list. Further, changes in the approved codes occur each January, and it takes nearly 2 years for an approved test's code to be printed in the AMA's official CPT-4 Code Book, published on an annual basis (12). Therefore, patient care may be well served because a provider received helpful diagnostic information, but the clinical laboratory manager never recovers operating costs if providers continue to order medical diagnostic tests that have not been assigned an approved procedure code (CPT-4).

PARADOX OF ASTOUNDING TECHNOLOGICAL EXPERTISE (HIGH COST) IN THE ERA OF SYSTEMATIC PRICE CONTROLS BY THIRD-PARTY PAYERS AND MANAGED CARE ORGANIZATIONS

Part of "2 - Business Management of the Clinical Laboratory"

Added to the problem of lost operating dollars is the control of reimbursement rates by Medicare, third-party payers, and managed care organizations (MCOs). This problem has escalated to the point of a formal review of Medicare laboratory test payment practices in 2000 by the National Academy of Sciences.

FLEXIBILITY, ADAPTABILITY, AND ALTERNATIVE OPTIONS

Part of "2 - Business Management of the Clinical Laboratory"

In combat, when chaos surrounds even the well-trained and seasoned soldier, the leader must have more than one alternative strategy to survive during the noise and confusion of battle. Leaders of diagnostic laboratories in similar chaotic, confusing periods must be flexible and adaptable and have more than one plan to survive, using the following recommendations.

TABLE 2.1. Distribution of Hospital Costs for Diagnostic versus Therapeutic Medical Technologies

Clinical Diagnostic Costs (Medical Diagnostic Technologies)	% of Total Clinical Costs	% of All Hospital Costs	Clinical Therapeutic Costs (Medical Therapeutic Technologies)	% of Total Clinical Costs	% of All Hospital Costs
Radiology	6.7	3.88	Radiology	2.1	0.12
Nuclear medicine (isotope)	0.3	0.16			
Clinical laboratory	8.7	5.05			
Transfusion medicine (blood bank)	0.3	0.17	Packed red blood cells and whole blood	2.1	0.12
ECG (cardiology)	0.9	0.51			
EEG (neurology)	0.1	0.05			
Endoscopy (internal medicine)	0.02	0.01			
Laparoscopy (ambulatory surgery)			Intravenous therapy	0.26	0.15
			Respiratory therapy	2.5	1.43
			Physical therapy	2.5	1.43
			Occupational therapy	0.6	0.35
			Speech pathology	0.2	0.13
			Renal dialysis	0.3	0.17
			Home dialysis program	0.02	0.01
			Operating room	6.2	3.57
Total diagnostic costs		9.83%	Total therapeutic costs		7.48%

Data source: 1994 Medicare cost reports (Worksheet A).
N = 6,235 hospitals. ECG, electrocardiogram; EEG, electroencephalogram.
From Woolhandler S, Himmelstein D. Costs of care and administration at for profit and other hospitals in the United States. *N Engl J Med* 1997; 336:769, with permission.

"The Mission No Longer Drives the Costs"

The altruistic mission of patient care is no longer in effect, except in a very few privately endowed, wealthy HCOs—institutions that can care for patients without having worries about meeting expenses. In all other HCOs in which medical diagnostic laboratories are a vital and indispensable component of the patient care, the bottom line is what matters.

Reemergence of the Hospital Laboratory as a Competitor

Although the future role of the hospital-based diagnostic laboratory was undefined and tenuous in the early 1990s, it became a major competitor in the diagnostic testing marketplace in the late 1990s. The massive shift from independent to consolidated hospital systems during this era created larger core laboratories within large health care institutions for the purpose of keeping more tests on-site (13). One of the reasons for this significant shift is the expertise of hospital-based professionals and the need to interface on a timely basis with clinical staff who require the expert assistance of laboratory physicians and professionals. A second reason is the steady erosion in the median price that U.S. hospitals pay for tests sent to outside reference laboratories. In 1996, the median price was $28.73, and in 1998, it was $23.19 (13,14). A third reason is the evolutionary change seen with tests introduced in the mid-1990s, formerly deemed esoteric and complex, now becoming routine tests. For example, polymerase chain reaction testing was a cutting-edge procedure and has migrated into the category of routine testing as a result of increasing hospital integration and consolidation.

Further, the percentage of tests sent out by hospitals is decreasing, while the average complexity of each referred test is increasing (as estimated by the relative value units per CPT code). Portugal estimates that 20% to 40% of the referred tests sent to other laboratories by community hospitals could be performed in-house at lower costs (15).

Another factor that should increase the likelihood of more test performance in hospital laboratories in the future is the 1999 proposal by the Health Care Financing Administration (HCFA) to eliminate the practice of global billing to Medicare from free-standing laboratory sites for hospital inpatients. This portends a doubling of billing costs for independent laboratories because they must make arrangements with the hospital to be paid for the technical components of the services performed for hospital inpatients (16).

Integrated Delivery Systems, decentralization and networking

Laboratory directors and managers in both the public and privates sectors need a strategy to continue to confront challenges associated with the need to contain costs and provide greater value. One strategy is to build and manage integrated delivery systems, also known as organized delivery systems. These terms are used to denote systems that are trying to organize to achieve a greater degree of coordinated care. Integration is needed to appropriately coordinate the laboratory's operating units (the end point at which tests and services are delivered). Organization of the network of operating units is necessary to achieve a coordinated continuum of tests and services to a clinical population (17,18). Integrations are further subdivided into clinical and functional integrations. Clinical integration is the degree to which patient care services are coordinated across people, functions, activities, processes, and operating units to maximize the clinical care delivered to patients. Functional integration is the degree to which key support functions and activities, such as financial management, human resources, information systems, strategic planning, and continuous quality improvement, are coordinated across the spectrum of operating units.

The goal is to achieve the greatest possible value for patients (high quality of care) and the organization with lower costs, greater efficiency, and productivity. Although no single organizational structure exists for an integrated/organized delivery system, some key elements are necessary for its development Table 2.2.

TABLE 2.2. KEY ELEMENTS NECESSARY FOR DEVELOPMENT OF AN INTEGRATED DELIVERY SYSTEM (IDS)

An IDS requires these qualities to allow it to be clinically and fiscally responsible for the health of the defined population:

- The organization must be patient centered.
- Physicians and nonphysician providers who order tests and service must be directly involved.
- The hub of the IDS needs to be the whole system, not one site.
- Information systems must link each operating site.
- The organization must be able to assess the needs of the population it serves.
- The system must adapt quickly when conditions warrant change.
- The system must be based on a cross-function/cross-service line structure.
- Those affiliated with the system must be multiskilled.
- The organization's focus needs to be on the continuum of care, not in the acute care hospital.
- The system must be able to improve continuously.
- Clinical decision making must be based on the clinical staff's approved evidence-based protocols.
- Incentives and compensation must reward appropriate behavior.

Creative directors and managers often experience barriers to change and reorganization, including:

- lack of understanding of the new health care environment
- lack of understanding of the merits of integration
- fear of losing control after an integration
- continued focus on the central operating unit as most important
- lack of commitment in the operating units to the goals for the integrated delivery system
- loss of geographic concentration of operating units (distance between units precludes good communication and cooperation)
- lack of trained staff
- inability to identify the key elements required to integrate (Table 2.2)
- inadequate information systems
- lack of demographic information on the population served and their specific customer needs

Integrated delivery systems require decentralization and networking to flourish—two difficult tasks for any director or manager who has to face the reluctance of older, inflexible staff members to move a centralized workstation or implement new workstations at distant sites that require daily or more frequent travel. Because the clinical staff treating the patient population in a decentralized or remote site determines the timing and workload type and volume for the laboratory, a prudent director or manager will first assess the needs of the clinical staff before deciding to move an established workstation.

Once a decision is made to integrate and decentralize, the barriers to change noted previously should be reviewed and an education program for the involved staff members should be developed. At the education sessions, the leaders of the organization should be present to explain why it is necessary and how the staff will be affected. The most valuable asset of any director or manager is to realize that some, but not all, of the staff will have a desire to participate in the decision-making and planning process. Those who buy into the plan must be encouraged, supported, and rewarded for their ideas and participation as the integration progresses and after it is implemented.

MERGER AND ACQUISITION SHOCK

Part of "2 - Business Management of the Clinical Laboratory"

Unlike integrations, in which the parent organization remains intact, many laboratory directors and managers have had to face a different set of crises when mergers and acquisitions were effected by their organizations. A merger is a fusion or absorption of one entity into another, usually where one of the entities ends up with less status than before the merger. The less important entity ceases to have an independent existence. An acquisition is the takeover of one entity by another; however, both parties retain their legal existence after the transaction (19).

MANAGEMENT COPING SKILLS

Part of "2 - Business Management of the Clinical Laboratory"

Change is a natural and inevitable part of organizational life. Whether driven to reactions to external market demands or regulations (e.g., managed care, the federal government) or internal desires for continuous improvement in patient care and financial success, change is constant.

Regardless of whatever corporate or legal transaction takes place to integrate, merge, or acquire a HCO, of which the clinical diagnostic laboratory is an integral part, the laboratory director and his/her managers must be able to cope with change and continue to lead the staff to assure that patient care does not suffer in the aftermath. In fact, laboratory directors, managers, and even the scientific, medical, and administrative staff are so frequently required to be involved in matters involving integrations, mergers, acquisitions, professional and financial survival that it is difficult to realize that they are still the organization's only experts in providing pathology and laboratory services. They need to fill the dual role of versatile and perceptive leaders at the same time they are needed for their intellectual skills and professional knowledge.

Key Concerns for Managing Transitions

Figure 2.3 illustrates a “triangulation” of the conditional states that directors and managers face when the organization's legal status changes. Note that there is an equilibrium and bidirectional flow between the “present state” condition (Step 1) and the “transition state” (Step 2), as well as between the transition state (Step 2) and the “new or desired state” (Step 3). This is in contrast to a static state between the “present state” and the “new or desired state.”

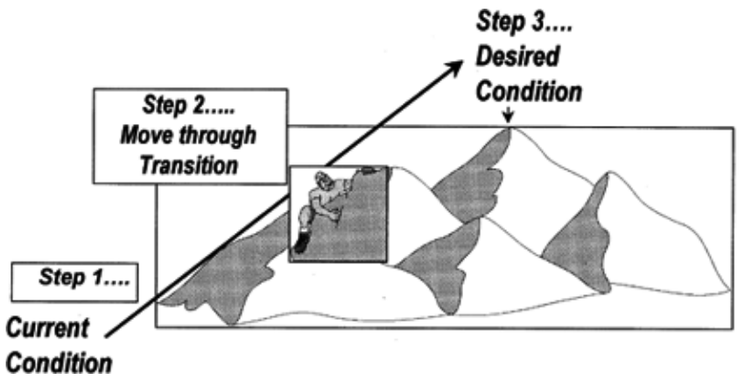


FIGURE 2.3. Key concerns for managing transitions depicted in the diagram are how to assist people effectively in letting go of the present ways of doing things; how to direct and manage people effectively in their movement through the transition period; how to provide the necessary support for people to accept, adopt, and execute new ways of doing tasks. (From Costello SJ, *Managing change in the workplace*. Burr Ridge, IL: Richard D. Irwin, 1994, with permission.)

The key concerns for managing transitions should be focused on the following (20):

- how to assist effectively the staff in letting go of the current ways of doing things
- how to direct and manage people effectively in their movement through the transition period
- how to provide the necessary support for people to accept, adopt, and execute new ways of doing tasks

Technimanagement

Unlike the management culture and environment of the traditional business or commercial organization, the director and managers of the clinical diagnostic laboratory are concerned with mastering the principles of managing human relationships to improve the productivity of a complex scientific/technical/medical organization. The traditional business organization sets its goals as production goals with product(s) produced at a competitive cost and quality. The traditional business manager relies on methods to get others to do what the manager wants (manipulation). Although this concept is now a common goal in the commercial, for-profit laboratory environment, there is a different perception in the noncommercial clinical diagnostic laboratory health care setting. Here, the treatment of illness and other unpredictable

events make the traditional production-process management style unworkable. The major difference between technimanagement and traditional management is primarily one of control. A technidirector or technimanager will see the organization as having a life of its own, with the organization responding to the needs of the individual.

Basically, the increased productivity and efficiency of persons employed in the scientific/technical/medical organization require a major transition from the old to the new, wherein the entire management structure itself undergoes constant and adaptive improvement (21). This is mentally and physically hard work for the director and manager, requiring constant attention to the internal culture and needs of the staff members. Basically, directors and managers need to use a mechanism that works to cope with the fear and anxiety caused by change. The director's or manager's skill is needed to establish the means and environment where each staff member's personal goals are consistent with those of the organization (21). When goals are not consistent, staff will resign, ask for a transfer to a different part of the organization, or make their discontent obvious.

Control of the organization's future success is not gained by controlling every person in the organization. It is gained by allowing the staff who wish to do so the opportunities to express their personal goals, blend them as a team with their co-workers' goals, and apply them as a group that supports and buys into the concept that working together improves value for the patient, offers a secure future for the individual (job security), and allows the organization to be financially successful.

Success in managing the complex scientific/technical/medical organization lies primarily with the director's and manager's understanding of a few basic management principles that generally hold, while recognizing their limitations (21). Management cannot be forced into a formula or equation, and those who try it do not succeed. The reason for this is that people, not products, are center of our success. People are not as simple as machines and procedures—they are totally different from each other, and they change constantly to meet their own internal needs and external threats.

LEGISLATIVE FACTORS

Part of "2 - Business Management of the Clinical Laboratory"

The Balanced Budget Act of 1997

The Balanced Budget Act of 1997 enacted the most significant changes to the Medicare and Medicaid programs since their inception in the 1960s. Many negative implications and operational changes for laboratories resulted from this legislation, including issues of implementing medical necessity, incurring reimbursement decreases, and a proposal for competitive bidding for tests and services provided to Medicare and Medicaid beneficiaries.

While increasing the laboratory's accountability for record keeping and documentation, it also increased the likelihood of major increases in workload by mandating improved screening for the most common diseases found in the U.S. population (e.g., diabetes, cervical cancer, prostate cancer, breast cancer) (Table 2.3). As a direct result of this legislation's implementation, laboratory directors and managers anticipated higher operating

TABLE 2.3. MEDICARE: BUDGET RECONCILIATION ACTION OF THE 105TH CONGRESS

Area Affected	Description	Implementation Date
Medicare-covered preventive services		
Mammography screening	Covered yearly for women age 40+. Part B deductible is waived for this service.	January 1, 1998
Screening pap smear and pelvic examination (including clinical breast examination)	Covered every 3 years. For women at high risk of developing cervical or vaginal cancer, coverage expanded to every year.	January 1, 1998
Prostate screening	For men age 50+, annual test includes one or both of the following: digital rectal examination and prostate-specific antigen blood test.	January 1, 1998
Additional provisions for prostate screening	Secretary of Health & Human Services (HHS) can deem other procedures screening appropriate.	2003
Colorectal screening	Coverage provided for the following procedures; fecal-occult blood test, flexible sigmoidoscopy, colonoscopy for high-risk individuals, and other tests or procedures and modifications to tests and procedures with and at a frequency the Secretary of HHS determines appropriate after consultations with appropriate organizations.	January 1, 1998
Other colorectal screening coverage	Secretary of HHS is required to publish a determination on the coverage of a screening barium enema.	Within 90 days of enactment
Diabetes self-management training	Medicare coverage will include diabetes outpatient self-management educational and training services, but only if physician managing diabetic's condition certifies that services are needed under a comprehensive plan of care to provide the individual with necessary skills to manage his/her condition.	July 1, 1998

This is an overview of some of the changes to Medicare as a result of the 1997 Balanced Budget Act. Listed are the changes that directly affect laboratory workload.

costs to pay for the provider's use of preventive and screening MDT. This is likely because of (a) more intense efforts to establish diagnoses in outpatient, rather than inpatient, settings and (b) more emphasis on preventive health care by Congress. At the same time, HCFA (and private-sector third-party payers) will be paying a much lower percentage of the actual cost for a laboratory to perform these mandated tests (Fig. 2.4).

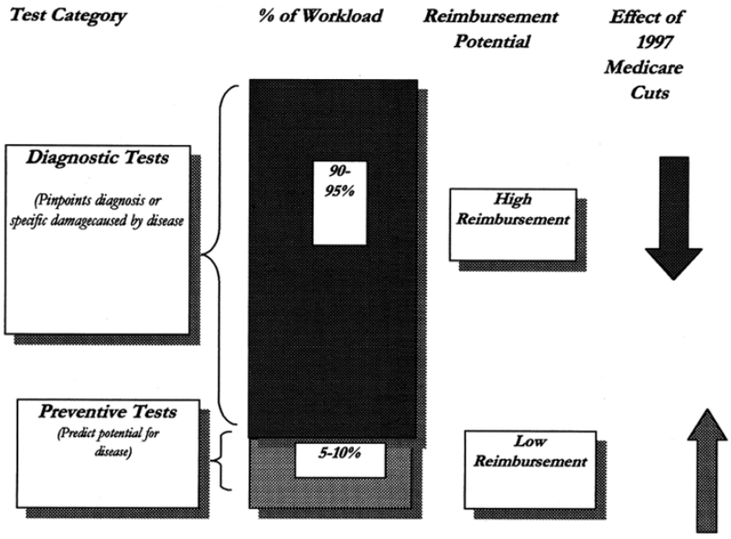


FIGURE 2.4. Potential for Medicare reimbursement for diagnostic versus preventive tests. (This figure provided as a courtesy by Dr. E.M. Travers, Annapolis, Maryland.)

GOVERNMENT'S PUNITIVE INDIFFERENCE

Part of "2 - Business Management of the Clinical Laboratory"

After the Clinical Laboratory Improvement Amendments of 1988 (CLIA ‘88) affected the entire laboratory world in 1993 with its final regulations designed to perform more quality measurement, follow new personnel regulations, and regulate the type of laboratory tests performed in each patient care setting based on complexity levels. Laboratorians reengineered, regrouped, and redesigned their operations—facing certain extinction by loss of licensure if they failed to comply.

Five years later, a massive change occurred in the organizational structure and operation of laboratories within HCOs. Costs and job tasks—not only in response to CLIA ‘88, but to dictatorial policies of MCOs requiring compliance with their contracts—changed radically and added new variable (labor and materials) costs to the cost of each test.

The quality and timeliness of patient care were also affected at the point of care, especially in clinics, group practices, and physicians' offices, where laboratories were closed in response to strict federal CLIA ‘88 and Occupational Safety and Health Administration (OSHA) regulations and the mandates of health management organizations (HMOs) that required laboratory tests and services on their beneficiaries to be sent only to their own contractees—large reference laboratories contracting with the MCO.

These and other burdensome external influences, however, make it difficult for laboratorians to remember that improved patient care was the original reason to provide high-quality, timely test results. These influences included regulatory issues, training, oversight, cost, fraud prevention, and reimbursement.

At the top of the list is the impact of medical necessity legislation, which became effective in 1996.

After the 1993 implementation of CLIA ‘88 and OSHA regulations, the federal HCFA issued instructions to Medicare intermediaries in 1997 intended to reduce government payments for unnecessary laboratory tests. Created by the need to quell fraudulent activities with Medicare reimbursement for laboratory tests on the part of a few commercial and corporate organizations, these necessary but intrusive, costly regulatory measures created a true economic dilemma for hospital and other laboratory administrators.

Fraud prevention is laudable and necessary to avoid misuse of taxpayers' dollars; however, Medical Necessity legislation, instituted in 1996, is causing massive additional administrative and oversight responsibility and increased labor costs for a laboratory's monitoring and verification of the physician's narrative

statement, medical documentation accuracy, and conformance with both the ICD-9 and CPT-4 code books and/or the actual diagnosis codes required before a claim can be made for reimbursement of a diagnostic test or service.

Because reimbursement and capped rates for capitated contracts have already lowered the laboratories' (or parent HCOs) incoming revenue streams, Medical Necessity regulations add more costs with hidden labor and consumables costs for a test or service. The hidden costs are required to account for the valuable time spent by technical personnel and additional administrative/clerical staff who must screen, verify, and act as the final clerical reviewer of the physician's test order (Table 2.4).

TABLE 2.4. MATRIX FOR ESTIMATING THE COST IMPACT OF FEDERAL REGULATORY, COMPLIANCE, INSPECTION, AND SAFETY REQUIREMENTS

	Increased Cost Impact—Requires Additional													
	Labor Effort					Nonlabor								
Federal Requirement	Administration	Supervisory	Technical	Clerical	M.D.	New ADP	New SW	Supplies	Training/Education	Sales	Marketing	Legal	Consult	Potential Revenue Losses
Balanced Budget Act 1997	x			x						x	x	x		Medicare may choose to contract with a competing “center of excellence”
Competitive Bidding														Hospitals lose 10% of HCFA funds by 1999
Medical necessity	x	x	x	x	x	x	x	x	x			x		“No diagnosis, no dollars” “No correct code, no reimbursement”
Operation Restore Trust	x			x	x	x	x	x	x			x	x	Stiff monetary penalties for coding and billing errors
Physicians in Teaching Hospitals (PATH) Audit	x			x	x				x			x	x	Residents not allowed to sign out cases; attending M.D. must sign out and verify
Safety (OSHA & Blood-Borne-Pathology) Inspection/accreditation	x		x	x				x	x					
JCAHO	x	x	x	x				x	x					
CLIA	x	x	x	x				x	x					CLIA failures may close lab/section
CAP/AABB/FDA	x	x	x	x				x	x					
NIDA	x	x	x	x				x	x					
NRC	x	x	x	x				x	x					
States	x	x	x	x				x	x					

This table is provided as a courtesy by Dr. E.M. Travers, Annapolis, Maryland.

With a seemingly punitive and indifferent attitude, the HCFA administrators to date have refused to completely accept the position of clinical laboratories that they should not be held accountable for the medical necessity of tests ordered by providers of medical care who order laboratory diagnostic tests. In fact, HCFA has deemed that the clinical diagnostic laboratory should act as a policing body to assure that providers are accountable for correct medical documentation.

THE COST OF COMPLIANCE

Part of "2 - Business Management of the Clinical Laboratory"

Greed and poor judgment on the part of a few (the need for more power and money) has caused an unstoppable overreaction of the federal government's investigation and compliance arm to find new instances of fraud and abuse. Outside the clinical laboratory world, 10% of the nation's trillion dollar health tab is said to be tainted by fraud or abuse, and prosecutors are following the money trail with highly specific software on computers linked throughout the nation. Computer printouts identify billing anomalies and investigations are born. Sudden popularity of laboratory tests and aberrant increases in the amount of patient billing have all spawned federal inquiries (22).

Since the 1990s, a greedy, unethical few have committed Medicare fraud and abuse to increase federal reimbursement for laboratory tests. Their crimes now incriminate and invoke guilt by association for the entire clinical laboratory profession.

The government has also developed an auditing mandate known as the Physicians in Teaching Hospitals (PATH) program and has placed all teaching hospitals on notice that significant penalties can be avoided when hospitals voluntarily disclose billing discrepancies before the government's actual investigation (22).

Written compliance programs are now nationally required in both public and private HCOs as another control on fraud and abuse, and many HCOs have created new full-time positions for compliance officers to manage their compliance program. A compliance program is a formal regime that methodically identifies reports and resolves billing errors. The structure for government-required compliance programs actually took shape in the 1980s as various defense contractors committed fraud and abuse with Department of Defense moneys.

Now, an independent compliance officer must be put in place when a health care facility receives federal funding through

HCFA (Medicare). This officer must have the authority to report directly to the organization's Board of Directors. With this, the compliance officer must have the freedom to investigate and act on incidents of fraud and abuse (F&A). All employees and contractors may be required to sign a pledge of commitment to avoid F&A. HCFA mandates signatures to assure absolute individual accountability for each person involved in the test ordering, production, and billing process.

Violations and indictments of failure to comply with CLIA, OSHA, and HCFA coding and billing regulations may result in the institution's loss of Medicare reimbursement for a minimum of 5 years. In areas of the United States where 75% to 80% of the patient population treated is in the Medicare age group, risking denial of Medicare funding is a preamble to a HCO's financial failure (22).

Because this additional administrative expense is mandated, the HCO parent organization, which administers the laboratory, or the laboratory which is independently owned and operated, must incur additional operating costs to carry out the mandates. The federal government's joint effort between the Department of Justice, the HCFA, the Office of the Inspector General, and the Federal Bureau of Investigation (FBI) to identify health care fraud and abuse is thriving. As of late 1999, even the FBI has developed a new Health Care Fraud program, and agents are visiting HCOs to review medical records.

Now laboratory and hospital administrators must balance the need for patient care and value with the additional variable costs of compliance and regulation created by CLIA ‘88, OSHA, and Medical Necessity guidelines (Table 2.4).

The decision becomes even more difficult when the cost of providing services increases, and the director or manager must choose between maintaining existing functions or establishing new ones. During the budget cycle, directors and managers must always plan to cover the additional expense for compliance with legislative and regulatory mandates, especially if their organization receives even a small amount of federal Medicare or state Medicaid funding.

BALANCING CARE MANAGEMENT WITH BUSINESS MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

Despite the threats and challenges of the external and internal environments, there must always be a concern for the patient's welfare. Care management has become an important and visible issue in the late 1990s with the advent of legislation in Congress and attempts on the part of the Executive Branch departments and agencies of the federal government to assure patient rights, medical records, and safety. This is a welcome turn of events for the professionals who work in the nation's clinical diagnostic laboratories.

In December 1999, the National Academy of Science's Institute of Medicine issued a seminal report incriminating the nation's health care system as responsible for the deaths of nearly 98,000 patients per year from medical errors and mistakes. Clinical diagnostic tests and services were included in the list of reasons enumerated by the their expert panel (23).

Although there is currently no conceivable way to totally manage the provider's use or interpretation of a test result once it has left the laboratory's sphere of control, there is a way to help protect the patient. The electronic information link between the laboratory and the provider is a major safety factor.

The following sections on managing the clinical interface should be consulted for information on how to balance the need for timely and accurate diagnostic information with the need to maintain direct communication with the patient's provider. This, in fact, is the business link that is critical to both the laboratory's compliance with federal medical documentation regulations (the National Correct Coding Initiative) and the laboratory's entitlement to file a claim if the provider has properly assigned a diagnosis at the time the request (order) for the diagnostic test or service was given.

PROMOTING THE VALUE OF THE LABORATORY

Part of "2 - Business Management of the Clinical Laboratory"

The future survival of any paraclinical health care support service (e.g., laboratory, radiology, nuclear medicine), within or linked to a HCO in the first 10 years of the millennium, depends on how valuable the organization's administration and providers think it is to improve their ability to do their job. Even more important than positive financial margins is the measure of a laboratory's team effort to partner with their clinical and administrative counterparts in the HCO.

The clinical laboratory that installs and supports the programs discussed in the following sections is likely to be viewed as a responsible contributing partner in a HCO.

Each category described is actually in place somewhere in a HCO in some part of the North American continent. Each program has proved its value to increasing the value of the clinical laboratory in the eyes of those who decide whether to keep test production in-house or send it to an outside source.

PATIENT OUTCOME MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

Top-down support from administrators and executive staff of a laboratory organization to encourage better methods for patient outcome management programs is critical to the success of the organization's accreditation, inspection, and reimbursement efforts. With the assimilation of physician practices into provider and health system structures, administrators, directors, and managers need to view physicians as equal partners in charting their organization's course (24).

Even reimbursement is now tied to quality of care initiatives, and clinical physicians are subject to audit in health care venues that receive federal moneys, based on the use of the False Claims Act against entities suspected of providing poor quality care. The audits are conducted by nonmedically trained federal agents who use a standards template to assess whether a problem exists. The central issue is a legal argument. Poor quality care equates to no care provided. If a claim for payment was submitted and paid by a third-party payer to the provider and it was false (no care was provided), the provider should have known this and should not have been paid. The lesson for laboratory directors and managers is that they can also be indirectly involved in a fraudulent claim if they do

not check to verify that a test/service was directly attributed to a real patient with a verifiable disease or condition (25).

This suggests that it is more important that laboratory professionals and managers take a strong, active role to protect not only the patient, but also themselves from incrimination if there is an accusation of failure of quality involving a laboratory result or procedure. Directors and managers can no longer be passive and wait for the clinical staff to initiate an outcome management program. They must make resources available and reward improvement. Accountability is the key element, and the organization's leaders are responsible for initiating ideas and establishing a culture that views measurement and comparison with (preferably) a locally developed gold standard as routine (24).

When the potential exists for a laboratory diagnostic result to be misused or inappropriately used by a provider who has insufficient knowledge of how to use the test(s) result(s), then it is vital that there is an electronic query or other rapid, bidirectional information feedback system between the laboratory and the provider to prevent patient injury.

The following factors summarize the current universal problems that face all HCOs trying to improve patient outcomes:

1. There is too much diagnostic information.
2. There is no universal sorting and matching process in the laboratory that ranks and evaluates the importance (value) of a group of related tests to a patient's expected outcome for a provider making a medical decision, other than the flagging of MDT results to identify critical (life-threatening) abnormals.
3. Patient outcome is improved by ranking, classifying, and interpreting diagnostic data against known clinical findings before it reaches the provider of medical care.

PUBLISHING EVIDENCE-BASED DIAGNOSTIC REASONING PATTERNS FOR PROVIDERS

Part of "2 - Business Management of the Clinical Laboratory"

The Physician's Reasoning Process in Diagnostic Testing: The Role of Logic, Memory, and Intuition

The logic that is used in medical problem solving combines the systems of formal principles of deduction or inference, much like the arguments and concepts of Socrates and Aristotle, with the management of information derived from the patient's database and physician's own information base (1).

Even though the physician is unaware of it, much of his/her logic is not intuitive but is based on syllogisms, strategies, and information that have been programmed into the “silicon chips” of short-term, long-term, and submerged memory. By pressing the correct punch-key in the doctor's brain, the internal logic and natural intelligence of this system are activated; what is displayed on the printout is euphemistically called “medical intuition.”

Data processing in medicine is the method by which a database is transformed into a problem list. It is the important step whereby information gathered about a patient is filtered and clues are selected and grouped into meaningful problem sets. The clustering of clues into meaningful groups is also part of data processing, e.g., symptoms and abnormal physical signs in the same organ system. The cluster of clues that seem to belong together may be stated as a problem, even if there is not enough evidence to make a diagnosis.

The Role of Age and Experience: Clinical Judgment

Studies have shown that an experienced physician forms early hypotheses, tracks a key clue, forms a cluster, spots a triad or tetrad, considers a differential diagnosis right away, pursues only one subset of the database, and rapidly zeroes in on a diagnosis. Younger physicians have many more hypotheses than older physicians and generate more clinical events to prove their hypotheses than older physicians (1).

DIAGNOSTIC PATTERN RECOGNITION OF SIGNIFICANT ABNORMALS AND NORMALS

Part of "2 - Business Management of the Clinical Laboratory"

Similarly, there is a need to provide a clustering or a sorting and matching process in the data management process in the laboratory, ranking and evaluating the importance (value) of a group of related tests to a patient's expected outcome. These diagnostic outcome clusters need to be published in a user-friendly, legally acceptable electronic format that can be easily accessed by the patient's provider as soon as the data are analyzed and interpreted by a laboratory professional.

PUBLISHING EVIDENCE-BASED SENSITIVITY AND SPECIFICITY DATA FOR PROVIDERS

Part of "2 - Business Management of the Clinical Laboratory"

There is a central hypothesis that simplifies the task of integrating the laboratory information base with the clinical user: Providing data-based evidence for clinical providers about the sensitivity and specificity of medical diagnostic tests will improve the quality and speed of patient care by showing the provider which MDT best reflect the patient's disease or condition and/or measure the progress in a course of treatment

Integrating Clinical Findings with Medical Diagnostic Test Significant Abnormals

The value of the clinical laboratory to a HCO is its potential to condense and relate the results of the medical diagnostic tests it produces with data obtained from the physical examination and history.

However, integration of clinical findings from the physical examination and history with laboratory findings has rarely been used owing to the difficulty of accessing the information in the medical record and the failure of direct care providers to enter patient-specific information on laboratory request forms.

Using relational databases to correlate clinical with laboratory findings is now possible because the electronic medical record can be accessed in most modern medical care facilities, if a computerized patient record system with a graphic user interface is installed in the organization's mainframe computer (26).

Data can now be accessed in a format that allows the laboratory professional to view both tentative and final diagnostic conclusions reached by a provider, even before the results of the medical diagnostic tests ordered are completed.

It is critical to the laboratory's operational efficiency to use decision analysis as a major tool for the construction of optimized, effective test groups that correspond to the stages (or intensity) of a disease process.

Impediments to Integrating Clinical and Laboratory Findings

The greatest impediment to providing easily readable, integrated, and interpreted results is the lack of economic incentive to stop using older, outdated laboratory information systems (LIS) that have severely dated characteristics. Simply stated, it is not worth it for software developers to invest millions to write the ideal integrative software when it can only be sold to 100 clinical laboratories rather than 100,000 laboratories.

Medical Necessity and Other Standardized Federal Regulations

Although it may not seem evident, there already is a rules-based standardization process that has had and will continue to have a profound effect on the current and future design of each HCO's and clinical laboratory's information system. This rule-based system is known as *medical necessity* and requires exact compliance with federal regulations to receive Medicare reimbursement.

Redundancy in testing is frowned on by government and other third-party payers, and laboratory managers are now responsible, by federal regulation, for establishing a system of checks and balances that acts as a warning system to identify providers who have ordered redundant tests.

If the clinical laboratory is ready to spend the capital and budget the operating funds to support the installation of a system that will support the information needs to comply with medical necessity, correct coning initiatives and other third-party payer documentation mandates, it must select a system that will grow with the vastly unpredictable information requirements of the third-party payers, both private and public.

Third-party payers are the most important entities in a HCO's financial management structure because they provide the organization with the major portion of its revenue from reimbursement.

User-Friendly Information Systems

Diagnostic test result information databases linked with information management software systems offer clinicians and laboratorians greater ability to manipulate data, interpret results, provide interpretive comments, view graphic displays, analyze trends, and isolate significant abnormals.

The report that is issued must be able to be changed whenever the laboratory's payer's rules and regulations for filing claims for reimbursement change.

DISEASE MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

There are facts that are emerging in the new medical culture that have emerged since federal regulations were loosened to allow professionals, other than physicians trained in the traditional M.D. program, to provide their services for patients: (a) the number of nonphysician providers is increasing and (b) nonphysician providers often do not have adequate training or experience in interpreting medical diagnostic test results.

It follows logically that these facts support a need for a new order in the culture of medicine, to manage more carefully diseases and conditions to avoid misuse and inappropriate use of medical diagnostic tests and prevent patient injury. There is simply too much information for the provider of medical care to have adequate knowledge of how to properly use the nearly 2,000 tests and procedures in the laboratory section of the CPT-4 code book. Laboratory professionals are the logical, trained experts in the specialties and subspecialties of laboratory medicine.

Therefore, there is an emerging role for the laboratory professionals to assure that there is an educational channel or method to provide appropriate education in the use of diagnostic tests and services for all users of the laboratory—regardless of the extent of their training. There is a central, underlying theme that supports this role: Interpretation of significant abnormal and normal medical diagnostic test results before they leave the laboratory will improve the speed and quality of medical care. It is now possible to use an electronic interface between the laboratory and the clinical user (the provider) to make this educational effort successful.

DEFINITION OF “PROVIDER”

Part of "2 - Business Management of the Clinical Laboratory"

The previous paragraph referred to the nonphysician provider as a new and emerging challenge for the laboratory's administration to master and control to avoid potential injury to patients from

providers of all types who misuse or inappropriately use medical diagnostic tests. It is necessary to describe the new order of things, as deemed by reimbursement regulations—not necessarily founded in the same logic.

Table 2.5 describes the federal regulations' broad scope of the definition of provider. The reader should note that there are many variations of providers listed, and they all have variable degrees of education in the fields of laboratory medicine and pathology. The presumption that each provider is capable of making both medically necessary and appropriate decisions when using diagnostic tests and services is not a comfortable thought for experienced professionals in the field of laboratory medicine. Therefore, it is reasonable and ethically correct for laboratory experts in all specialties and subspecialties to provide an assistive and consultative mechanism for all clinical providers in the form of disease management.

WHAT IS NECESSARY TO PERFORM DISEASE MANAGEMENT PROPERLY?

Part of "2 - Business Management of the Clinical Laboratory"

Disease management is a term that emerged during 1990s, the decade of managed care, and denotes that interaction and cooperation are required in a modern medical setting to care for patients with chronic diseases. Further, disease management means that the solo physician is no longer totally responsible for maintaining all the activities of patient care and follow-up him/herself to provide care for a patient with a chronic disease or condition. The team approach is now very common in clinical practice, and physicians, nurses, and specialists of many backgrounds all care for the patient in a continuum.

Although the laboratory professional, both physician and nonphysician, has long been excluded from the hands-on delivery of patient care by some archaic tradition, it is no longer reasonable to exclude the experts who have the correct answers from the continuum of care. The following recommendations are provided for the active, challenged laboratory professional who has no fear of cultural barriers in the clinical sphere and cares about helping their clinical colleagues deliver patient care more effectively.

Interpretation of Significant Abnormals for Nonphysician Providers

Because there is so much variation in the educational preparation of nonphysician providers and because the loosening of the Clinical Laboratory Improvement Act's supervisory rules for monitoring the quality of test performance, this is a new and important area of laboratory professional expertise to be transmitted to those who may not have had sufficient training in laboratory medicine.

Direct Consultation of Laboratory Professionals with Clinician Specialists

For example, in the management of diabetes mellitus, the most critical test that allows the provider to measure the patient's progress is glycohemoglobin, yet providers persist in ordering a wide variety of tests in many areas of clinical chemistry to follow patients. Table 2.6 is an example of a protocol that allows the laboratory professional who is an expert in diabetes test interpretation and management to consult with diabetologists, endocrinologists, and the diabetes treatment team by reviewing a wide spectrum of test results. The objective of the format presented in Table 2.6 is to allow the laboratory professional to teach the provider how to use parsimony and economy in test selection to avoid unnecessary workload and higher costs for the laboratory. This is necessary because many clinical laboratory tests, ordered especially for hospitalized patients, appear to be redundant. Investigators in 1998 found that eliminating redundant tests in just one hospital would have saved nearly $1 million (27).

TABLE 2.6. COST BURDEN FOR DIAGNOSTIC TESTING AND MONITORING OF DIABETES CASES

		Initial Diagnosis			*Diagnosis and Management of Complications and Treatment*								
Clinical Stage		**Detection**	**Confirmation**				**Monitoring of Treatment of Complications**						**Followup**
					Uncomplicated		**Complicated**						
Episode of Care		1	2	3	4	5	6	7	8	9	10	11	12
Cost/Episode of Care		$	$	$	$	$	$	$	$	$	$	$	$
Cumulative Cost		$	$	$	$	$	$	$	$	$	$	$	$
Type of Diabetic Condition	Prediabetes						Cardiac	Renal	DKA	Infection	Immune Sys & Drug Res.	Metabolic	Screening
	Impaired Glucose Tolerance	SM(WB)G UA	FPG OGTT										FPG/yr
	Gestational	SM(WB)G UA	FPG OGTT										FPG/yr
	Primary												
	Type 2 Insulin Resistant (formerly NIDDM)		FPG OGTT HbA1c		UA EKG Lipid Prof Tot/HDL Chol		EKG Card Enz Lipid Prof Tot/HDL Chol	UA Ur Albumin Cr BUN	UA pH HCO3 PO4 Mg Ca Acetone Ketones Osm	CBC Diff Gram Stain Cult/Suscept blood urine wound	Vit B12 (Metformin users)	FPG HbA1c/qtr CBC Diff	FPG/qtr HbA1c/qtr CBC/qtr
	Type 1 Insulin Dependent (IDDM)		FPG OGTT HbA1c Antibodies to: Insulin anti-glutamic acid decarboxylase(GAD)		EKG UA Cr BUN Lipid Prof Tot/HDL Chol		Ur Albumin	Lactic Acid Na/K/Cl/CO2 PO4 Mg Ca Acetone Ketones Osm Lactic Acid Na/K/Cl/CO2			Insulin C-peptide Insulin levels Anti-beta cell ab Anti-insulin ab	pH HCO3 PO4 Ketones Na/K	FPG/qtr HbA1c/qtr CBC/qtr
	Secondary	FPG	OGTT HbA1c				Lipid Prof Tot/HDL Chol				AutoAb to: Parietal cells Thyroid Stomach		

This table is provided as a courtesy by Dr. E.M. Travers, Annapolis, Maryland.

Provide Information on the Actual Costs of Tests and Services

Most providers authorized to order medical care procedures, tests, and activities do not know about the operating costs of their decision-making process. They have a plethora of information about the technical and safety requirements but do not have instant cost information delivered to them about the actual cost of an operation, procedure, test, new equipment, or medical activity.

Most doctors and caregivers have no idea how to begin to know their costs or the costs of their practice group, whether it is in a hospital, clinic, or office setting. Although articles on cost-effectiveness and sporadic articles focusing on cost-effectiveness (not operational costs) appear regularly in current subspecialty medical and nursing journals, there is no generic, single source for anyone with training and background in health care to quickly grasp the importance of knowing the costs that they incur when caring for a patient.

The national (and international) variation in health care provider activities has become a target for (a) standardizing medical care and (b) denying reimbursement when it does not meet the requirements for reimbursement of third-party payers.

Reduce the Variation in Testing Using Evidence-Based Medicine Protocols

Attempts to reduce test variation have been made using evidence-based medicine. Evidence-based medicine (EBM) is a discipline focused on the avoidance of the use of personal opinion, practical experience, and intuition in the medical care process. Instead, it is a style of medical practice in which the caregiver bases a practice decision on evidence from well-designed studies.

EBM should be used by physicians and other caregivers to provide the best standard of care for each patient.

Manage the Clinical Information Interface

The emergence of the electronic information technology platform has allowed progressive HCOs to thrive, especially with the advent of software that bypasses the manually written requests for tests. Similarly, an electronic return of information to the

TABLE 2.5. DEFINITION OF A PROVIDER

An entity authorized to receive direct (or indirect "incident to" reimbursement) such as a	
Hospital Rural care primary hospital	with an agreement to participate in Medicare
Skilled nursing facility	
Home health agency	
Hospice	
or a	
Clinic Rehabilitation agency	with an agreement to furnish outpatient services
Public health agency	
or a practitioner who is a Physician	
Nonphysician practitioner (e.g., optometrist, podiatrist)	with a provider number furnished by HCFA or a Managed Care Organization
Nurse (clinician, practitioner, specialist)	
Physician's assistant	
Clinical psychologist	
Clinical social worker	
Medical nutrition support	
Physical occupational therapist	

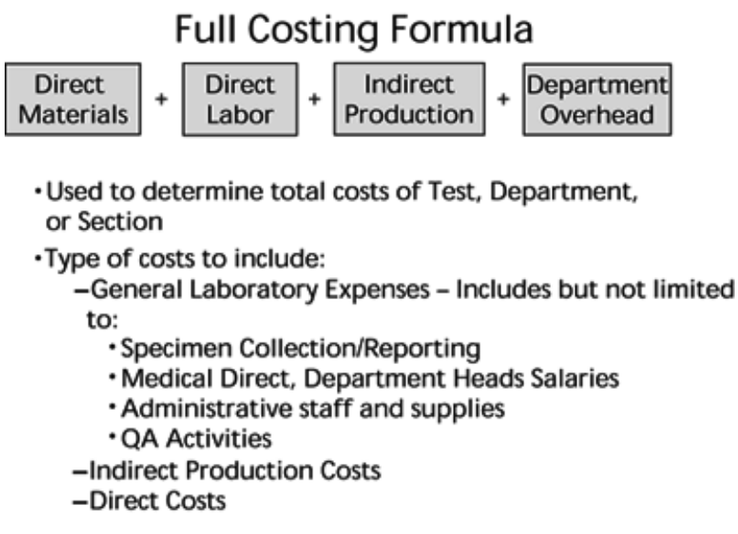


FIGURE 2.5. Full costs, also known as *Total Costs* measure the *actual* direct and indirect costs of providing a service or product.

provider through a bidirectional electronic medical record allows laboratory professionals to interact rapidly in two high-cost, intense areas of hospital practice: transfusion medicine and pharmacy. Hospital spending for operational costs continues to consume the largest portion of the health care dollar (3). In addition to pharmaceuticals, one of the highest categories of operating cost in tertiary care hospitals is for transfusion medicine, blood products or derivatives, and the clinical support services for transfusion, infusion, or therapeutic pheresis of patients (Table 2.1).

More than any other tertiary care setting services, these services consume the major portion of hospital's operating budget and cause the organization and the laboratory more difficulty than most services because their products and services ordered by providers are so expensive to purchase and are often in short supply.

They are also high-risk services and often add to the organization's legal burden for allergic, adverse, transfusion, and other reactions. Prompt, direct consultation by any means, verbal, electronic, or telephone, can save lives and avoid adversity, waste, and inappropriate ordering and dispensing of scarce, expensive products.

LEADERSHIP AND BUSINESS RELATIONSHIPS

Part of "2 - Business Management of the Clinical Laboratory"

Outside the laboratory-clinical interface arena, the four most important management issues for the laboratory director or manager of the 21st century are leadership, communications, politics, and management of human resources.

Leadership

A successful leader recognizes and controls the forces that affect the laboratory and its staff. This begins with reading about, listening to, and participating in new educational ventures that instruct on the external and internal changes, trends, and forces that alter the course of the laboratory's daily mission.

The laboratory's leader should encourage the staff to attend courses and seminars that teach the truth and realities of national and regional health care economics. For instance, to know how to cope with rapid, unexplained change in operations, the laboratory's leaders in operations should be encouraged to attend management courses and workshops at local colleges or at national meetings. The compliance staff should attend educational activities promoted by medical and laboratory professional societies and visiting lecturers; attorneys who are specialists in the regulatory affairs that affect laboratories should be invited to speak to the staff who are most likely to interact with the providers who order tests and services.

Leaders may find themselves reading and seeking out information far beyond the traditional references and journals published for laboratory professionals. The health care administration literature is especially thorough in its coverage of global issues affecting hospitals and other health care delivery sites. Journals in the fields of health affairs and health economics concentrate heavily on reimbursement issues that often affect clinical laboratories.

Communications

The interactive skills we learn before we become laboratory professionals often are inadequate to support us in the later years of our professional activities. This is because laboratory professionals traditionally identify themselves as scientists and/or medical caregivers and not managers or politicians. The personality of scientists is one that favors introversion, focus, and concentration, avoiding confrontation at all costs in the best interests of promoting the pure thought needed to enhance the value of the scientific contribution to be made. The personality of the business leader is one that may not enjoy confrontation and conflict, but he/she recognizes that it is the normal daily activity that occurs in every workplace.

Managers who are scientists as well as leaders must reject their nonconfrontational instincts nurtured in the hallowed scientific halls of their early experiences. Leader managers must be willing to develop a new management persona—one that works as a player on a team and not as an individual striving for recognition of solitary efforts.

Successful managers learn to overcome their inhibitory aspects of their “early” personality, accepting change as a given to succeed in the task of management.

Politics

Also recognizing that all American hospitals are now no different than the traditional business workplace, a laboratory's leader understands the politics of the workplace. Unpopular as the term politics has become in the American vernacular of the 1990s, it usually negatively connotes a person or organization who has chosen a lifestyle/method to gain power and personal recognition by compromising a value system.

To the contrary, one of the definitions of politics is defined as “the art or practice of administering public affairs” (19). A hospital or HCO certainly fits the description of a public entity—as a provider of medical care services to the public. The leader in a public organization is expected to carry out the responsibilities and duties that assure the safety and value of services provided to the public.

Because our country is so successful at preserving life, we are experiencing the heavy, complex clinical workload associated with the aging of America. To deal with the cost of caring for our aging parents and relatives, Congress has appropriated money to be administered through a public system of administration. Because the process of allocating funds is totally connected to politics (as we have just defined), we are captured in the process of administering public affairs.

Because so many of our aging citizens are Medicare and Medicaid dependent, the practice of health care in the United States has become, by definition, a public process. Because the majority of American hospitals (more than two thirds) are currently financially nearly totally dependent on federal Medicare funding to maintain their operations successfully, the administration of the process of receiving the dollars (reimbursement) has also become subject to the laws and regulations required to justly and fairly administer the distribution of the taxpayer's dollars.

Therefore, the laboratory's leaders must recognize that they now work in the political system of the health care arena, regardless of where they practice their profession, and must follow

the law of the land to receive adequate compensation for the services performed by the laboratory's staff.

For these reasons, it is more important to understand both the politics of the health care workplace and the politics of the administration of public funds—even though we may personally despise the unscientific nature of the process.

Ultimately, the laboratory's leaders must become skilled in political maneuvering to

- protect the rights and health of a patient who may be denied tests and services that are justifiably needed,
- form alliances with other laboratory professionals in hospitals, clinics, and distant sites where medical care is delivered to assure that adequate tests and services are provided, even though they may not be fully reimbursed by third-party payers,
- become vocal and vibrant supporters of the laboratory's duty to provide the highest quality of care even though it may be opposed by administrative and nonmedical professional decision makers who control the resources (budget) for the laboratory.

INFORMATION MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

Implementation of a responsive, flexible, bidirectional information system is the first line of defense for any laboratory management team determined to increase the value of the laboratory for patient care, comply with regulations, and reduce its operating costs. It will profoundly affect daily operations, both positively and negatively for many years after its installation.

It will either permit you to improve your quality, clinical effectiveness and outreach, business, and revenue position or result in chaos and frustration for clinical staff, your own staff, and the organization's top management.

A simple structured approach to understanding the most important elements of an effective information system in a modern laboratory, integrated or not, include:

I. Primary (active) database
   Interfacing the clinical laboratory with the electronic medical record
   Database connectivity
II. Secondary (passive) database
   Workload
   Coding
   Provider verification of diagnosis
   Financial
   Billing
   Compliance
   Accreditation
III. Networks
   Intranet
   Internet

TOTAL COST MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

Total cost management in a progressive and successful clinical diagnostic laboratory encompasses much more than finding the actual cost of a test or accumulating costs to prepare an annual budget. Total cost management in the clinical diagnostic laboratory is a comprehensive planning process that:

- anticipates all the expected categories of cost to produce and deliver high-quality medical diagnostic information to its clinical users,
- includes a method to estimate when to stop producing products that are not cost-effective or medically needed,
- includes a method to predict when to start producing more products that can be sold to gain new revenue for the laboratory,
- includes the hidden costs required to professionally certify, accredit, and inspect the laboratory,
- includes the support costs to meet the legislative and regulatory mandates of federal and other third-party payers who require verification of the accuracy of medical documentation and medical necessity,
- provides a flexible framework to meet and cover unanticipated costs caused by emergencies, failures, and other unplanned contingencies without financial deficit,
- supports the business expenses of sales, marketing, and outreach to introduce the laboratory's products to new customers,
- supports the information management, Intranet, Internet, and telemedicine linkages needed to communicate with all clinical users in the network.

Achieving efficiencies and productivity improvement depends on understanding the fully allocated (total, true) per unit cost of delivering the result of a test or service performed in the clinical diagnostic laboratory (Fig. 2.5). If a manager does not have a baseline from which to measure cost, it will also be impossible to determine whether there was gain or loss after delivery of a desired level of production.

To improve productivity in laboratory operations, whether it be at the level of the workstation or at the section, division, or department level, a knowledge of fully allocated costs is needed before a manager can begin to reengineer the work process that produces deliverable, billable information to a provider. Precise and meaningful cost management is critical for determining the effectiveness of both human and fiscal resource use—a prerequisite for effectively managing global risk.

In addition to knowing the full cost of the laboratory's tests and services produced, the director or manager has the responsibility to assure the organization's top management that the margin of revenue gained after the receipt of reimbursement and other revenue streams is greater than the cost of producing all the laboratory's products and services. This is known as gaining a positive contribution margin for not-for-profit organizations and gaining a positive profit margin for for-profit organizations.

MANAGEMENT RESPONSIBILITIES AND CONTRIBUTION/PROFIT MARGIN

Part of "2 - Business Management of the Clinical Laboratory"

In a traditional business, during times of unrestrained resource availability, profit-maximizing decision makers will elect to produce all the products they can as long as their marginal revenue

(MR) (the price or charge) exceeds marginal cost (MC) (MR > MC).

However, in times of limited resource availability such as mandatory decreases in laboratory reimbursement caused by the Balanced Budget Act of 1997 (see Legislative Factors section), both contribution and profit-maximizing decision makers should (and will only) select those tests and services for market expansion that will have the greatest effect on producing a high contribution or profit margin.

High-Cost Areas of the Laboratory

There are two high-cost areas of the laboratory that need to have more attention paid to them to reduce operating costs (expenses)—complex, high-cost, low-volume laboratory medical tests and all the elements that comprise the practice of transfusion medicine.

High-Cost, Low-Volume, Esoteric Tests and Services

A generic example of the difficult decisions that have to be made at all levels and in each section of today's clinical diagnostic laboratory environment is given in the following.

Example

A clinical specialist has asked the laboratory to use a series of very sensitive and specific, but very costly screening tests. The justification is made that regular receipt of the results from these tests will greatly improve patient outcome. There are only a few patients in the laboratory's demographic database who will benefit from these tests. The workload volume generated by the provider's request is also very low. We assume that from what the provider has justified in writing that providing a more sensitive test than what is currently being used will greatly improve the quality of care and the outcome for the small group of affected patients.

Solutions

- Use evidence-based medicine to come to a decision. This is a style of medical practice in which the caregiver bases a practice decision on evidence from well-designed studies. The solution can be provided by the maxim “trust but verify.”
- Use the “substitution rule.” In classical economics, a substitute is a good (product) that can be used to replace another good product, with a positive effect on the attitude of the consumer of the good (28). For example, a laboratory professional's search of the Internet and other databases to find a similar test or group of tests to prove or disprove the clinician's supposition that the requested test is truly more specific needs to be done as part of the management decision process. This is classical economics applied to a practical situation. The laboratory professional is best qualified to perform the search and act as a consultant to the provider to find a product (test) that at least partly satisfies the needs of the provider. The laboratory professional finds the “substitute good that satisfies the needs of the provider, while assuring that the good (test) can be performed with a high level of specificity and sensitivity for the patient's disease/condition at a lower cost than the high-cost test that was requested by the provider.
- Use “make or buy” analysis (next section).

As often happens in clinical practice, the provider ordering the tests does not have the time or on-site resources to use evidence-based medicine, does not have a global view, and operates within a narrow, subjective frame of reference directly related to the provider's experience and the current teachings and literature on the clinical specialty or subspecialty. When individual variation is combined with clinical progress in diagnosis and use of the latest medical technologies, physicians and other caregivers blindly opt for the best standard of care and often ask the clinical diagnostic laboratory to resolve the problem objectively with a specific test or tests.

Because of the trend toward more esoteric testing on-site, as discussed in the introduction to this chapter, directors and managers need to have three additional production control methods in their management tool kit in addition to evidence-based medicine and the substitution rule, tools to know when to start and stop testing (make or buy decisions).

Make-or-Buy Analysis (Break-Even Analysis)

After the decision is made regarding the sensitivity, specificity, and applicability of the test to the patient's disease or condition, laboratory directors and managers have the responsibility to decide when and where it is most expedient to perform tests: in-house (“make”) or in another laboratory (“buy”).

This problem arises particularly in connection with the use of idle or underutilized equipment, idle space, or idle labor or with insufficient equipment, space, or labor to perform tests. Faced with a “make-or-buy” decision, the manager should (a) consider the quality of the tests offered by the outside source of testing, (b) compare the cost of buying the test with performing it in-house, (c) determine the medically necessary turnaround time, (d) evaluate consultation and clinician review capabilities, and (e) estimate the problems with cost and the logistics of sample transport.

If the decision is made to send tests to another laboratory, the laboratory's financial manager should prepare a statement that compares the laboratory's cost of performing needed tests with the potential vendor's price. The statement should present the differential costs of the tests as well as a share of existing fixed expenses

and a profit figure that compares the costs on a comparable basis. The laboratory's budget should also be revised to indicate the effect of the change when the transfer of tests is effected. The following worksheets and decision-making aids can be used to assist in the process of evaluation:

"Make" Decisions: When to Start (or Continue) Testing

Table 2.7 is a checklist for determining whether the capacity exists to add a new (or additional) workload to an existing workstation configuration, section, or department. It is a rough estimate of whether there is sufficient support within the laboratory to add more tests to an existing test menu. Simply stated, the laboratory needs to know what its current workload output is by workstation, section, and department before it can make a decision to add new workload. The "normal" rate of utilization (workload) must be determined for the area where the new test(s) will be added, and if the workload is below normal (excess capacity exists), it may be feasible to add the new test.

TABLE 2.7. CRITERIA FOR DETERMINING IF EXCESS CAPACITY EXISTS IN YOUR LABORATORY TO ADD NEW WORKLOAD FOR REVENUE GENERATION

	Yes	No
• Equipment		
Incomplete (not utilized at capacity) utilization of equipment	____	____
Duplicative equipment items not used to full capacity	____	____
The majority of high volume automated equipment is not labor intensive and has "walk-away" features	____	____
Equipment is highly reliable and supported by excellent repair services	____	____
Equipment is adaptable to adding/deleting from menu for low vol/high cost tests	____	____
• Human Resources/Consultation/Test Interpretation Skills		
Expert medical knowledge/experience is on site (pathologists)	____	____
Expert scientific knowledge/experience is on site (scientists)	____	____
Expert technical knowledge/experience is on site (technologists)	____	____
Expert information systems expertise/experience is available	____	____
Expert business/billing expertise/experience is available	____	____
Idle capacity exists in latter half of day shift (technical)	____	____
Idle capacity exists on evening and night shifts	____	____
The skill mix for the medical technologist work force is high/varied	____	____
Laboratory management has an interest in, supports and provides resources for cost-per-test analysis	____	____
• ADP/Information Systems		
ADP system is currently adaptable to modern/telepathology and other outreach connections	____	____
Test turnaround time is exceptionally good for the 25 most commonly requested tests	____	____
Bar code and other time-saving techniques are used to access, process, inventory, and report	____	____
• Quality/Accreditation		
The laboratory's quality is consistent and regarded highly in the community, is defensible and has been recognized by accepted accrediting bodies	____	____
• Teaching/Research		
The laboratory offers tertiary-level, highly-specialized tests often needed by teaching/research physicians in university settings or by clinical specialists practicing in the local community (when a university hospital is not nearby)	____	____
Interpretation: A predominance of "Yes" answers indicates that your laboratory has excess capacity to add new workload. Whether you can afford to do it depends on cost analysis (see Chap. 27, Appendix 27A.)		
TOTAL		

Major assumptions must be explicitly defined and evaluated quantitatively for incremental cost assessment as shown in Table 2.1. As an adjunct to this table, information on the incomplete utilization of equipment and labor capacity can be readily determined using objective assessment techniques. This is often true for high-throughput routine sections with automated instruments. The "bottom line" is that to be able to add new workload, these criteria must be used to establish that excess capacity exists in the laboratory. From *Increasing Revenue Using Excess Capacity and Decreasing Costs*, © 1997 by E. M. Travers.

TABLE 2.8. USING MARGINAL REVENUE AND MARGINAL COST CONCEPTS TO EXPAND TEST PRODUCTION

(1) Quantity Per Day, Q (given)	(2) Total Cost, TC (given)	(3) Price Per Unit, or Average Revenue, P = AR* (given)	(4) Total Revenue TR (3) × (1)	(5) Average Total Cost, ATC (2) ÷ (1)	(6) Marginal Cost, MC Change in (2) / Change in (1)		(7) Marginal Revenue, MR Change in (4) / Change in (1)	(8) π Net Revenue NR (4) − (2)
0	$ 25	$ 10	$ 0	$ —	$ 10		$10	−25
1	35	10	10	35.00	6	What it costs to produce	10	−25
2	41	10	20	20.50	4	1 extra unit	10	−21
3	45	10	30	15.00	2	"	10	−15
4	47	10	40	11.75	2	"	10	− 7
5	49	10	50	9.80	3	"	10	+ 1
6	52	10	60	8.67	5	"	10	+ 8
7	57	10	70	8.14	8	STOP PRODUCING	10	+13
8	65	10	80	8.13	14	HERE	10	+15 MAX π
9	79	10	90	8.78	21		10	11
10	100	10	100	10.00				0

* average revenue

$MC = \frac{\Delta \text{ cost}}{\Delta \text{ vol}}$

marginal revenue (MR) is greater than marginal costs (MC)

From Leiken A. Lecture notes and handouts. State University of New York, Stony Brook, NY, 1996.

"Buy" Decisions: When to Stop Testing (or When Not to Test)

Table 2.8 illustrates a method used to determine when to cease testing, based on the break-even point method. Even if there is excess capacity, it may not be prudent from a fiscal point of view

to perform the new tests in-house. This is when it is important for the director or manager to know:

- the projected volume of tests that will be needed over time,
- the actual (job order, “true”) direct and indirect production costs of the test,
- the rate of reimbursement (charge, fee, price).

Again, the important relationship between MR and MC is illustrated in this table. The goal of the director or manager is to reach the point where MR > MC.

If MC and the laboratory's indirect (fixed) costs are known and they are completely covered by reimbursement, then the total revenue from the production of the test can be maximized by increasing test production. If the costs to produce the test exceed the reimbursement or other revenue received, then it is not wise to expand test production.

Event Costing

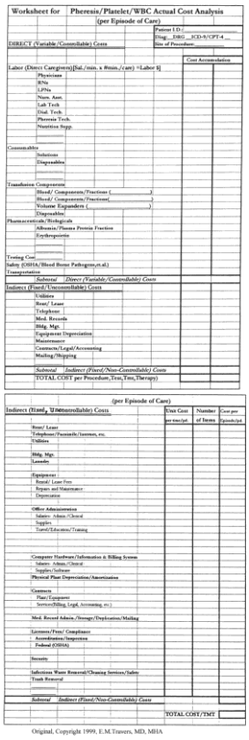
Event costing is the technique used to find the cost of any activity, procedure, test, or event that incurs expenses for the laboratory during its production by the laboratory staff.

Two rules are important for event costing:

- Finding the actual (true) total cost of the activity, procedure, test, or event to be produced is critical. It cannot be estimated (it must be calculated to assure accuracy).
- The full array of costs required for the production of the activity, procedure, test, or event must be included in the calculation.

Event-Costing Worksheet

A worksheet for calculating in-house total test costs (TC) for one activity, procedure, test, or event to aid in making make-or-buy decisions can be found in Table 2.9.



Worksheet for Pheresis/Platelet/WBC Actual Cost Analysis

(per Episode of Care)

Patient I.D.:

Diag.: ICD-9/CPT-4:

Site of Procedure:

DIRECT (Variable/Controllable) Costs — Cost Accumulation

Labor (Direct Caregivers)[hd./min. x #/min./Case) =Labor $]

Physicians
RNs
LPNs
Nurs. Asst.
Lab Tech
Diad. Tech.
Pheresis Tech.
Nutrition Supp.

Consumables

Solutions
Disposables

Transfusion Components

Blood/ Components/Fractions (_____)
Blood/ Components/Fractions(_____)
Volume Expanders (_____)
Disposables

Pharmaceuticals/Biologicals

Albumin/Plasma Protein Fraction
Erythropoietin

Testing Cost
Safety (OSHA/Blood Borne Pathogens,et al.)
Transportation

Subtotal Direct (Variable/Controllable) Costs

Indirect (Fixed/ Uncontrollable) Costs

Utilities
Rent/ Lease
Telephone
Med. Records
Bldg. Mgt.
Equipment Depreciation
Maintenance
Customers/Legal/Accounting
Mailing/Shipping

Subtotal Indirect (Fixed/Non-Controllable) Costs

TOTAL COST per Procedure,Test,Test,Therapy)

(per Episode of Care)

Indirect (Fixed/ Uncontrollable) Costs

	Unit Cost per item/pkt	Number of Items	Cost per Episode/pt
Rent/ Lease			
Telephone/Pagers/Intercom, etc.			
Utilities			
Bldg. Mgt.			
Laundry			
Equipment :			
Rental/ Lease Fees			
Repairs and Maintenance			
Depreciation			
Office Administration			
Salaries: Admin./Clerical			
Supplies			
Travel/Education/Training			
Computer Hardware/Information & Billing System			
Salaries: Admin./Clerical			
Supplies/Software			
Physical Plant Depreciation/Amortization			
Customers			
Plant/Equipment			
Services(Billing, Legal, Accounting, etc.)			
Med. Record Admin./Storage/Duplication/Mailing			
Licensure/Fees/ Compliance			
Accreditation/Inspection			
Federal (OSHA)			
Security			
Infectious Waste Removal/Cleaning Services/Sales			
Trash Removal			
Subtotal Indirect (Fixed/Non-Controllable) Costs			
TOTAL COST/TMT			

TABLE 2.9. AN EXAMPLE OF A WORKSHEET FOR CALCULATING IN-HOUSE TOTAL TEST COSTS FOR ONE ACTIVITY, PROCEDURE, OR EVENT

Cost Management in Transfusion Medicine

State-of-the-art transfusion medicine practices in academic health care settings, responsible for developing and implementing complicated, cutting-edge transfusion medicine technologies and treatments, are difficult to support and finance. This is primarily because of escalating acquisition costs for blood products and an increasing volume of the sickest oncology patients referred from primary care sources. A third cause of high costs is the increased use of commercially prepared high-cost biologicals and factors derived from blood plasma used to treat rare diseases or promote precursor cell growth for transplantation or platelet generation. In addition, high operating costs for each of these are augmented if the hospital permits the practice of transfusion/infusion or therapeutic pheresis (T/I/P) in multiple locations within the HCO's physical plant, i.e., decentralized satellite T/I/P sites. The most important principles for the laboratory director and transfusion medicine chief to remember include (29):

- Find the cost to procure + test + T/I/P each blood or blood-derivative product line/service
- Estimate the revenue to be received from each payer for this product line/service
- If the cost exceeds the reimbursement, decide whether the product line should continue to be produced in-house (make) or purchased (buy) from an outside source

UTILIZATION MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

Utilization management is a control process used to alter the rate of use of a service, product, or the process of production. In medical care, the production process is primarily a reaction to the orders given by providers on the clinical side of the organization.

TABLE 2.10. DIFFERENCE IN COMPLEXITY OF TESTING PROCEDURES BETWEEN LABORATORY SECTIONS

Sections Producing One Result/One Cost	Sections Producing One Result/Multiple Costs		
Routine chemistry	Special chemistry		
Therapeutic drug monitoring	Immunochemistry		
	Endocrinology		
	Toxicology		
Routine hematology	Manual urinalysis		
Routine coagulation	Blood bank		
	Immunology		
	Microbiology		
	Mycobacteriology		
Automated urinalysis	Mycology		
	Virology		
	Immunopathology		
	Surgical pathology		
	Cytology		
	Autopsy		
	Electron microscopy		
Reportable = result	Preliminary result	Interim result	Final reportable result
	↓	↓	
	[Cost 1 +	**Cost 2 +**	**Cost 3] = Total cost**

(*Note*: There may be more than one interim result, depending on method.)

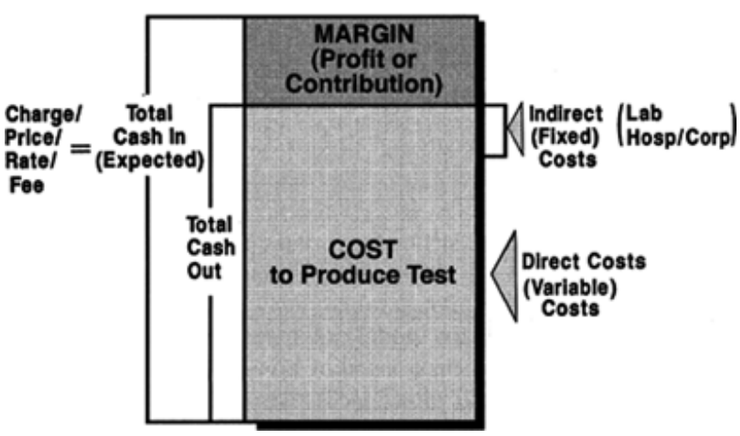


FIGURE 2.6. Normal anatomy of a billable test: the relationship of margin to cost with the anatomy of a billable test, the prime source of revenue/income for a laboratory. The emphasis on the left side of the diagram is placed on the difference between cash in and cash out, showing that cash out is the cost to produce a test. On the right side of the diagram, the cost is composed of two basic parts: direct (variable) and indirect (fixed) costs. Direct costs are controlled by the director/manager and account for nearly 95% of a laboratory's operating expenses.

Often there is no communication between the laboratory and the clinical utilizers (providers), and this usually precludes control of utilization. The facts and hypotheses important to understanding the successful implementation of utilization management include:

1. There is no interpretation of diagnostic information before it reaches the direct care provider.
2. Clinical users of information provided by the laboratory must receive information about their own, unique clinical diagnostic test utilization patterns.
3. Selection of disease-related abnormal and normal medical diagnostic tests and prerelease interpretation reduces utilization.
4. Laboratory professionals must institute an active process of control diagnostic economy over the inappropriate utilization of tests and services if the goal of financial success is important.

Definition of Diagnostic Economy

The definition of the word *economy*, “the efficient and sparing use of the means available for the end proposed” blends nicely with Webster's definitions for the noun/adjective diagnostic. The definitions “a careful, critical study of something to determine its nature or importance” and “the art or acts of identifying a disease from its signs and symptoms” describe the special skill that emerges only after a physician has devoted his/her intellectual and cognitive life during medical training to the accumulation and assimilation of normal and abnormal facts about the function of the human body and its pathologic conditions. Then again, it is not only the training period, but also years of experience using this process that makes the “careful, critical” part of the definition of “diagnostic” meaningful.

Indeed, the two words placed together in sequence are most significant in today's medical economic environment, where the federal government's regulations for reimbursement demand the precise, exact use of descriptive terms to designate diagnoses (ICD) and procedures (CPT), terms to qualify a provider for reimbursement with Medicare or Medicaid moneys (see Coding section).

If the definition of diagnosis means “careful and critical,” then the addition of the word economy to it further enhances the meaning of the phrase *diagnostic economy*. The new meaning becomes “the efficient and critical choice of careful, diagnostic studies that identify a disease to determine its nature or importance.”

Why must we emphasize economy in the diagnostic process? Because previous generations of physicians never had to face the reality of knowing about the costs of the procedures that they needed to establish, confirm, or monitor diseases and conditions. Although it has not entirely been their fault, clinicians have not had the opportunity to learn about test costs. Laboratory professionals in the past rarely took the initiative to present the cost problem to clinicians, because hospitals depended on the income generated from laboratory and other tests. Therefore, a multitude of test results was always passively produced, uncomplainingly, and without question.

Similarly, for-profit hospitals have always emphasized charges—none had to be concerned with actual costs until 1984 when the Tax Equity and Fiscal Reform Amendment produced the paradoxical shift from income-generating revenue centers to cost centers. Market-oriented reformers argue that more care—more tests, more treatment, longer hospital stays—is not necessarily better care.

PRODUCTIVITY MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

It has already been established that to improve productivity in laboratory operations, whether it be at the level of the workstation or at the section, division or department level, a knowledge of fully allocated costs is needed before a manager can begin to reengineer the work process that produces deliverable, billable information to a provider (see Total Cost Management section).

Benchmarking

Benchmarking is the process of setting standards for future performance based on analysis of data and trends from past production processes.

Measuring trends in daily workload and quality control is expected in daily laboratory operations; however, measuring the progress of productivity and efficiency in the laboratory is also needed to improve production and enhance cost management. In quality control, laboratorians use a gold standard, purchased from a manufacturer of laboratory reagents, to assess whether a test result is on the mark, i.e., does the result received come close to what the manufacturer's standard says it should be.

In productivity management, benchmarking is done by the director and manager to measure current production processes and compare them with a gold standard they have established as a goal for future production levels and characteristics. Unlike the reagent manufacturing industry that prepares standards with known values, there is no national benchmarking “manufacturer” to set the goal for what a laboratory's productivity and efficiency (and staffing levels) should be. Therefore, the art of management enters into the process, and the director or manager becomes the leader who sets the goal, the standard, the pace, and the direction in which to head to achieve high-quality patient care and ensure financial equivalency with MR = MC, if not better.

Operational Productivity

All sections and workstations must do more with less to increase operational productivity. This is defined by the equation:

$$\text{Productivity} = \frac{\text{Output}}{\text{Input}} = \frac{\text{Products produced}}{\text{Resources consumed to produce products}}$$

Using the ratio Total Billable Tests/Total Technical Full-Time Equivalents (FTE), to measure productivity shows how technical labor productivity, the laboratory's largest operating expense, is improving or declining. An increase in the number of Billable Tests with the same or fewer Technical FTE is the desired productivity improvement.

Note that the terms billable and technical are italicized for emphasis. It is important to note that the only billable tests and services that can be claimed by the laboratory for reimbursement by third-party payers, even though there are many nonbillable tests, procedures, and activities that have to be “paid” for by the laboratory's budget dollars. Therefore, to increase productivity, it is critical to always produce more billable tests and services than nonbillables.

Departmental “C4” (Consolidation, Constriction, Constraint, Cross-training) and Integration

In 1999, the newly elected president of a major professional laboratory organization summed up the impact of the need to reorganize the way laboratories produce their products in his acceptance speech:

> Pressure to reduce costs will continue to play a major part in how testing is performed. We will respond by consolidating small operations into regional laboratories. Hospital laboratories will either expand testing volume to support appropriate automation or reduce testing menus by outsourcing—to higher volume laboratories—tests that cannot be cost-justified. Such actions will be required to increase efficiency but will also be necessary as a result of a decreasing supply of qualified laboratorians (30).

Corporations, airlines, hospitals, and businesses are all merging to increase productivity. Not only has it worked to reduce operational costs and encourage the economy, but also it provides an example for the reorganization of the “business” of the laboratory. Similarly, the consolidation and integration of the laboratory's “small businesses,” its sections, subsections, and workstations, can greatly improve the volume and quality of products produced (output) per staff member.

However, as the leaders of corporations and businesses agree to merge and consolidate their attributes, the laboratory's managers of departments, divisions, sections, and subsections, must agree to relinquish a certain degree of control and “turf” (space and equipment) to achieve a better result for the laboratory as a whole. This is not an easy task and may not be able to be realized unless the director of the laboratory and the department or section manager display confidence, assurance, and the belief that consolidation and integration are a common goal that projects a “win/win” result for the staff. It is critical to remember that the real issue that defeats any good idea is lack of “buying into” and cooperation from the staff who perform the tasks on the production line.

It is important to remember that the most brilliant strategy and best management theories from famous universities and teachers never reflect real situations on the production line. A time-honored cliché in the senior management cadre is always heeded: “the culture eats the strategic plan every day for breakfast.”

Consolidation, Constriction, Constraint, Cross-training, and Integration

The director's or manager's decision to institute an internal change in the work pattern of the laboratory is certain to add additional stress and concern for the staff who will be directly involved in the change. Because many staff persons may be unwilling to change or may accept change but fear the consequences of the change, it is crucial to involve them in the planning process. Adequate preparation for any merger, consolidation, or integrative activity always starts with the informed consent and briefing of the staff (the culture) to be involved in the proposed change.

Current Trends for Internal Reorganization

The internal and external laboratory restructuring trends across the nation that have emerged in response to management's need to reduce operating costs and increase efficiency center primarily around services that share joint support systems and can easily cross-train staff. The most common specialties involved are general chemistry and hematology, since both have simpler, “one order-one result” work-flow patterns. Some laboratory management structures have consolidated their chemistry and hematology sections into one service, named “chematology.” The concept is sound, because the work processes, time for test result production, and automated nature of the “one order-one result” lends itself to merging these functions. Further, the high volume of screening and monitoring tests most needed by providers to establish a diagnosis and to monitor the progress of treatment are performed by chemistry and hematology, more than any other specialty section of the laboratory.

Table 2.10 illustrates the work complexity of the laboratory's specialties and subspecialties, separating the “one order-one result” specialties from those that require two to six complex processes before a usable test result is produced. Directors and managers desiring cost reduction need to consider consolidation of

sections and workstations where there are similarities in both work flow and work processes, even if the instrumentation differs greatly from section to section. The basis for this action is the more efficient and effective use of the labor time of the staff needed to produce the test output. Whenever test volume can be increased by using the same number of staff, the variable costs for reagents and supplies and labor will increase, but improved productivity and efficiency of the staff will decrease the cost per test produced. If the cost per test is multiplied by the increased volume of *billable* tests (tests for which reimbursement can be claimed) produced, and the laboratory's billing system accurately captures and submits a claim for all billable tests, then the laboratory gains revenue per test because the variable cost has been reduced.

Current Trends for External Reorganization

Unless the laboratory is independently owned and operated, there is often little control of the decision-making process that decides to consolidate, merge, acquire, or integrate laboratory functions across a health care system network, or the continuum of care in which the laboratory is involved. However, there is one area where the laboratory director and manager can have a profound effect to improve efficiency and reduce costs for the organization, because medical management's complex characteristics require special laboratory expertise to assure quality patient care delivery.

This area, with great potential for reorganization and consolidation, is termed transfusion, infusion, pheresis service.

Transfusion, Infusion, Pheresis (T/I/P) Service

As noted in a previous section (see Cost Management in Transfusion Medicine), this is one of the laboratory's highest areas of highest operating cost. One of the recommendations to resolve the problem of increasing costs is to consolidate the various internal and external sites in the HCO where transfusion medicine is called on to provide support for clinical services.

Support for this management approach is found in the application of business principles to hospital operations. The most important of these include:

- Consolidation and merger of similar functions performed in multiple sites
- Reduction in physical plant to decrease fixed costs
- Reduction in redundant, wasteful, and unnecessary functions

When these principles are successfully applied by managers in actual tertiary situations, they result in the following administrative actions:

- Close decentralized sites (T/I/P) used for outpatient T/I/P on inpatient care units
- [*Business reason*: Inpatient fixed costs are always greater than for T/I/P than in an outpatient setting owing to higher fixed labor and indirect support costs]
- Create a central T/I/P center in an outpatient setting for all patients requiring T/I/P who cannot ambulate or be moved from an acute care unit because of severe illness or incapacitation

REVENUE MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

To understand this section, the reader needs to review two equations:

$$\text{Revenue} - \text{Cost} = \text{Margin}$$

$$\text{Revenue expected} = \left\{\begin{array}{l}\text{Billable}\\ \text{product}\\ \text{volume}\end{array}\right\} \times \left\{\begin{array}{l}\text{Price,}\\ \text{rate,}\\ \text{charge, or}\\ \text{fee}\end{array}\right.$$

As always in an equation, the right and left sides must balance.

If there were any written rules to follow to assure the financial viability of a clinical diagnostic laboratory as a part of an organization or as an independent entity, they would probably be:

- Rule 1: Add new revenue streams.
- Rule 2: Avoid loss of revenue.
- Rule 3: Reduce operating costs while adding new revenue streams and avoiding loss of revenue.

Because there are no formalized, written rules for revenue management such as there are for methods and procedures in science and medicine, the director or manager must learn to read and interpret financial statements to understand where the clinical diagnostic laboratory ranks in the order of services within the parent organization. Directors and managers are judged by how well they save budgeted operating dollars and are ranked in order of positive or negative balances at the end of each fiscal reporting cycle.

The organization that owns or is responsible for providing funding for the laboratory's budget usually sets its own rules. They are simple rules:

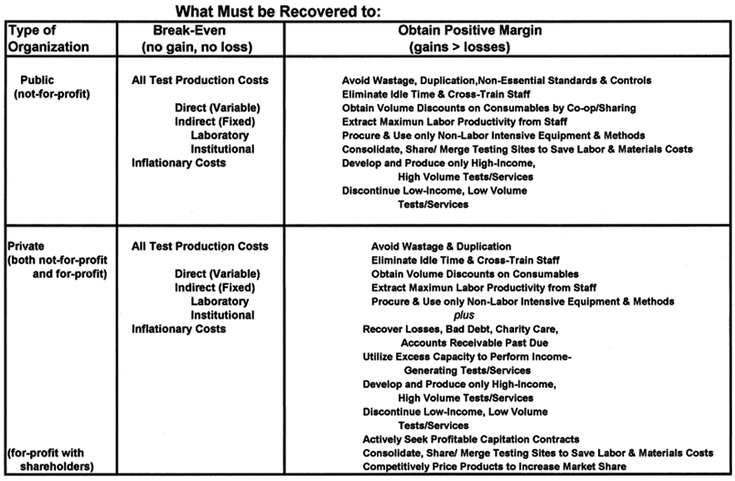
- Rule 1: Never exceed your allocated budget.
- Rule 2: Plan for unanticipated expenses when submitting annual budget requests.
- Rule 3: Never ask for “new money” in the same budget year.

MARGIN

Part of "2 - Business Management of the Clinical Laboratory"

In the Management Responsibilities section (p. 64), it was noted that in addition to knowing the full cost of the laboratory's tests and services produced, the director or manager has the responsibility to assure the organization's top management that the margin of revenue gained after the receipt of reimbursement and other revenue streams is greater than the cost of producing all the laboratory's products and services. Margin is the term applied to designate the amount left over after all operating, administrative, and selling costs are deducted from the revenue gained from reimbursement and other sources of revenue gained by selling tests and services, sharing agreements, contracts, and other business arrangements. Figure 2.6 illustrates this concept. Margin can be captured for profit-making to provide distributions to shareholders or for making contributions back to the organization that produced the products that generated the revenue for the organization.

A summary of what is necessary to assure positive margins is provided in Figure 2.7. The reader should note that there is a significant difference between the requirements for the public not-for-profit and the private for-profit organizations, where the profit-seeking goal of the private organization engenders the high costs of competition, marketing, advertising, and logistics to bring their products to market.



	What Must be Recovered to:	
Type of Organization	Break-Even (no gain, no loss)	Obtain Positive Margin (gains > losses)
Public (not-for-profit)	All Test Production Costs Direct (Variable) Indirect (Fixed) Laboratory Institutional Inflationary Costs	Avoid Wastage, Duplication,Non-Essential Standards & Controls Eliminate Idle Time & Cross-Train Staff Obtain Volume Discounts on Consumables by Co-op/Sharing Extract Maximun Labor Productivity from Staff Procure & Use only Non-Labor Intensive Equipment & Methods Consolidate, Share/ Merge Testing Sites to Save Labor & Materials Costs Develop and Produce only High-Income, High Volume Tests/Services Discontinue Low-Income, Low Volume Tests/Services
Private (both not-for-profit and for-profit)	All Test Production Costs Direct (Variable) Indirect (Fixed) Laboratory Institutional Inflationary Costs	Avoid Wastage & Duplication Eliminate Idle Time & Cross-Train Staff Obtain Volume Discounts on Consumables Extract Maximun Labor Productivity from Staff Procure & Use only Non-Labor Intensive Equipment & Methods *plus* Recover Losses, Bad Debt, Charity Care, Accounts Receivable Past Due Utilize Excess Capacity to Perform Income-Generating Tests/Services Develop and Produce only High-Income, High Volume Tests/Services Discontinue Low-Income, Low Volume Tests/Services Actively Seek Profitable Capitation Contracts
(for-profit with shareholders)		Consolidate, Share/ Merge Testing Sites to Save Labor & Materials Costs Competitively Price Products to Increase Market Share

FIGURE 2.7. Requirements for positive margins. This figure is provided as a courtesy by Dr. E.M. Travers, Annapolis, Maryland.

PRICE (CHARGE, RATE, FEE)

Part of "2 - Business Management of the Clinical Laboratory"

Although the terms noted in the title of this section are not truly synonymous and interchangable, it is important to cluster them together to indicate to the reader that they are terms currently used to designate the dollar value desired by the producer of a product that designates the expected monetary return of revenue to the producer of the product.

Each term differs in its true meaning. *Price* is usually

1. what must be given in exchange for something,
2. the value of expected revenue to be received in the future, starting with the cost of production of the test or service, and adding what dollar amount is desired to add new revenue streams,
3. the amount of money received to allow a business or organization to meet expenses and provide a positive margin of revenue above the break-even point (the point at which expenses to produce the product are equal to the revenue received for its sale)

Charge has the same definition as price.

Rate is the predetermined dollar value of a test or service to be reimbursed to the provider of the test or service, as determined by

the Board of Directors of a managed care organization (MCO), a contract negotiated with a HMO, or another contractural arrangement that specifies a predetermined maximum amount in writing. The rate usually lasts over a specified time period, as agreed on by the parties who negotiate and sign the contract.

Fee is a term usually reserved for a professional service provided, which may or may not have a preset limit, depending on the policies of the MCO, third-party payer, or HMO.

It is important to note that, before a price, charge, rate, or fee is finalized in a business relationship, the producer of the product knows that the true, actual cost needs to be established first for each and every product (test or service) to be sold. This assures that a fair and equitable return of revenue is received just to cover the product's cost of production (operating cost). This subject is discussed in the Event Costing section and is summarized in the following equations:

1. What **cost** should be recovered?

Price, charge, rate, fee or reimbursement	=	Direct costs of production	+	Indirect (fixed) costs of production	+	Indirect hosp/corp costs

2. What *margin* must be recovered and gained in price/charge?

Price, charge, rate, fee or reimbursement	=	Direct costs of production	+	Indirect (fixed) costs of production	+	Indirect hosp/corp costs	+	Margin (profit or contribution)

RISK MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

What are the risks for the laboratory? There are many new risks, in addition to losing accreditation or not passing an OSHA or FDA inspection. There are now serious risks to financial stability and the risk of being contracted out to a larger laboratory when financial or management failure is detected by top management officers in the HCO's annual financial statement review. There is also the risk that the passive, silent laboratory director or manager will not speak up and be heard in contract negotiations with MCOs and third-party payers and lose out on an opportunity to assure a better reimbursement rate for the laboratory's services.

Risk management is the management process used to protect an organization or entity from the chance of loss. Managers need to be realistic about the task of assuming risk—especially financial risk. They need to understand the complexity of accurately predicting utilization and understanding the cost of delivering clinical diagnostic services. Managers should also realize that the current trend of Medicare, other third-party payers, and MCOs toward relinquishing the danger of financial risk and turning it over to those who deliver health care services (providers of hands-on medical care) does not take into consideration the clinical diagnostic laboratory's losses when providers order tests that cannot be reimbursed.

In fact, the clinical diagnostic laboratory is rarely considered by these entities when they meet to decide the reimbursement rate for an inpatient diagnosis-related group (DRG) or outpatient current procedure terminology (CPT) procedure code. Rather than designate a portion of the reimbursed amount to cover laboratory services, the laboratory function is considered as a necessary support function, an intermediate product, or a commodity in the health care delivery process.

This is why a commitment to the development of a dedicated management plan to manage the clinical diagnostic laboratory's global risk is so important. No one outside the professional organizations that represent a laboratory's professionals is concerned at all about its risk—except the person(s) in charge—the manager and his/her staff. Therefore, the director and manager are responsible for following the notices of meetings and contract negotiations that will involve the laboratory. A powerful voice representing the laboratory's interests must be heard in the boardroom. Otherwise, the rate setting will be conducted by the HCO's fiscal and contract negotiators. This means that the cost of providing laboratory services will not be accounted for in the reimbursement rate unless the laboratory representative can produce data on the true and accurate costs of testing for the population to be served in the managed care or capitation contract being negotiated.

Managers that can demonstrate the following characteristics are more likely to succeed at assuming risk:

- *Control over a comprehensive set of tests and services*. The more services that a laboratory contracts for, the less control it has over the production, accuracy, cost, and use of services.
- *A focused management infrastructure*. Managing risk effectively means devoting adequate resources to the onerous tasks of utilization management, quality assurance, information systems management, and contracting with providers and MCOs in your community. Funding up front for these tasks is necessary to realize a return on investment in the future. Managers must have sufficient understanding of the local market environment to be assured there is enough business in the community to justify the expenditure.
- *Effective relationships with physicians and employees*. Physicians and other nonphysicians who are hands-on providers of medical care will continue to control the use of your laboratory's tests and services unless you are able to build relationships with them to advise and educate them about appropriate utilization and costs of tests and services. The successful manager will get both providers and key clinical and laboratory employees involved in all aspects of risk assumption, share economic incentives for its success, and listen and act on their ideas and recommendations for its success. The ultimate goal is to reengineer the work process to reduce unnecessary utilization by clinical users of the laboratory.
- *Knowledge of costs*. If a manager does not have a baseline from which to measure cost, it will be impossible to determine whether profit or loss has occurred after delivery of a desired level of production.
- To improve productivity in laboratory operations, whether it be at the level of the workstation or at the section, division, or department level, the manager's knowledge of the unit cost per test or service makes the difference between a work area that is financially successful and one that is not.

- *Negotiating savvy*. Many directors and managers underestimate their marketplace strength and must understand that payers compete for business based on reduced risk, healthier patient populations, and their ability to use their “black box” editing software to deny claims for laboratory tests and services. Value and quality are important, but more important is the willingness of the laboratory's contracting officials (or the HCO's officials who speak for the laboratory) to be willing to accept a 20% lower rate than what it actually costs to produce a test or service.

REIMBURSEMENT MANAGEMENT

Part of "2 - Business Management of the Clinical Laboratory"

The increased threat to HCOs in the late 1990s from the federal government's “Operation Restore Trust” efforts to prevent fraud and abuse in Medicare intensified the importance for clinical laboratories of avoiding questionable, missing, erroneous, or faulty documentation practices and understanding correct coding use for filing claims for reimbursement. The following discussion of the categories documentation, coding systems, and billing systems represents the primary areas of management control and oversight for managers and administrators responsible for correct reimbursement practices.

Documentation

In the late 1990s, the slogan “if it wasn't documented, it didn't happen” became the most frequent admonition of third-party payers in government and private-sector organizations. The motive of payers responsible for providing reimbursement (payment) to HCOs and providers for the use of their services, facilities, and raw materials was to avoid potential litigation and penalties for incorrect documentation of medical care activities.

Payers reacted strongly to the news of stiff federal fines paid by three major U.S. HCOs for violating federal Medicare regulations. Scathing indictments of these incriminated HCOs followed the fines, and compliance programs emerged throughout the United States to train all medical care providers and administrative and billing persons, not just laboratory staff, in correct coding and billing.

Payers constitute the “business” of medicine and follow the principles established by the financial and accounting professions to ensure the survival of their organizations. They too face severe civil and criminal penalties if they allow laboratories and health care providers to submit incorrect, inaccurate, or not medically necessary claims for reimbursement.

Coding Systems

CPT Codes

Laboratories are currently totally dependent on the CPT coding system (12) to receive reimbursement for the tests, services, and procedures they perform; however, there are major deficiencies in the coding system that penalize laboratories and other entities that depend on inaccurate existing CPT code numbers and descriptions of tests to claim reimbursement.

Coding of CPT claims for reimbursement is somewhat of an arcane science in which a syntax-based standard word structure created by medical publishing house committees is combined with structured, hieroglyphic-like numerical sequences. This rigid, inflexible nomenclature/code system is endorsed and mandated by the federal HCFA, without frequent scientific review by recognized bodies such as the World Health Organization, the reviewing body for ICD-9 diagnosis codes. In fact, the Evaluation and Management series of CPT codes has been roundly criticized for its inadequacies and inaccurate representation of the actual work performed by physicians when evaluating and managing patients (31).

Lack of frequent review by the committees established by the AMA to oversee the CPT coding methodology (in conjunction with HCFA) leaves the laboratory profession without a court of appeal when either:

1. new tests and procedures are approved by the FDA, yet have not been approved by the HCFA for code assignment (preventing reimbursement) (this often leads to a 2-year delay and/or a loss of operating expenses when a laboratory test or service is performed that is not assigned a CPT code),
2. old, obsolete tests and procedures are no longer used and cannot be quickly or easily removed from the coding book.

ICD-9 Codes

Similarly, the current ICD-9 coding system has flawed, incomplete categories. However, there has been a significant effort contributing to a massive and comprehensive reorganization and revision of the ICD-9 coding system. This new system, ICD-10, was introduced in the year 2000.

Its entire numbering system was replaced with both alphanumeric and Arabic number designations, which means that for all diagnosis codes used in 1999 to claim reimbursement, there now has to be an entirely new process taught to the laboratory's claims, coding, and billing staff. The positive attributes of the ICD-10 system are its greater flexibility and more comprehensive nomenclature of diseases, operations, conditions, and classifiers.

Laboratory directors and managers are required to support new training efforts for the staff after the ICD-10 is released.

Billing Management

The subject of billing management could easily require an entire textbook to explain its complexity and difficulty. Simply stated, it is critical that a laboratory director or manager instruct the claims and billing staff to never send out a claim for reimbursement if they are not absolutely sure that:

- the documentation provided on the order form by the provider of medical care is totally correct and agrees with the documentation in the progress notes of the medical record,
- the diagnosis (ICD) for the patient for whom the test or service was performed matches the CPT code for the test, procedure, or service performed by the laboratory,
- the provider has certified that the test, service, or procedure was medically necessary,
- the provider who orders the test, service, or procedure is a certified attending physician and not a teaching physician, medical student, intern, resident, or fellow.

The reason for these safeguards is that the laboratory may be penalized by the federal government with monetary and/or civil penalties if there is any discrepancy in the items noted above.

Compliance Management

This is an entirely new category of management, necessitated by the federal government's "Operation Restore Trust" effort started in the late 1990s to correct the widespread practice of fraud and abuse in claiming reimbursement for work never performed for a patient. As a direct result of federal indictments and documented incidents caused by major laboratory corporations and HCOs involving laboratory tests, even as early as 1990, the federal government instituted new regulations to prevent future false claims. In addition, the National Correct Coding Initiative in 1998 standardized the software for all 4,000 to 6,000 third-party payers in the United States who are authorized by the HCFA to act on their behalf to disburse federal moneys through the Medicare and Medicaid programs.

This action resulted in a uniform process for review and unannounced auditing of any or all medical records, laboratory data, or other data available in a laboratory's or HCO's database. The audits are performed by authorized FBI agents, Department of Justice agents, or other agents of the federal government, based on random selections or whistle-blower allegations of suspected fraud or abuse in claims for Medicare reimbursement.

The auditing process is based on reviews of:

- Medical necessity
- Correct documentation and coding of medical records and laboratory test requisition completeness
- Justice
- Safety and quality
- Accreditation, inspection, and compliance with the CLIA, OSHA, and FDA regulations

SUMMARY AND CONCLUSION

Part of "2 - Business Management of the Clinical Laboratory"

Because the nature of the business learning process is so different from the scientific/technical, it is important that the reader understand that there is no one standard method for performing many business tasks, such as there is in many scientific, medical, and technical fields. In business, how you got there is a combination of many varied approaches. Combinations that work depend on the balance and "chemistry" of the team players in a business, who combine their backgrounds, experience, and education to make the business successful. The only rigid plan that they follow is a business plan, drawn up by them, that sets goals and a way to reach them within the business itself.

In science, medicine, and technical professions, there is often only one standardized approach to reach a goal because of the need to apply the rigors of the scientific method plus legal, regulatory, and other external influences. Although it is a different way of thinking, the participant who is just beginning to understand the value of using business concepts to better manage a clinical laboratory should be willing to accept the unscientific, loose arrangement of business survival tactics.

Mastery of the topics in this chapter is critical for directors and managers in any clinical diagnostic laboratory setting that sends claims for reimbursement to third-party payers. Sound management education constitutes the basic elements of a director's or manager's strategy for reducing costs and optimizing reimbursement—keys to survival for health care specialties and related industry suppliers and partners. These subjects are not the only important topics in a vast array of business topics that the well-informed, proficient director or manager should know, understand, and be able to apply to daily operations. As the basic sciences of biology, chemistry, and physics are the keystones to the practice of scientific, medical, and technical professions, the basic tools of management briefly summarized in this chapter are the keystones to continued survival in an uncertain health care business environment.

REFERENCES

1. Cutler P. *Problem solving in clinical medicine*, 2nd ed. Baltimore: Williams & Wilkins, 1985:98.

2. Zuger A. Are doctors losing touch with hands-on medicine? *New York Times* 1999 Jul 13:sect D, 1.

3. Iglehart JK. The American health care system: expenditures. *N Engl J Med* 1999;340:70–76.

4. Wilensky GR, Newhouse JP. Medicare: what's right? What's wrong? What's next? *Health Affairs* 1999;18:92–106.

5. Geisler E, Heller O. *Managing technology in health care*. Norwell, MA: Kluwer Academic Publishers, 1996.

6. Fuchs VE. Health care for the elderly: how much? Who will pay for it? *Health Affairs* 1999;18:11–21.

7. Graber DA, Johnson RW, Ott RF. In vitro *diagnostics in the millennium*. Minneapolis, MN: U.S. Bancorp Piper Jaffray Equity Research, 1999:5–6.

8. Auxter S. IVD industry is largest in medical technology. *Clin Lab News* 1999;13:18.

9. Woolhandler S, Himmelstein D. Costs of care and administration at for profit and other hospitals in the United States. *N Engl J Med* 1997;336:769.

10. Landauer PR. *In managed care, what is the value of laboratory services*. Chicago: Abbott Laboratories, Diagnostics Division, 1996:12.

11. Iglehart JK. The American health care system: expenditures. *N Engl J Med* 1999;340:70–76.

12. American Medical Association. *Physicians' current procedural terminology*, 4th ed. Chicago: American Medical Association, 1999.

13. Root James Hospitals. Commercial labs vie for esoteric volume. *Lab Indust Rep* 1999;8:1.

14. Steiner JW, Root JM. The battle between hospital and commercial labs. *Clin Chem News* 1999;25:2.

15. Weissman D. Dedication to research and development a must in fast-moving esoteric market [Editorial]. *Lab Indust Rep* 1999;8:5.

16. Washington G-2 Reports, industry buzz. *Lab Indust Rep* 1999;8:12.

17. Gillies RR, Shortell SM, Young GJ. *Best practices in managing organized delivery systems*. Leading Organizational Transformation Conference, U.S. Dept. of Veterans Affairs, Management Decision and Research Center, Health Services Research and Development Service, Boston, April 1996.

18. Shortell SM, Gillies RR, Anderson DA. The new world of managed care: creating organized delivery systems. *Health Affairs* 1994;13:46–64.

19. Black HC. *Black's law dictionary*, 6th ed. St. Paul, MN: West Publishing Co., 1983.

20. Costello SJ. *Managing change in the workplace*. Burr Ridge, IL: Irwin Professional Publishing/Mirror Press, 1994:54.

21. Brown DB. *Technimanagement: the human side of the technical organization*. Englewood Cliffs, NJ: Prentice-Hall, 1995:1.

22. Mustokoff M, Yecies M. America's most wanted. *Health Syst Rev* 1996;29:10–12.

23. National Academy of Sciences/Institute of Medicine. *Report of the Committee on Quality of Care*. Washington, DC, 1999.

24. Lanser EG. Outcomes and performance management: refining how health care is strategized and delivered, healthcare executive. *Am Coll Healthcare Exec* 1999;14:20–24.

25. Myers MT. Is quality of care a compliance issue? Presented at the Health Care Compliance Association Annual Meeting Workshop on “Compliance 101,” Chicago, October 27–28, 1999.

26. *Computerized Patient Record System (CPRS)*. Washington, DC: U.S. Dept. of Veterans Affairs, Veterans Health Administration, Office of the Chief Information Officer, 1988.

27. Bates DL, Rittenberg E, Rittenberg MA. What proportion of diagnostic tests appear redundant? *Am J Med* 1998;104:361–468.

28. Pearce DW. *The MIT dictionary of modern economics*, 4th ed. Boston: MIT Press, 1997:414.

29. Travers EM. Worksheet for pheresis/platelet/WBC actual cost analysis. In: Travers EM, ed. *Medical technology cost analysis*. Chicago: American Society of Clinical Laboratory Scientists, II.4, VIII.3, IX.6, IX.7, X.2, 1999.

30. Parker KM. President's message, AACC president's acceptance speech. *Clin Lab News* 1999;25:1.

31. Lasker RD. The intensity of physician's work in patient visits. *N Engl J Med* 1999;341:337–341.

3

Laboratory Safety

Gerald A. Hoeltge

Clinical laboratories are complex environments. People are surrounded by chemicals, specimens, equipment, and technology. The chemicals may be toxic, specimens may be infectious, equipment can malfunction, and any technology may be misunderstood. Most technological processes and laboratory procedures expose the operator to the risk of injury. Every laboratory hazard can be contained. No clinical laboratory can afford to operate without a safety program, and the best build safety into the culture of the laboratory.

An organized safety program begins with a written plan that defines the scope of the program, its goals, and the individuals responsible for it (1). The administration must provide the necessary resources including time, equipment, and assistance. There should be a comprehensive document that summarizes the laboratory's current safety policies (2). Personnel need to be familiar with its contents. Individual technical procedures may define the specific safety precautions that are relevant to each analytic method. For example, the procedure for the preparation of bacteriologic media should cross-reference the portions of the safety manual that describe the operation of the autoclave as well as specify the types of gloves and face protection needed during the sterilization of culture media. The phlebotomy manual should be specific about the use of gloves and the method for discarding used needles. The measures for the control of hazards associated with flammable and caustic reagents that are handled may be usefully expressed within the chemistry laboratory's analytic procedures manual. All documents that list safety policies and procedures should be reviewed on a regular basis. Annual review is suggested.

One of the most practical ways to ensure regular review of one's safety program is through the establishment of a laboratory safety committee that includes representatives from each major analytic discipline. The membership should reflect the degree of experience and involvement of the average worker in the facility; that is, the committee should be composed primarily of bench-level technologists. All members of the committee should lead by example and promote the principles of safe work practices. The most effective safety programs are found in institutions that value and reward safe work practice.

The appointment of a safety officer (SO) is appropriate for larger laboratories and may be required by regulation (3). This individual helps administer safety policy and serves as an advisor on the management of safety-related matters. The SO may be responsible for maintaining accident records, conducting or coordinating the safety training program, and surveying the workplace for hazards. The SO or another official in the organization sees to it that the facility is in compliance with federal, state, and local regulations.

DEFINITIONS

Part of "3 - Laboratory Safety"

Hazard containment is the system of routine control processes that keeps the incidence of accident and injury within reasonable expectations. Compliance with the appropriate and established safety policies of the facility should limit the risks that are associated with clinical laboratory practice. A determination of the potential dangers that require containment is part of a hazard assessment program. After a safety audit has identified opportunities for improvement in safety policies, new controls may be instituted as part of a hazard abatement program. Many efforts toward hazard abatement are instituted after investigation of a reported accident or after the external review of the facility by an accrediting or licensing agency. A hazard minimization program is a continuous and regular effort that seeks to reduce risk to a level as low as can reasonably be attained. Programs such as these should not be confused with risk management activities, which are focused attempts to reduce a facility's indemnity expenses and liability. Hazard containment, abatement, and minimization programs are directed toward the goal of laboratory safety; risk management strives to mitigate the consequences of unsafe action.

Control systems can be grouped into three broad categories. Engineering controls include all the features and systems that are built into the design of the facility and its fixed equipment. Examples include fume hoods, biosafety cabinets, steam pressure regulators, electrical circuit breaking switches, and the safety devices that are part of the design of technical tools and movable equipment. Sheaths to protect lancets (4) and phlebotomy needles, replacement of glass capillary tubes with shatterproof plastic, centrifuge lids that lock when the motor is energized, immobilizing chains for storage of compressed gas cylinders, and double-walled containers for shipment of etiological agents (5) are additional types of engineering controls. Personal protective devices are barriers from potential hazards that the employee must select before they will be effective. Gloves, respirators, and mechanical pipetting aids are obvious examples. Safe work practices encompass learned skills and techniques such as the avoidance

of aerosol formation, segregation of hazardous wastes, and the proper operation of hazardous equipment. An engineering control is always preferred over any approach that requires a deliberate choice on the part of the worker.

PATIENT AND VISITOR SAFETY

Part of "3 - Laboratory Safety"

Access to hazardous areas should be limited to individuals who have been trained to avoid injury. Patients and visitors to the laboratory should not be allowed free access to technical areas. The laboratory design should incorporate a reception and waiting area that protects untrained individuals from any appliance, fume, or aerosol that may injure them.

The phlebotomy area should be designed with patient safety in mind. The patient arriving in the laboratory for a blood specimen collection should be escorted to a comfortable chair that includes an armrest and a passive restraint system to protect against syncopal falls. Ammonia inhalants should be within easy reach of the phlebotomist.

The patient should be asked about iodine allergies if iodophors are to be used as skin antiseptics. Each step of the procedure must be announced to avoid the patient's making involuntary movements. All needles, lancets, and capillary tubes should be kept out of the reach of patients. Needles and syringes should be kept under lock and key whenever unattended.

Gloves must be worn at all times during the phlebotomy and when handling the resultant specimen. A secondary container, such as a sealed plastic bag, is mandatory whenever the exterior of the container is soiled with blood. All used materials must be properly discarded. Items soiled with blood must be considered contaminated and biohazardous and placed in a properly labeled container.

Phlebotomists are more likely than any other hospital worker to be injured by a needlestick (6). Blood collection devices that minimize the risk of needlestick injury should be used in preference to those with exposed needles. Devices that connect directly to an intravenous line and ones that have retractable or sheathed needles and even self-blunting needles are available. Such devices can be highly effective in reducing the rate of needlestick injury (7).

As a general rule, used, exposed needles are never to be broken nor recapped. Clipping or breaking of needles can spatter blood and spread the area of contamination. Resheathing risks puncture injury of the hand that holds the cap. Although there may be acceptable alternatives, such as specially engineered caps (8) or a one-handed recapping technique (9), the simplest option is to drop the entire collection device into the waste container that is reserved for sharps. (Containers for sharps should never be more than two thirds full to avoid the temptation to force one more needle into the container.)

All spills and spatters of blood or body fluids must be cleaned with aqueous detergent, after which the contaminated surface must be decontaminated. Papers that may have become bloodstained must be replaced with clean copy. Gloves are to be changed regularly and whenever visibly contaminated. Washing hands between every glove change and is encouraged because it reduces the risk further. Regular use of skin emollients helps prevent chapping and fissures in the skin.

ATTIRE

Part of "3 - Laboratory Safety"

The proper attire for laboratory work minimizes exposure to infectious and chemical agents. A distinction is made between personal attire [including uniforms and cloth laboratory coats (lab coats)] and personal protective equipment (PPE). Except in areas where the risk of contamination is extraordinary (such as necropsy), personal attire may be worn in all areas of the facility and outside the building. PPE is designed to protect the employee and his or her personal attire from harmful exposure. Each item of PPE (whether gloves, face protectors, gowns, aprons, or shoe covers) must be constructed of a material through which the infectious or chemical agent of concern cannot penetrate. Protective equipment that is sufficiently damaged to permit entry of harmful materials must be replaced immediately. All items of PPE are removed before leaving the laboratory.

Attire should be matched to the task. For example, phlebotomists must be identifiable as laboratory professionals. A clean, neat uniform can help reduce a patient's anxiety. At the same time, the technician must protect him- or herself from the chance that blood might spray or leak from the collection device. A long-sleeved, buttoned lab coat worn with disposable gloves meets both needs.

The performance of a necropsy demands a higher level of containment and therefore more effective personal barrier protection. Water-resistant gowns that extend from the neck to below the level of the table are needed. The sleeves should be tight at the wrist and tucked inside vinyl or latex gloves. Facial protection (goggles and mask or full face shield) and shoe covers are necessary. Slash-resistant gloves made of steel mesh or a specially selected polymer may be needed (10).

Most tasks in the clinical chemistry laboratory can be safely performed while wearing the same laboratory attire that protects workers elsewhere from common infectious hazards. A fluid-resistant lab coat, gown, or tunic combined with disposable gloves is usually sufficient. Face protection (either worn by the worker or interposed as a barrier shield) must be used whenever splashing or spraying with droplets of blood or other potentially infectious fluids is a risk. The objective is to protect all exposed skin surfaces including the neck and wrists. Some tasks will require additional protection. The handling of caustics and corrosives obligates the use of gloves, aprons, and face protectors that are made of materials known to be resistant to the liquid being manipulated. The handling of even small amounts of carcinogenic substances requires the use of gloves and forearm protection (as well as facial protection if splashing or spattering is anticipated).

Shoes should be rubber soled for effective traction and to decrease the shock hazard precipitated by an electrical accident. Shoe covers made from appropriately resistant materials are needed if splashing is a risk. In particular, individuals who must transport or decant corrosive liquids must not do so wearing unprotected cloth or porous shoes.

Between uses, reusable protective clothing should be hung away from sources of heat and potential ignition. Uncontaminated garments (such as street clothes) should be stored separately from laboratory wear. All items of personal barrier protection should be changed immediately once they are found to have been contaminated with blood or toxic chemicals. Because of the

chance of insensible contamination, protective clothing should be changed regularly. Obviously contaminated, single-use (disposable) gear should be discarded in properly marked containers. Reusable lab coats and gowns should be placed in leak-resistant bags to be laundered in a manner that ensures decontamination. To contain the hazards as much as possible, the laundering of contaminated clothing at home should be prohibited.

Hazard containment applies to hairstyles and personal adornments. Hair should be worn behind the head and off the shoulders to preclude its contact with contaminated surfaces. Disposable hair covers may be indicated for some activities. Hair, beards, and jewelry should not be allowed to overhang into contaminated areas or risk entanglement in moving equipment.

GOOD LABORATORY HYGIENE

Part of "3 - Laboratory Safety"

Handwashing is one of the most effective means to minimize personal exposure. Thorough lavage with a good detergent removes external bacteria and chemicals before they can be ingested or absorbed. Antibacterial agents are not required for a good laboratory hand detergent. Hands should be washed frequently throughout the day. Washing is especially important after removing gloves, before and after contact with patients, before leaving the laboratory, and before eating, applying cosmetics, and using lavatory facilities. The well-designed handwashing sink is operable with elbow blades or foot pedals. Paper towels can be disposed of with the general refuse.

All areas of the laboratory should prohibit eating, drinking, and manipulating items that may contact mucous membranes (e.g., contact lenses, cosmetics, and lip balm) within the technical work areas. Any such activity risks exposure to both infectious agents and toxic substances. Food and drink should never even be brought into the technical work area. Refrigerators used for storage of reagents and specimens off off-limits for food items. Dining areas and refrigerator are to be reserved exclusively for food items.

Individuals who wear contact lenses must take special precautions. Any activity that might splash a toxic, irritating, or infectious fluid into the eyes presents a special hazard for contact lens wearers. Tightly fitting contact lenses can inhibit the naturally protective effects of tearing, permitting substances to contact the conjunctivae for longer periods of time. Irritants can induce immediate inflammation and swelling, rendering the contact lens difficult to remove. Scleral injection hastens the absorption of toxic substances. Although some may prohibit the wearing of contact lenses in the laboratory, goggles or face shields provide an acceptable level of protection. Special precautions also need to be taken by lens wearers who are manipulating solvents that might dissolve the plastic of the corrective lens. In formulating a contact lens policy, it must be remembered that individuals who do not wear such lenses are also at risk for eye injury; the safest strategy is to require the wearing of full facial protection for all employees engaged in tasks that chance splashing or spraying.

Mechanical devices should be always used for pipetting tasks. The pipetting of fluids by mouth must be strictly prohibited throughout the laboratory. Even innocuous fluids such as water should be pipetted mechanically for consistency.

SPACE AND VENTILATION

Part of "3 - Laboratory Safety"

The laboratory should be allotted sufficient space in which all work processes can be done safely (11). The arrangement of equipment on countertops should leave plenty of working area. Having to reach across hot plates or to brush past containers of caustics as a routine invites an accident. Power cords or hoses that cross walkways indicate poor and unsafe laboratory design. Lighting must be bright enough to illuminate all areas of the laboratory.

The laboratory should be designed to accommodate the special needs of handicapped workers. Ramps, wide aisles, and lowered work surfaces are needs of wheelchair-bound workers, for example (12).

Air handling must be engineered to the task. Fume hoods and biosafety cabinets need to be available in appropriate locations. Operations on open benches that generate occasional or low-level fumes should be positioned such that fumes are drawn away from workstations and not across them.

HOUSEKEEPING

Part of "3 - Laboratory Safety"

Cleanliness and good housekeeping are attributes of the best laboratories. Proper storage of laboratory implements when not in use keeps work areas clear to facilitate disinfection, decontamination, and hazard containment. A cluttered aisle is a nuisance that can lead to falls and lower leg injuries. Accumulations of chairs, equipment, and boxes that block exits or otherwise interfere with emergency egress must be forbidden. Smoke doors must remain completely unobstructed so that they will close automatically in the event of fire. Free access to all emergency equipment (such as portable fire extinguishers, showers, and eyewash fixtures) must be ensured at all times. Clothing, signs, and decorations should never be hung over doorways or safety windows in any manner that obstructs the view. Floors and walls are to be cleaned on a regular schedule. All cabinetry, including refrigerators, freezers, incubators, and water baths, should be cleaned and disinfected regularly (usually at the end of each work shift) as well as immediately after spills.

PERSONAL BARRIER PROTECTION

Part of "3 - Laboratory Safety"

Gloves

Gloves, the most commonly used personal protective device, protect the wearer from contact with noxious substances and dangerous materials. Gloves should fit as comfortably as possible. Glove length must be sufficient to extend over the cuffs of the sleeves to protect the wrists.

Heavy rubber gloves made of neoprene, nitrile, or butyl rubber are designed for chemical protection. Choose a material known to resist the specific chemicals to which the worker will be exposed. Glove materials are rated (in hours) as to the duration

of protection provided and/or the rate of chemical permeation through the fabric.

Thin latex or vinyl gloves provide short-term protection against dusts and contaminated aqueous solutions. Minute and insensible defects can develop easily in such gloves, and viruses can slowly penetrate intact glove wear. Such gloves must be changed and discarded regularly to be effective (13). Thin gloves are designed to be single-use devices and must be replaced immediately when torn or contaminated.

Natural latex rubber is an allergen. Cornstarch and talc particles increase the hazard by assisting the delivery of latex proteins to respiratory epithelium. Thin, flexible gloves made of vinyl plastic should be offered as an alternative for those employees (and patients) who are latex sensitive. Powdered latex gloves should be prohibited from the workplace, and reduced-protein latex gloves made available for routine work (14).

Heat-reflective and insulated gloves protect against thermal injury when handling items at blistering temperatures or when working in frigid conditions. Cryostat operators can develop chronic peripheral neuropathies and Raynaud's phenomenon if inadequately protected (15).

Stainless-steel mesh gloves provide shielding when handling sharp items such as knife blades and broken glass.

Outerwear

Lab coats are long-sleeved jackets of sufficient length to extend below the level of the workbench when the worker is standing and to fully cover the individual's lap as he or she sits at the workstation. Lab coats should always be worn fully buttoned and should be made of water-resistant fabric. Coats that are constructed of absorbent cloth must usually be overlaid with a plastic apron or tunic. Gowns provide a greater measure of protection. Gowns fit snugly at the wrists and neck and have either no front openings or a fully fastened front opening. Jumpsuits are fitted to the ankles as well as the wrists; they are appropriate in areas where hazardous dusts or liquids may contaminate stockings or trousers. Water-resistant garments are to be used when performing tasks that risk splashing with infectious or otherwise potentially infectious fluids. When chemical exposure is anticipated, outerwear must be constructed of a material that is resistant to the chemical that is handled. All single-use, protective outerwear should be changed as soon as possible after contamination; PPE designed for reuse should be cleaned or decontaminated as soon as possible after soiling.

Eye and Face Protection

Safety glasses are spectacles made of high-impact plastic; imperforate side shields at the temples are standard. Safety glasses offer a minimal and usually inadequate level of protection. Goggles, which fit snugly around the face, are preferred and should be combined with masks that protect the mouth and nose. A face shield, a shatterproof plastic device that wraps around the face to safeguard all exposed facial skin (including the neck), provides more complete protection. Small, freestanding area shields rest on or are suspended over the workspace, allowing the technologist to manipulate hazardous fluids with an acceptable level of eye and face protection. Activities that risk splashing or spraying with infectious materials or fluids that may irritate mucous membranes should be performed behind shields or while wearing goggles and a mask. Corrosive liquids must always be manipulated with shield protection.

Respiratory Protection

Many types of devices are manufactured for respiratory protection. Such devices must be used to prevent the inhalation of fumes, dusts, or aerosols that may be harmful. The use of properly designed air-handling cabinetry (such as chemical fume hoods) is best, but personal protective equipment should be used whenever engineering controls may offer inadequate protection.

The choice of a device for respiratory protection is based on the nature of the hazard that is targeted for control. Dust masks are simple cloth or paper devices that are nonetheless effective in preventing the inhalation of aerosols and particulates when properly fitted. Fume masks cover half the face and employ replaceable cartridges that contain absorbent chemicals. Each cartridge, chosen to reduce the exposure level of a specific group of noxious fumes, must be changed periodically. An airline or self-contained breathing apparatus (SCBA) offers the greatest level of protection because all inspired air is delivered from an uncontaminated source. Each device must be used in accordance with its manufacturer's directions.

Respirators that fit tightly around the mouth and nose impede ventilation. Although their use is sometimes necessary, only employees who are physically able to breathe properly through the device should be allowed in a work environment in which their use is required. Safe use of the more complicated devices may depend on proper fit and training.

An SCBA is the device of choice whenever no simpler device would suffice to protect the worker. Examples include levels of toxic vapors that exceed safe limits, vapor hazards of unknown toxicity, deoxygenated atmospheres (<19.5%), or any other acute hazard to life or health (16).

Chemical fume hoods that are ducted to the roof of the building provide an engineering alternative that is better than relying completely on personal protective equipment. The rating of the device is based on the airflow velocity at the face of the work area. Airflow is typically controlled by means of a sash made of laminated safety glass. A face velocity airflow of 80 to 120 feet per minute (fpm) is recommended (17). A perchloric acid hood is a special type of chemical fume hood and ducting system that is constructed of corrosion-resistant stainless steel. Some limited tasks that generate minimally hazardous but noxious fumes may safely be performed on a downdraft platform or at a rear-plenum exhaust workstation; both systems draw vapors away from the face of the operator. The absorbent chemical filters of such devices are similar to those in portable fume masks.

WARNING SIGNAGE

Part of "3 - Laboratory Safety"

Regulatory and Advisory Systems

Several systems of signs have been devised to warn individuals of hazards and potential injury. Each addresses a specific target audience,

and a combination of several systems will be necessary for most laboratories.

The Occupational Safety and Health Administration (OSHA) has defined one system for area signs for accident prevention (Fig. 3.1) (18). The colors of the signs indicate the hazard level. Red letters indicate a dangerous location, the nature of which is identified with black letters on a white background. A yellow background with black letters prompts one to use caution. Green signs are advisory and display helpful safety information. Pictographs may be included. By virtue of their clarity and simplicity, OSHA-style signs convey minimal detail. They are intended to alert anyone who is approaching the area, even those with little training. Additional information (such as emergency procedures) may need to be posted nearby.

FIGURE 3.1. Informational signs for accident prevention formatted according to regulations of the Occupational Safety and Health Administration.

The National Fire Protection Association (NFPA) developed a hazard identification system (Table 3.1) specifically for fire fighting (Fig. 3.2) (19). A diamond-shaped emblem is divided into four fields that are color coded to distinguish the type of hazard. Three of the fields display a numeral from 0 (negligible hazard) to 4 (severe hazard). At the pinnacle is a red diamond that indicates the flammability risk of the chemicals stored in the room or cabinet. On the left is a blue diamond for the health hazard posed by acute exposure through inhalation or dermal contact to those chemicals or to their combustion products. On the right, a yellow diamond reflects the instability risk. These are relative rather than absolute ratings and are assigned according to the professional judgment of the SO. The white diamond at the low point on the emblem may have a W with a line through it (do not use water) or the letters OX (oxidizers present). It is important to understand that NFPA 704 is designed for the protection of fire fighters. It is an appropriate system to post at eye level in entryways and on cabinet doors. It is not a good system for individual reagent vessels because the effects of chronic exposure to a chemical are not considered in the health rating. NFPA 704 should be used only with the consent and cooperation of the local fire department.

FIGURE 3.2. Example of fire hazard warning sign (from the National Fire Protection Association standard 704).

The U.S. Department of Transportation (DOT) requires the use of a different signage system (20). Partly pictorial and partly numeric, DOT signs are usually found on cartage vessels, cartons, and containers. The diamond-shaped signs can be useful when choosing the proper response to contain and decontaminate a leaking and otherwise unidentified shipping container.

Additional emblems have been promoted for specific tasks. Figure 3.3 displays standardized symbols that may be helpful in clinical laboratories.

The labels on reagent vessels should clearly delineate the hazards of the contents and their adverse effects on specific target organs. The identity of the chemical is best described in accordance with the International Union of Pure and Applied Chemistry or the Chemical Abstracts Service rules of nomenclature. The label may also specify the personal protective equipment required. Commercially labeled reagent containers often display additional information such as advice on first aid and spill cleanup. Secondary containers into which the contents are transferred should always be labeled with the key safety information. The only exception to this rule is when the individual transferring the chemical is the sole user of the chemical.

Warning labels are not to be removed or obscured unless a revised label will be applied immediately. Unlabeled containers should remain unopened and discarded as an unknown chemical that is presumed to be hazardous.

It is good laboratory practice to indicate on the label the date that each container was opened. Such a practice is also a safety precaution when dealing with chemicals for which the hazard level increases upon exposure to air. For example, diethyl ether slowly but spontaneously can form explosive peroxides when mixed with atmospheric oxygen; because of this, storage must be brief after opening.

CHEMICAL HAZARDS

Part of "3 - Laboratory Safety"

The smaller the quantity of hazardous materials that a laboratory has on hand, the better. The hazards of a laboratory chemical can be described in terms of ignitability, corrosivity, instability, and toxicity. The characteristics of many chemicals belong to more than one hazard class. Containment strategies are based on the nature of the hazards involved. A chemical hygiene plan (CHP) is a strategic approach to reagent and waste chemical hazard control (3). The CHP specifies the safety precautions appropriate for each class of chemical that is in use in the laboratory. The goal is to promote the combination of engineering controls and safe work practices that best protects employees from the hazards of the chemicals that they are using.

Exposures to fumes must be kept within permissible limits (see the section on toxicity). The CHP will identify (a) the criteria that will be used to determine the need for special control measures (such as personal protective equipment), (b) the requirements for chemical exposure monitoring (i.e., anticipation of circumstances during which permissible exposure limits may be exceeded), and (c) the provision for medical attention when dangerous exposures do occur.

Corrosivity

The U.S. Environmental Protection Agency (EPA) defines corrosives based on pH or the ability to destroy steel. Any aqueous waste that is extremely acidic (pH 2.0) or extremely alkaline (pH

12.5), or that can corrode SAE 1020 steel more than 6.25 mm (0.250 in.)/year at 55°C is corrosive (21). As a warning on a reagent label, however, the term may be generalized to refer to any substance that can cause visible destruction or irreversible alteration in human tissues at the site of contact.

Corrosives should always be stored near the floor and below eye level. A catchment basin on the floor near the point of use and the regular use of protective bottle carriers minimizes contact with a spilled chemical. Mutually incompatible chemicals, e.g., hydrocarbons, including organic acids, and strong oxidizers, such as sulfuric, nitric, or perchloric acid, must be stored separately. Personnel must wear the proper equipment when handling corrosives. Gloves, aprons, and other protective garments should be constructed of materials that are known to be resistant to the corrosives in use. Vapors of corrosives are often highly irritating, and the pouring and pipetting of concentrated corrosive

TABLE 3.1. HAZARD RATINGS FROM NFPA 704

Identification of Health Hazard Color Code: Blue		Identification of Flammability Color Code: Red		Identification of Instability Color Code: Yellow	
Signal	**Type of Possible Injury**	**Signal**	**Susceptibility of Materials to Burning**	**Signal**	**Susceptibility to Release of Energy**
4	Materials that, under emergency conditions, can be lethal	4	Materials that will rapidly or completely vaporize at atmospheric pressure and normal ambient temperature, or that are readily dispersed in air and will burn readily	4	Materials that in themselves are readily capable of detonation or explosive decomposition or explosive reaction at normal temperatures and pressures
3	Materials that, under emergency conditions, can cause serious or permanent injury	3	Liquids and solids that can be ignited under almost all ambient temperature conditions	3	Materials that in themselves are capable of detonation or explosive decomposition or reaction but that require a strong initiating source or must be heated under confinement before initiation
2	Materials that, under emergency conditions, can cause temporary incapacity or residual injury	2	Materials that must be moderately heated or exposed to relatively high ambient temperatures before ignition can occur	2	Materials that readily undergo violent chemical change at elevated temperatures and pressures
1	Materials that, under emergency conditions, can cause significant irritation	1	Materials that must be preheated before ignition can occur	1	Materials that in themselves are normally stable but that can become unstable at elevated temperatures and pressures
0	Materials that, under emergency conditions, would offer no hazard beyond that of ordinary combustible material	0	Materials that will not burn	0	Materials that in themselves are normally stable, even under fire-exposure conditions

Adapted from NFPA 704. *Standard system for the identification of materials for emergency response*. Quincy, MA: National Fire Protection Association, 1996, with permission. This reprinted material is not the complete and official position of the National Fire Protection Association on the referenced subject, which is represented only by the standard in its entirety.

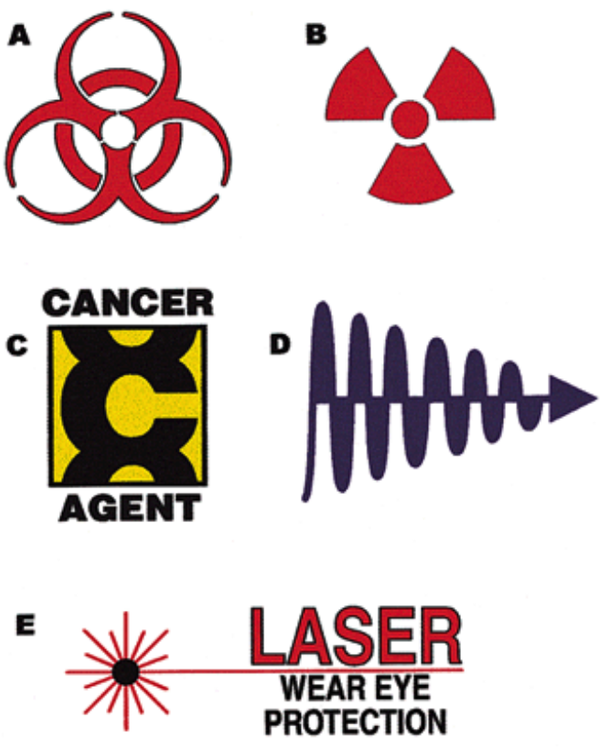


FIGURE 3.3. Standard symbols for biohazards **(a)**, radiation hazards**(B)**, carcinogenic hazards**(C)**, ultraviolet light hazards**(d)**, and laser devices**(E)**.

liquids within a chemical fume hood are highly recommended. The area of the laboratory in which corrosives are used must have an adequate source of water, such as an emergency shower and eyewash station, for rapid decontamination.

Ignitability

The NFPA developed a useful classification system based on hazard potential (13,22). Flammable liquids are those with a very low flash point (the temperature at which a liquid gives off vapors in sufficient concentration to form an ignitable mixture with air) or boiling point (at normal atmospheric pressure) or both. Flammable liquids are divided into three groups. Class IA designates the most flammable liquids:

- Class IA: flash point <22.8°C and boiling point <37.8°C
- Class IB: flash point <22.8°C and boiling point ≥7.8°C
- Class IC: flash point ≥22.8°C and boiling point <37.8°C

For example, diethyl ether (with a flash point of -45°C) and benzene (with a flash point of -11.1°C) are class IA liquids. Absolute ethanol has a flash point of 12.8°C and is a class IB liquid.

Combustible liquids have a flash point that is greater than 37.8°C. Combustible liquids are also classified into three groups:

- Class II: flash point ≥37.8°C and <60°C
- Class IIIA: flash point ≥60°C and <93.3°C
- Class IIIB: flash point ≥93.3°C

As an example, 37% aqueous formaldehyde (the solution usually referred to as undiluted formalin) has a flash point of 85°C and is therefore grouped within NFPA class IIIA.

The value of this classification system can be seen in Table 3.2. The ignitability risk of a stored chemical is a function of the nature of the substance, the type of storage vessel that contains it, and its shielding from sources of ignition. Specially designed flammable solvent containers are safer storage vessels than glass bottles. A safety can is made of metal or polymer. It has a spring-closing cap, a pressure-release valve, and a flame arrestor within the spout. Storing flammable liquids in an ordinary glass vessel is a greater hazard than storing the same amount of material in a safety can.

TABLE 3.2. MAXIMUM ALLOWABLE SIZES OF CONTAINERS OF IGNITABLE LIQUIDS

Ignitability Class	Glass	Metal or Approved Plastic	Safety Can
IA	1 pt (0.473 L)	1 gal (3.8 L)	2 gal (7.9 L)
IB	1 gal (3.8 L)	5 gal (19.0 L)	5 gal (19.0 L)
IC	1 gal (3.8 L)	5 gal (19.0 L)	5 gal (19.0 L)
II	1 gal (3.8 L)	5 gal (19.0 L)	5 gal (19.0 L)
III	5 gal (19.0 L)	5 gal (19.0 L)	5 gal (19.0 L)

Adapted from NFPA 30. *Flammable and combustible liquids code*, 1996. Quincy, MA: National Fire Protection Association, 1996, with permission. This reprinted material is not the complete and official position of the National Fire Protection Association on the referenced subject, which is represented only by the standard in its entirety.

These definitions are derived from the National Fire Code. Regulatory agencies may use slightly different definitions. For example, the EPA uses the term *ignitable* to refer to liquids that have a flash point less than 60°C, are strong oxidizers, or have some other property that contributes significantly to a fire hazard (such as the capability to cause a fire through friction, absorption of moisture, or spontaneous chemical change) (23).

A flammable storage cabinet (FSC) is designed to sequester its contents from heat during a conflagration. The purpose is to allow personnel a short amount of time (approximately 10 minutes) to evacuate the laboratory before the liquids inside ignite. An FSC can be constructed of metal or wood. The door should fit tightly and preferably be self-closing. An FSC may be either entirely enclosed or vented to the exterior of the building. Metal FSCs must be properly grounded. The volume of ignitable liquids that can be stored both outside such a cabinet and within an FSC is limited according to whether the laboratory is located within a health care occupancy (such as a hospital) (24) or in a business occupancy (such as a freestanding clinical laboratory) (17). The maximum volume that is loaded into any FSC must be within the manufacturer's specifications.

A flammable storage vault is a heavily insulated room capable of storing larger quantities of ignitable liquids safely. All storage areas must be chosen to avoid open flames, radiators and other sources of heat, and electrical devices capable of sparking. Solvents are best stored in an area that is separate from the general workflow. Routine storage in a chemical fume hood is usually inappropriate because the containers will be on a working surface.

Cold liquids are less hazardous than warm liquids of the same composition because the vapor pressure is lower. Storage of volatile solvents in refrigerators and freezers is best avoided, however, because of the possibility of accumulation of vapors. When the door is opened, the concentrated vapors are released and may travel considerable distances along the floor to a point of ignition. Refrigerators that must be used for storage of flammable liquids must be explosion proof, i.e., lack interior electrical switches and have motors shielded from vapors that might accumulate within the refrigerator storage unit.

Instability

Chemicals capable of violent decomposition at normal temperatures and pressures form a separate hazard class. This term is reserved for chemicals that can readily detonate at standard temperatures and pressures. It also applies to substances that react violently or form a potentially violent mixture in water or that are capable of explosive decomposition. Examples are hydrazines and elemental sodium.

Toxicity

This term can refer to almost any substance in quantity. Aspirin, vitamin preparations, and other commonly encountered substances have all caused fatal poisonings. Even water can cause an acute intoxication if sufficient quantities are ingested over a short interval. In the context of laboratory safety, the term is generally applied if a substance that is inhaled, ingested, or contacted in small amounts can cause serious biological effects. Toxicity may be short term or long term. Irritants cause a local inflammatory effect on living tissues at the site of contact. Irritants may be identified as either primary skin irritants or as primary eye irritants or both. A sensitizer is a chemical that can elicit an allergic reaction in a susceptible individual after repeated exposure. Asphyxiants are chemicals that exert an adverse effect by displacing atmospheric oxygen or by preventing the metabolic use of available oxygen. Many gases are asphyxiating in large quantity.

Among the long-term toxins, mutagens are chemicals that can cause a heritable change in genetic material. There are many different test systems to detect mutagenic potential. Those that employ living organisms usually measure the effect of the substance on lower organisms such as bacteria, fungi, or insects. Because of this, the term mutagen is frequently applied to substances that have not been shown to cause genetic changes in mammalian species. Reproductive effects are those that affect fertility, cause developmental abnormalities during gestation, or can have adverse effects on newborns. Tumorigens are substances that have been reported to be associated with benign or malignant tumors. Carcinogens are substances that have been shown to cause malignant neoplasms in valid animal tests (using control animals and appropriate statistical methods) or that have conclusively been shown to cause malignant neoplasms in humans. Authoritative sources include the International Agency for Research on Cancer and the U.S. National Toxicology Program. Neoplastinogens cause tumors that cannot definitely be classified as either benign or malignant.

The mode of contact is an important variable in estimating the toxicity of a substance. Some chemicals may produce serious biological effects after only short-term exposure; such chemicals are termed acutely toxic. Standard tests have been developed to measure and define such acute toxins. The toxicity is usually defined by the dose required to kill 50% of test animals (LD_{50}). The route of administration and the stated dose must be provided, followed by the species of animal in which the substance was tested. The lower the concentration at which toxic effects can be measured, the more toxic the chemical. For example, LD_{50} (oral) is 5 mg/kg, LD_{50} (dermal) is 40 mg/kg, and LD_{50} (inhalation) is 1,000 ppm are descriptions of extremely toxic chemicals.

Restrictions should be applied to the use of chemicals that precipitate the most serious consequences. Chemicals of special concern include acutely toxic chemicals as well as carcinogens and reproductive toxins.

Access to areas in which such chemicals can be found should be strictly limited to those who have specifically been trained to contain the hazards. Mandatory apparel may include gloves, tightly fitting goggles, dust masks, and jumpsuits. Disposable items are preferred. All personnel must thoroughly wash their hands and any areas of skin exposure thoroughly following accidental contact with restricted chemicals.

The American Conference of Governmental Industrial Hygienists (ACGIH) is one useful source for recommendations on chemical exposure limits (25). Sources used by the ACGIH include industrial experience as well as human and animal studies. Two major types of permissible exposure limits (PELs), designated as threshold limits values (TLVs), are recognized by the ACGIH. TLVs based on 8-hour, time-weighted averages are used to denote the permissible exposures for substances to which workers are exposed repeatedly, day after day and week after week, without adverse effect. TLVs based on short-term exposure limits are the concentrations of vapors to which workers can be exposed continuously for brief periods (15 minutes) without experiencing irritation, tissue damage, or narcosis. The OSHA has defined an action level (AL) for selected substances. ALs, generally set at one-half of the 8-hour PEL, are administrative decision points that trigger environmental monitoring or other occupational health undertakings. For regulatory details on a particular chemical, one should first determine whether a specific standard exists [as, for example, formaldehyde (26)]. If not, tables of regulated chemicals exist that list PELs (27).

Chemical spill control materials should be located strategically throughout the laboratory. The simplest material is sand (often mixed with soda ash) that can be applied around a spill as a dike to contain the hazard. The CHP should specify the appropriate decontamination procedure for each class of hazardous chemical. Only emergency response personnel should be allowed in the vicinity. If the vapor level of the spilled substance is dangerous, all persons in the room must wear the proper respirators.

An important source of information about the hazards of a chemical substance is the Material Safety Data Sheet (MSDS). These are OSHA-required documents that are supplied by the vendor. MSDS information includes the chemical name and synonyms, its hazards, proper PPE, suggested medical intervention following acute exposure, and spill cleanup details. The MSDS may be the only source of information for some mixtures and proprietary formulations, but MSDSs should not be regarded as complete sources of authoritative data. Other sources are the Registry of Toxic Effects of Chemical Substances (National Institute of Occupational Safety and Health, Washington, DC) (28), TOXLINE (National Library of Medicine, Bethesda, MD), and Chemical Abstracts Service (American Chemical Society, Columbus, OH).

COMPRESSED GASES

Part of "3 - Laboratory Safety"

All compressed gases are hazardous because of the potential of rupture of the container. Many are also toxic or ignitable. The breakage of the regulating valve of a compressed gas cylinder can propel the cylinder like a missile. A heated cylinder will explode when its ability to contain the growing internal pressure is exceeded.

All but the smallest cylinders should be secured upright. Each cylinder should be labeled indelibly as to its contents. Additional, temporary labels that read “in use” or “empty” may be applied. Cylinders should be moved only with handcarts to which they have been secured. The rolling or dragging of cylinders must be strictly forbidden.

Both gas cylinders and piped systems have pressure-regulating devices. These valve systems are designed separately for each family of gases and must not be interchanged. Valve safety covers must be in place until pressure regulators are attached. Threaded fittings must be tightened carefully with the properly sized wrench. Cylinders and connections should always be tested with a soap solution for leaks after attaching, adjusting, or disconnecting the system. Valves should never be lubricated.

The typical valve system has two components: a high-pressure valve at or near the tank to release the contents and a low-pressure valve at the point of use to regulate the flow. Valves should be opened slowly, and personnel should stand to the side of the gauge as a precaution against blowout. Valves should be maintained in the closed position when the cylinder is not in active use (29).

Flammable gases must be used with special caution. Supplies should be minimal. Only cylinders in active use should be in the work area. The storage facility should be a secured room or enclosure reserved exclusively for that purpose. The location should be away from combustible materials, elevators, stairs, and passageways. Oxidizing gases must be separated from flammable gases. Sources of ignition should be carefully insulated from the emplacement of the cylinders.

Empty cylinders awaiting return to the vendor must be treated carefully because considerable pressure may remain within.

MICROBIOLOGICAL HAZARDS

Part of "3 - Laboratory Safety"

All clinical specimens are potential sources of infection. Human blood and body fluids may carry viruses such as the hepatitis B virus (HBV), hepatitis C virus, and the human immunodeficiency viruses (HIV). Control of hazards of this type is a concern throughout the health care industry (30). Workers in microbiology laboratories face the additional risks of the pathogens that are isolated and cultured.

Development of an exposure control plan (ECP) (31) designed to contain the infectious hazards of the laboratory is highly recommended. Each task or procedure that involves the potential for exposure to infectious agents is to be identified in the ECP. Engineering controls should be introduced wherever feasible to minimize the hazards. Work practice controls should be specified. Required PPE should be clearly identified in the ECP for each hazardous task.

Containment systems must be designed with an understanding of the virulence and the route of introduction of the agent into the body. Ingestion of agents may complicate mouth pipetting. Failure to wash one's hands thoroughly when leaving the laboratory before eating or smoking may introduce agents into the mouth. Parenteral exposure can result from direct inoculation or absorption through the mucous membranes of the eyes, nose, or mouth. Even small scratches and abrasions can allow entry of agents that would otherwise be excluded by intact skin. Contact leading to ingestion of agents or parenteral exposure may be either direct or indirect. Direct contact includes splashes and sprays of infectious fluids. The inadvertent transfer of contaminants from inanimate surfaces such as telephones or keyboards to the face or eyes is an example of indirect contact. Scrupulous attention to good work practices and consistent use of personal barrier protection minimizes the risk of ingestion and contact exposure.

Inhalation of aqueous aerosols and fomites can cause respiratory infections. Aerosols may be formed by dropping solutions onto hard surfaces, by the rapid heating of liquids (including moisture on inoculation loops), and by many other inadvertent motions. Containment is possible with careful technique and by performing all operations that predictably generate aerosols (such as centrifugation and sonication) within a biological safety cabinet.

Biological Safety Cabinetry

Microbiological hazards are best contained within a properly engineered biological safety cabinet (BSC) (32). Class I cabinets operate at negative pressure with a minimum face airflow velocity of 75 ft/min. The air in the chamber is exhausted through a high-efficiency particulate air (HEPA) filter. The front of the class I BSC may be either open or fitted with arm-length rubber gloves. A class II BSC features vertical airflow and interior air that is recirculated through HEPA filters. The chamber operates at a negative pressure with respect to the room identical to an open-front class I cabinet, but the purified air within minimizes contamination of the cultures under study. Class I and II cabinets provide a similar level of personnel safety. Class III cabinetry must be used for the most virulent agents. The chamber is entirely

enclosed. The contents must be manipulated by means of arm-length rubber gloves fitted to the chamber wall. All materials entering and exiting the class III BSC must be autoclaved or decontaminated.

Class I and II cabinets are commonly found in clinical laboratories. Class III BSCs are needed in specialized facilities that culture agents such as *Mycobacterium tuberculosis* or systemic fungi and when production-level quantities of less infectious agents such as human immunodeficiency virus are involved.

Biosafety Levels

The Centers for Disease Control (CDC) and the National Institutes of Health have codified a system of increasing levels of containment for microbiological and clinical laboratories (32). Biosafety level (BSL) I is designed for student laboratories using agents not ordinarily infectious for humans. Work may be conducted on open bench tops. Good laboratory practices include use of pipetting devices, prompt spill cleanup, daily disinfection, and proper waste disposal. Clinical laboratories should routinely follow BSL II precautions. BSL II differs from BSL I in that access to the area should be restricted to exclude untrained individuals and that some procedures such as those that generate infectious aerosols are to be conducted in BSCs. BSL II is effective at containing the infectious hazards of blood-borne agents in clinical laboratory specimens. Routine bacteriological procedures such as plating and preparing smears for staining may be conducted at BSL II. Parasitic investigations, most bacterial studies, and some fungal and viral culture manipulations may safely be contained with BSL II precautions. BSL III is appropriate for laboratories that work with agents that may cause fatal disease after inhalation. Access to the laboratory and airflow is carefully controlled. All procedures must be performed in BSCs or similar devices. Personnel must wear fully protective laboratory clothing. Few clinical laboratories other than those culturing systemic fungi and tubercle bacilli need to follow BSL III containment.

Universal Precautions

Most clinical laboratory activities require the handling of blood or body fluids (9,33,34). The CDC and the National Committee for Clinical Laboratory Standards (NCCLS) have developed a system for worker protection commonly referred to as universal precautions. This set of engineering controls, work practice controls, and use of PPE collectively shields the laboratory worker from the potential of exposure to blood-borne infectious agents.

The primary agents of concern are the human hepatitis and immunodeficiency viruses. The HBV is a DNA virus that, before widespread adoption of universal precautions, was responsible for approximately 12,000 infections in health care workers each year. During acute HBV, the virus is in high concentration in the blood and serous fluids (108 to 109 infectious units per milliliter) (9). One percent (or more) of hospitalized patients are chronic carriers of the virus (35). The virus can survive dried on work surfaces, and contaminated benches and equipment must be properly cleaned and disinfected or sterilized before reuse.

The HIV-1 and HIV-2 are RNA retroviruses spread by means of blood, semen, vaginal secretions, cerebrospinal fluid, serous fluids, and breast milk. The concentration of HIV in clinical specimens is far lower than that of HBV (100 to 102 infectious doses per milliliter) (9). The virus is slowly inactivated by drying (36). Any disinfectant that is capable of inactivating HBV will be effective against HIV because the latter agent is far more susceptible to such treatment.

Other agents are transmissible in blood and body fluids. The viremia associated with hepatitis A is brief and precedes symptoms by several days. There is no chronic carrier state associated with hepatitis A. The hepatitis C virus may be acquired from contact with blood or body fluids in the laboratory. Delta agent or HBV is an incomplete virus found only in the blood of chronic carriers of HBV. Epidemiological studies suggest that other uncharacterized hepatitis viruses transmissible by blood exist. Other blood-borne illnesses that could be transmitted through exposure to clinical blood specimens that contain viable agents include malaria, syphilis, babesiosis, brucellosis, leptospirosis, arboviral infections, borreliosis, Creutzfeldt-Jakob disease, human T-lymphotropic virus type I, viral hemorrhagic fevers, and cytomegalovirus infection.

The concepts of universal precautions apply to all human blood and tissues. Serous fluids such as pleural, peritoneal, pericardial, amniotic, cerebrospinal, and synovial fluids are included. Semen and vaginal secretions are equally hazardous. All other clinical specimens (such as sputum, stool, sweat, urine, tears, gastric contents, and saliva) are of less concern; universal precautions need apply only if such substances contain visible amounts of blood (9).

Personnel must use proper barrier precautions when handling clinical specimens. Latex or vinyl gloves should be worn and changed regularly. Water-resistant gowns, aprons or smocks, and face protection should be added when there is the possibility of splashing or splatters. Additional, watertight, occlusive dressings should be worn over any area of nonintact skin of the hands or forearms. Frequent handwashing throughout the day (especially whenever gloves are changed) is fundamental. All specimens of blood and body fluids should be collected and transported in containers that prevent leakage; any such container that has the potential for exterior contamination must be transported in a leak-resistant secondary container such as a plastic bag. Workers should be careful not to contaminate work surfaces, containers, requisitions, and reports. Visibly blood-tinged stains and spills must be decontaminated as soon as they occur or are discovered. All work surfaces must be decontaminated at the end of each work shift as well. Techniques that avoid the formation of spatters and droplets should be a regular part of new-employee training and continuing education programs. All pipetting tasks must be performed with the use of mechanical devices. Biohazard warning labels should be applied to all containers that contain contaminated items. IW must be packaged and properly destroyed. All sharps must be handled with respect and discarded in puncture-resistant containers. Selection of technologies that employ less hazardous alternatives (such as avoiding the use of sharp implements or substitution of automated for manual methods) is also part of universal precautions.

Human tissues require special handling. All surgical specimens

should be sealed in leak-proof containers within the operating room. Containers with contaminated exterior surfaces must be placed in plastic bags before transport to the laboratory. Frozen sectioning of unfixed tissue requires special care. Most infectious agents are not inactivated by freezing. Freezing of tissue must be performed cautiously; the spraying of tissue with freezing gases under pressure should be prohibited because it may spatter infectious materials. The interior of the tissue cryostat must be decontaminated regularly.

The entire necropsy suite should be considered a biohazardous area. Only those individuals essential to the performance of the examination should be allowed in the area. The individuals at the prosection table should wear full-barrier PPE including face and eye protection, masks, water-repellent clothing, aprons, shoe covers, and two sets of rubber gloves. Steel-mesh gloves are highly desirable. An assistant whose clothing remains uncontaminated should be assigned to circulate in the clean areas of the morgue. All procedures that may spray or spatter blood or tissue must be performed carefully. Electric bone saws, for example, naturally generate aerosols and particulates; the connection of vacuum exhaust to the saws or the performance of such procedures under plastic shrouds help to contain the hazard. All tissues should be fixed at the necropsy table to limit the number of work surfaces that will inevitably be contaminated with blood.

Special precautions apply to the handling of tissues derived from patients with Creutzfeldt-Jakob disease (37). The agent is resistant to inactivation by formalin and other aldehyde fixatives, alcohols, and heat (38). Strict adherence to universal precautions is required. Aqueous sodium hypochlorite (5.25%) (undiluted household bleach) or 2 N sodium hydroxide is effective for inactivating the agents on surfaces. Contact time should be at least 15 minutes (preferably 60 minutes). Discarded fluids and contaminated instruments may be treated by soaking in 5.25% sodium hypochlorite for 1 hour or by autoclaving at 132°C for 1 hour (39).

Decontamination

Any procedure or technology that reduces the infectivity of a substance or material to a safe (noninfectious) level is a decontaminant. Germicide is a general term for any substance that can kill pathogenic agents. A sterilant completely destroys all infectious agents (including viruses, mycobacteria, and bacterial spores). A disinfectant is effective against selected microorganisms. The spectrum of activity of commercial disinfectants can be found on the product label or in the accompanying literature. Disinfectants may not be effective against bacterial spores or mycobacteria. The effective action of all disinfectants requires adequate contact time; therefore, the manufacturer's instructions should always be followed. Sanitizers reduce microorganisms to levels considered safe by public health codes. They are used primarily in the food industry. An antiseptic is a chemical germicide appropriate for use on skin, tissue, or mucous membranes. Sanitizers and antiseptics are not to be relied on as laboratory disinfectants. Table 3.3 lists some common disinfectants and their spectra of activity.

Spilled blood or tissue samples must be cleaned up and the area decontaminated. Cleaning personnel must wear and use proper PPE. Forceps or scoops should be employed to remove broken glass without the need for manual contact. The bulk of the spill is absorbed onto paper towels, and the towels are properly discarded as IW. Proteins and lipids in the stain can inactivate chemical disinfectants or form a barrier around infectious agents. Therefore, the residuum is next cleaned with detergent and water. After all visible blood has been removed, an appropriate disinfectant is applied. A freshly diluted, 1:10 solution of household bleach (5.25% aqueous sodium hypochlorite) suffices. Iodophors formulated as hard-surface disinfectants may also be used. Aldehydes (as aqueous glutaraldehyde or formaldehyde) and phenols are also effective but are toxic; such substances should only be used in areas of adequate ventilation or with the aid of a chemical fume mask.

The decontamination of laboratory instruments should be a regular event. The frequency of application depends on the intensity of use. Maintenance personnel should wear gloves during all such activities. Additional equipment should be donned for special tasks. Steel-mesh gloves are needed for cleaning microtome knives, and facial protection is important when working with pressurized devices. Accidental spills within instruments must be cleaned and disinfected promptly. Exposure potential can be minimized by sensible technique. Breakage of sample tubes within a centrifuge is an example. The O ring on the lid and the buckets should seal tightly. The rotor must be allowed to come to a full stop and 30 minutes or more should elapse before opening the lid; this allows airborne droplets to settle. Glass shards are removed with forceps. The interior may then be wiped down with detergent and disinfectant. The preferable routine is to use centrifuge cups that are equipped with tightly fitted lids; broken glass and blood will be confined to a single cup (40).

Accidental Exposures and Immunizations

All susceptible clinical laboratory workers who frequently are exposed to blood or tissue should be vaccinated against HBV (41). HBV vaccine is synthesized from recombinant DNA. Three intradeltoid injections are required, the second and third after 1 month and 6 months, respectively. The vaccine is not required for individuals who are known to have adequate titers of antibodies to the hepatitis B surface antigen.

Vaccines for immunoprophylaxis are available for workers who are regularly exposed to specific agents of virulence. Inactivated or live, attenuated vaccines may be considered for veterinary laboratory workers (e.g., anthrax, tularemia, Q fever) or for workers in specialized public health laboratories (e.g., polio, yellow fever, equine encephalitis). All susceptible women of childbearing age who work with rubella should be vaccinated against this virus.

Plans for medical treatment after accidental exposures should be written in the laboratory's policy manual. Immediate treatment can prevent or mitigate an infection. The infectivity of the source and the susceptibility of the injured worker determine the type of therapy and the dosage schedule. For example, an unvaccinated individual exposed to HBV would be treated immediately with hepatitis B immune globulin and then begun on the HBV vaccine series.

There is no commercially available vaccine for HIV, but immediate

prophylaxis with zidovudine and lamivudine may be effective (42). (For this and all prophylactic and therapeutic decisions, the current recommendations of the CDC should be consulted. These are available through hospital departments of infection control and public health clinics.) After acute exposure, the infectivity for HIV should be determined by testing the source material for HIV antibodies (or HIV antigen or HIV DNA). Testing the exposed worker for antibodies to HIV at the time of exposure will allow detection of postexposure seroconversion. Seronegative workers should be followed for seroconversion for at least 6 months (e.g., at 6 and 12 weeks and 6 months). The results of all such studies must be followed by knowledgeable counseling.

RADIATION HAZARDS

Part of "3 - Laboratory Safety"

Ionizing radiation can be one of the most toxic sources of exposure to personnel in the laboratory. Acute exposures can lead to lymphocytolysis, bone marrow suppression, skin damage, and damage to enteric epithelia. Low-level exposure can produce genetic damage to germinal and somatic tissues leading to birth defects and neoplasms. Sources of radiation hazards in the laboratory include both medical devices and reagent materials.

Medical devices that generate ionizing radiation must always be used according to the manufacturers' instructions. Policies should be established to restrict the access of untrained individuals to the areas of radiation hazard. If the source is potent, continuous monitoring of the room to ensure that the shielding is intact is advisable. Moving or repairing the device must be done under the guidance of expert technicians who are skilled in radiation containment.

The radionuclides found in laboratories are generally limited to tracer chemicals that are of insignificant hazard unless ingested or absorbed. The same good laboratory techniques that effectively contain infectious hazards will limit personal exposure to soluble radionuclides. Gloves should be worn when handling radioactive materials, and handwashing after glove removal is to be emphasized. Any manipulation that might lead to hand-to-mouth or hand-to-mucous membrane transfer must be prohibited. Workbenches should always be covered with an absorbent paper toweling, to be changed at the end of each shift (or after overt contamination) and properly discarded.

Reagent nuclides include beta emitters (e.g., ^{3}H, ^{14}C, ^{32}P, and ^{35}S) and gamma emitters (e.g., ^{51}Cr, ^{125}I). Alpha particles are emitted by heavy elements and are unlikely to be found in clinical laboratories. Low-energy beta emitters such as ^{3}H or ^{14}C will produce adverse health effects only if ingested or absorbed; higher energy beta emitters such as ^{32}P can produce significant skin exposure. Gamma and radiographic emissions are penetrating and can cause biological damage without contact. The amount of radiation to which an employee may be exposed is specified by law (43), but the general principle is that exposures be maintained “as low as reasonably achievable” (ALARA) (44). The U.S. Nuclear Regulatory Commission expects employers to have an operational ALARA program.

Warning signs that indicate the presence of radioactivity should be posted in all areas where radionuclides are stored, used, or discarded.

Regulations of the U.S. Nuclear Regulatory Commission and

TABLE 3.3. DECONTAMINANTS AND THEIR SPECTRA OF ACTIVITY

Decontaminant	Concentration of Active Ingredient	Temperature (°C)	Contact Time (min)	Vegetative Bacteria	Lipo Viruses	Tubercle Bacilli	Hydrophilic Viruses	Bacterial Spores
Autoclave (saturated steam)	(15 lb/in^2)	121	50–90	+	+	+	+	+
Autoclave (saturated steam)	(27 lb/in^2)	132	10–20	+	+	+	+	
Dry heat oven	—	160–180	180–240	+	+	+	+	+
UV irradiation (253.7 nm)	40 $\mu W/cm^2$	—	10–30	+	—	+	±	—
Ethylene oxide	400–800 mg/L	35–60	105–240	+	+	+	+	+
Paraformaldehyde (gas)	0.3 g/ft^3	> 23	60–180	+	+	+	+	+
Quaternary ammonium compounds	0.1–2%	—	10–30	+	+	—	—	—
Phenolic compounds	0.2–3%	—	10–30	+	+	+	±	—
Chlorine compounds	0.01–5%	—	10–30	+	+	+	+	±
Iodophor compounds	0.47%	—	10–30	+	+	+	±	—
Alcohol (ethyl or isopropyl)	75–85%	—	10–30	+	+	+	±	—
Formaldehyde (liquid)	4–8%	—	10–30	+	+	+	+	±
Glutaraldehyde	2%	—	10–600	+	+	+	+	+
Hydrogen peroxide (liquid)	6%	—	10–600	+	+	+	+	+

+, very positive response; ±, less positive response; —, negative response or not applicable.
Adapted from Vesley D, Lauer JL. Decontamination, sterilization, disinfection, and antisepsis. In: Fleming DO, Richardson JH, Tulis JJ, et al., eds. *Laboratory safety: principles and practice*, 2nd ed. Washington, DC: American Society for Microbiology, 1995, with permission.

the Canadian Atomic Energy Commission specify procedures that must be followed when shipping and receiving radioactive materials. The individuals responsible for radiation safety in the clinical laboratory must be aware of the applicable regulations and incorporate them into the policies and procedures. Recommendations of this chapter for this reason may be superseded by governmental requirements. All shipments received should have the date and time received, the method of shipment, the condition of the package upon receipt, and the amount and type of radionuclide contained within recorded. Damaged packaging should be checked for contaminating radiation. A leaking package should be sealed in a plastic bag and secured. All areas that may have been contaminated should be surveyed and decontaminated as necessary.

Thermoluminescent dosimetry badges or film-badge dosimeters are suggested for all laboratory workers who are exposed to gamma irradiation, although their use with the usual low-level amounts of tracers found in clinical laboratories is optional. The type of badge device should be chosen by a knowledgeable radiation physicist or safety officer and should be based on the specific type of exposure anticipated.

All workbenches, refrigerators, and other storage areas should be checked regularly for signs of contamination. The most useful surveys are performed immediately after completion of a task; this allows estimation of maximal exposure. Surface activity should be below 200 disintegrations per minute (dpm) per 100 cm^2 (45). Work areas should be surveyed for contamination regularly. This may be performed with a typical survey rate meter. Acceptable levels are usually two to three times background. An alternative is to wipe the work areas with sponges soaked in alcohol or aqueous detergent. The sponges are wiped vigorously over a 100-cm^2 area of the contaminated surface, placed in a counting vial, and read on the usual analytic instrument.

Spills and Cleanup

Personnel engaged in decontamination must wear gowns or coats, aprons, gloves, and eye protection. Disposable items are strongly preferred over clothing that requires laundering. The contaminated area should be scrubbed with water and detergent (preferably using a cleaning compound designed for radiation decontamination). After the area is cleansed, it should be checked for residual radioactivity. The process must be repeated if the radioactivity is greater than three times the background or more than 200 dpm. All disposable materials involved in the cleanup must be discarded as radioactive waste unless survey measurements have indicated the amount of radiation remaining in the materials to be less than twice background. Each spill incident and decontamination episode should be carefully documented. Personnel involved should be advised to seek medical evaluation if the exposure was significant.

FIRE PREVENTION AND CONTROL

Part of "3 - Laboratory Safety"

There are multiple sources of combustible materials in the clinical laboratory. These include ignitable liquids and flammable gases as well as paper and wood construction materials. The laboratory must be designed to minimize these hazards. It is prudent to train personnel to intercede as soon as possible in a fire emergency because fires are most easily extinguished when still small. Emergency evacuation may be necessary to prevent injury to personnel; therefore, plans to clear each location within the laboratory should be made beforehand and explained to the staff. Laboratories located within hospitals are potential sources of ignition that may place bedridden patients and visitors at risk.

The working supplies of all ignitable materials should be minimal. Only modest amounts of reserve supplies should be kept, and these should be stored in proper flammable liquid storage cabinets or vaults. Safety cans should be used for all bulk storage of 5 gal or less in which the steel or polyethylene container will not compromise the purity of the solvent.

All potential ignition sources should be identified and controlled. Obvious sources include open flames and heating elements, but the most common sources are sparking within an electrical switch or motor (46). Static discharges are preventable sources of ignition; grounding of metal containers during transfer of combustible liquids will prevent sparking. Exothermic chemical reactions may also generate sufficient heat to ignite a fuel source. Clear separation of flammable materials from such ignition points is indicated.

The laboratory should be designed for fire protection. Walls and doors should be constructed to contain a conflagration until fire fighting equipment and personnel can be mustered. (In general, this specifies that walls be rated at 1 hour and door assemblies at 45 minutes.) Fire exits must be provided for emergency evacuation of personnel. Rooms that have an area larger than 1,000 square feet and all rooms in which flammability concerns are present should have at least two doors that open to exit corridors or to the outdoors.

A fire alarm system should be located in or near the laboratory. This may be a bell or public address system, but it must be audible in all areas of the facility, including storage areas, darkrooms, and lavatories.

Portable fire extinguishers should be located wherever significant fire hazards exist. Class A devices are appropriate for wood and paper fires, class B devices for ignitable solvents, and class C devices for electrical fires (33,47). Three types are commonly found in clinical laboratories: organic halogen gas, carbon dioxide, and dry chemical. Each is effective against solvent and electrical fires; the label on the device will specify the rating of the unit. Extinguishers using organic halogens are rapidly effective. They are most commonly found in electrical applications such as computer rooms because their vapor will not damage delicate equipment. Carbon dioxide acts by smothering and cooling a fire. Devices of this type are equally rapid but may cause thermal damage to instruments in the vicinity of use. Both types should be used only in well-ventilated areas or by a person wearing a SCBA. Dry chemical types are the least expensive. They are particularly effective against oil and grease fires, but cleanup is laborious and equipment damage can be expected. Water-based extinguishers are effective only against class A fires and should not be used in clinical laboratories. The size of the device should be large enough to extinguish a small fire without being too large for any of the personnel in the work area to manipulate.

Laboratory personnel should practice their responses to fire situations. Training sessions should emphasize sounding the alarm as the first action. Attempts to extinguish the fire should follow, but only if the proper devices are immediately available and the fire appears small enough to manage. Evacuation of the area may be necessary. Employees should be trained to close doors and windows as they leave by the designated exits.

All personnel working in hazardous areas should be trained to use the types of portable fire extinguishers chosen. There is no substitute for hands-on experience. Fire exit drills should be held regularly. Each employee should know the preferred and the alternate exit route for his or her workstation well enough to be able to find them in a darkened, smoke-filled atmosphere. Exit drills should be held at least once a year for the benefit of all employees.

ELECTRICAL SAFETY

Part of "3 - Laboratory Safety"

Electrical hazards are among the most difficult to detect in clinical laboratories. Shock hazards can develop in any electrical instrument or appliance. The laboratory should be surveyed at least once a year as part of an electrical safety inspection. All new instruments should be thoroughly inspected before being placed into service.

The preventive maintenance program should include regular safety inspections of all portable electrical devices by a qualified individual. The power supply cable must be checked for frayed or cracked insulation and for broken wires. Electrical equipment must be properly grounded, i.e., the power cable must be of three-wire construction and the neutral (green) wire must be in electrical continuity from the power supply chassis to the receptacle. (An exception may be made for a device that is doubly insulated or totally encased in plastic.) Attempting to ground electrical appliances to cold-water plumbing may be ineffective.

A qualified electrician should check instruments for current leakage with the ground open. The manufacturer's specifications should be used to determine the maximum allowable current for this test. When such specifications are not provided, a 100-mA current may be used to test most devices.

Electrical receptacles should be checked for ganged plugs, loose fittings, and broken faceplates (48). Electrical tests for polarity and ground continuity should be performed annually. Ground fault circuit interrupters must be installed in “wet” locations. Faulty and potentially unsafe equipment must be removed from service or locked and tagged to prohibit use until repairs have been completed.

Electrical equipment used in areas wherein ignitable vapors might accumulate should be designed and approved for such uses. Examples include flammable liquid storage vaults and applications that require heating of flammable and combustible solvents.

The safety-training program should anticipate electrical accidents. All personnel should know that the first response is to disconnect power to the arcing or burning equipment if it can be done without personal risk. This is especially important for individuals working in the vicinity of high-voltage devices, such as those used for electrophoresis.

SPECIALIZED EQUIPMENT

Part of "3 - Laboratory Safety"

Hazards may accompany any new item of laboratory equipment. Items that have moving parts, thermal elements, pressurized components, or radiant energy may be of particular concern. The manufacturer's instructions that accompany such devices must describe the particular risk and make recommendations for worker protection.

Centrifuges that have large rotors are mechanical hazards. Hair or clothing can become entangled in the mechanism if the centrifuge is allowed to operate with its cover open. Unbalanced cups and broken glass can be especially dangerous. All centrifuges should be fitted with a lid that is connected to a power-interruption switch to prevent operation with the rotor exposed. Spinning rotors must never be slowed or stopped by manual means.

Autoclaves must be operated by trained and knowledgeable individuals. Steam under pressure and heated to temperatures above the normal boiling point of water presents risks of scalding and explosion. The installation and fittings should be examined periodically by a stationary engineer. The pumps and valves should be inoperable when the door is unsecured. The door latch should be automatically disabled after operation until the interior temperature and pressure conditions permit safe access. The personnel who unload the chamber should wear heat-resistant gloves, aprons, and face protection.

A cryogenic liquid is any liquid with a boiling point of less than -130°F (-90°C). The extreme cold of such liquids can freeze tissues—including the fingers of careless workers—instantly. All give off large volumes of gas as they vaporize. For example, liquid nitrogen expands 700-fold as it boils at standard temperature and pressure. Inserting tissues or instruments into a cryogenic liquid will cause it to boil; careful technique and good ventilation are essential. All exposed skin must be covered when working with cryogenic liquids. Insulated gloves should be loose fitting so that they may be removed quickly if liquid is spilled within them (49).

Cell-sorting devices used for unfixed cells can expose the operator to infectious droplets and aerosols. Specific recommendations for minimizing this hazard are available (50). Laser devices may require the operators to wear special eye protection. Electrothermal heating devices may introduce radio frequency wave hazards to pacemaker wearers. It is not possible to list all of the possible hazards within this chapter. No piece of equipment should be placed into service until the special hazards have been defined, appropriate engineering controls have been installed, the requisite personal protective equipment provided, and personnel have been appropriately trained.

HAZARDOUS WASTE DISPOSAL

Part of "3 - Laboratory Safety"

Nearly all laboratory processes result in discarded solutions or materials. Much of it is hazardous (approximately 80 lb/1,000 ft^2 of chemical wastes and approximately 1 ton/1,000 ft^2 of biohazardous waste per year) (51), but all of it must be disposed of safely. Proper disposal implies protection of the workers who will handle the waste as well as consideration of the environment (52).

Waste may be discharged into water, buried, burned (or otherwise destroyed), or transformed for reuse. By U.S. law, the containment or destruction of a hazardous solid waste forever remains the responsibility of the generator regardless of any assistance in handling, transporting, treatment, or disposal of the waste.

Chemical Wastes

According to EPA regulations, chemical wastes may be hazardous because of the properties of ignitability, corrosivity, reactivity, or toxicity. Ignitable wastes include liquids (e.g., acetone, alcohols, toluene, xylene), gases (e.g., hydrogen, propane, butane), or solids (e.g., nitrate salts, peroxides). Corrosives include strong acids and alkalis. Examples of reactive wastes are dry picric acid and aged diethyl ether that contains peroxides. Toxicity is currently defined by the ability to generate dangerous leachates that can contaminate ground or surface water and by lists of specific chemicals.

Hazardous chemical waste (HCW) should not be mixed with each other unless one knows the compounds to be compatible. Likewise, they should not be mixed with nonhazardous waste, radioactive waste, or IW. Each type of HCW should remain segregated throughout transport, storage, and disposal. All workers must use appropriate PPE when handling HCWs.

A safe area for accumulation of HCW materials before transport will be well ventilated, secure from unauthorized access, surrounded by a berm or catchment basin, and provided with proper fire protection. HCWs should not be allowed to accumulate onsite for longer periods than necessary or in excessive amounts. The same principles of safe storage that are appropriate for the parent reagents apply to discarded chemicals: incompatible chemicals must be separated, and all waste containers must be properly labeled as to content and initial date of accumulation. Unlabeled containers of hazardous waste must be treated as hazardous.

Some HCWs may be treatable onsite. Examples include neutralization of acids and bases or precipitation of inorganic metal cations as sulfides. Such activities should only be performed under the direction of a knowledgeable chemist. Other HCWs may be reclaimed for reuse: organic solvents may be distilled; silver may be extracted from photographic darkroom wastes, for example. HCWs should never be discarded into ground water or in natural waterways. All chemicals introduced into the sanitary sewage system must be compatible with the treatment processes of the local publicly owned treatment works.

Infectious Wastes

Wastes capable of transmitting an infectious disease may come from (a) cultures and other discards from microbiological laboratories; (b) human blood, body fluids and tissues, and items contaminated with blood; and (c) contaminated animal carcasses and bedding. Control of the hazards of IW is an occupational health issue (52).

In microbiological laboratories, IW should be handled according to the BSL appropriate for the etiologic agents under study. In all clinical laboratories handling blood and body substances, universal precautions apply.

Packaging materials must be of sufficient strength and integrity to protect all who must handle the containers. Bags should be red or orange or display the universal biohazard symbol (Fig. 3.3). Contaminated sharps must be in puncture-resistant containers; liquid wastes must be in leak-resistant packaging.

Most IWs will be accumulated before disposal. The storage site should have limited access and be protected from rodents or other vermin that could become disease vectors. Steps should be taken to prevent putrefaction, including temperature control and short storage duration. Handling and transport methods must maintain the integrity of the packaging.

Some IWs, such as microbiological cultures, must be treated before disposal. Steam sterilization is preferred. The adequacy of decontamination depends on the exposure time; the temperature; the size, density, and water content of the load; and the penetrability of the containers. Autoclaving operations must be monitored periodically with a biological indicator (or chemical equivalent) that is placed inside a sealed package near the center of the load. Other treatment methods include dry heat sterilization, gas vapor sterilization, and chemical disinfection. The method chosen must be one that is effective against the agents known to be included in the IW.

Radioactive Wastes

All radioactive waste handling must be consistent with the regulations and license under which the laboratory operates. General principles include segregation of such wastes from the normal trash in a properly labeled, designated area. Radionuclides with a short half-life, such as ^{32}P and ^{125}I, may be allowed to decay in a protected area onsite and be discarded as nonhazardous refuse. Materials that cannot be allowed to decay in storage, such as ^{3}H, ^{14}C, and ^{35}S, may be released into the sanitary sewer system or into the atmosphere (as a gas or via incineration) in modest quantities as limited by law. Transfer of radioactive materials to offsite locations for burial or long-term storage is discouraged.

Sharps

Discarded needles, lancets, scalpel blades, Pasteur pipettes, capillary tubes, and shards of glass and plastic can cause penetrating injury (53). Infectious agents contaminating waste sharps can gain a portal of entry into the body through a puncture wound.

All sharps should be discarded immediately into a rigid vessel that resists both leakage and puncturing. Such containers should be located in every work area in which waste sharps may be generated. (This encourages proper disposal and precludes the need to carry pointed items any distance.) The containers should be clearly labeled to warn handlers of the potential for injury. Containers must not be allowed to overfill; injury can result from forceful insertion. Infectious sharps should be treated before disposal. All sharps should be destroyed to prevent injury to waste handlers. Destruction methods include incineration, grinding, shredding, crushing, and melting.

Waste Disposal

The laboratory director must be aware of the disposal processes that are used for all wastes generated by the facility. The methods must comply with all federal, state, and local regulations. With regards to HCWs, federal law identifies the generator as responsible for all personal and environmental injury that may ultimately develop, i.e., transfer of the waste does not transfer the generator's liability.

Incineration is the best overall disposal method, but the incinerator must be able to accommodate the materials in the waste mixture. HCWs should be burned only by appropriately licensed facilities. Incineration is both a treatment and a disposal method for IWs and sharps, but the large proportion of chlorine-containing plastic in the typical IW mix may require special incinerator design. Burial in a landfill should be considered only for nonhazardous wastes. Sewer disposal is appropriate for most aqueous liquid wastes but not for any materials that might interfere with processes in the water treatment facility.

Waste Minimization

All laboratories should develop a program for reducing the generation of hazardous waste. Waste minimization can be achieved either by reducing the volume of hazardous wastes generated or by reducing the degree of hazard of the waste generated. The volume of HCWs can be reduced by (a) substituting less hazardous chemicals in laboratory operations, (b) decreasing the amounts of reagents required by adopting micromethods, (c) recovering spent materials for recycling, and (d) redistributing surplus or unwanted chemicals to laboratories that can use them.

Reduction in the amount of IW generated is best achieved through proper segregation of such waste. The mixture of a small quantity of an IW with a large quantity of non-IW generates a large quantity of IW. Segregation must be performed at the point of generation. All IW must be inserted into the proper packaging, and noninfectious refuse should be kept physically separated to focus the waste handlers' attention on the materials of greatest concern.

Contingency Planning

Spill containment and cleanup procedures must be defined and specified for each type of hazardous waste generated by the laboratory. Fires, explosions, or other unexpected releases that may involve the health of individuals outside the facility or that may contaminate surface water in the vicinity must be included in the plan.

Record Keeping

Hazardous waste manifests must be completed and maintained as required by law. Injuries involving wastes must be reported as required to governmental agencies and to the facility's safety program for quality assessment and improvement.

EMERGENCY PLANNING

Part of "3 - Laboratory Safety"

The laboratory's safety policies manual should define the appropriate actions and responses in case of unforeseen emergency. The types of situations that must be included in the planning will vary with the nature of the laboratory, its equipment, and the materials that it stores. Emergency planning should take into consideration employees, patients, visitors, and the special needs of persons with disabilities.

Power Failure

Loss of electrical power incapacitates a modern clinical laboratory. Emergency lighting must be provided to enable workers to move about the premises safely and to facilitate egress if unsafe conditions prevail. Air-handling devices require electrical power. The exhaust fans in chemical fume hoods and BSCs may shut down in a power failure; all processes that must be performed in such cabinets must cease until power is restored. Laboratory policy should prohibit the use of any equipment that cannot be manipulated safely under conditions of emergency power.

Fires

Fire emergency preparedness is discussed in the section on Fire Prevention and Control.

Natural Disasters

Tornadoes, floods, and similar disasters may devastate the facility. Clinical operations must not be restarted until qualified inspectors have ensured the safety of the workers. Unauthorized individuals must not be allowed premature access to the premises. The building must be structurally sound. Electrical equipment that is not designed for wet locations must be disconnected until all water has been removed. Damaged containers of chemicals must be properly gathered and removed. Potentially infectious materials must be bagged, decontaminated, and safely discarded. In locations where earthquakes are known, movable equipment should be secured to walls or countertops.

Spill Cleanup

All hazardous materials can be spilled, and appropriate provisions must be made for each. All individuals who work in an area should have sufficient understanding of the basic medical, chemical, radiation-related, or infectious problems that can occur there to be able to recognize an emergency and to call for help. Small spills can often be safely cleaned by the spiller following written instructions. Larger spills require the intervention of specially trained personnel.

Medical Emergencies

Laboratories that deal directly with patients must be prepared for clinical crises. Individuals trained in basic cardiopulmonary resuscitation should be immediately available. If emergency supplies

of drugs and equipment are maintained onsite, a program to ensure the freshness of the medications and the proper functioning of the gear must be provided. If all emergency medical care is provided at or by remote facilities, the response time should be brief enough to minimize clinical complications.

Electrical shock injury may result in loss of consciousness. First responders should be trained to approach the victim with caution. The source of the electrical discharge must be inactivated if it can be identified. Power to the defective appliance should be cut, preferably at the junction box, or the device can be pushed away with a wooden broom or another nonconductive implement. Cardiopulmonary resuscitation should be initiated if the individual is not breathing, and emergency response personnel should be called.

The laboratory's CHP includes the assessment of the health effects of all hazardous chemicals. Spills or other acute releases may expose workers to toxic vapors. The CHP should specify the criteria for medical intervention. Such criteria may be based on signs or symptoms of overexposure or on environmental measurements. Splashes or sprays of toxic chemicals may result in exposure by absorption through the skin or mucous membranes. All such incidents should be documented, and medical treatment should be provided according to the CHP.

Similarly, the laboratory's infection control plan should specify the criteria for referral of employees who have been accidentally exposed to infectious agents.

Bomb Threat

Telephoned threats of bombing or other malicious acts should be included among the facility's emergency planning. All such calls should be taken seriously. The police or security services should be notified immediately. The caller should be kept on the line, and all details of the call recorded. The police may direct evacuation of the facility, and if so, the exit should proceed in an orderly manner.

EMPLOYEE TRAINING

Part of "3 - Laboratory Safety"

All personnel must have the information and skills to control the hazards in their environment. Training programs are required under many governmental regulations and for accreditation by various voluntary agencies. The basic elements of a safety-training program are (a) content, (b) target audience, (c) educational methodology, (d) documentation, and (e) monitoring for effectiveness.

Content

All safety policies should be included in the training program. The program for each employee should include the particular hazards applicable to his or her job. The program must give the employee sufficient information to understand the hazards. The controls provided by the laboratory director to contain those hazards must be included, and the requisite work practices must be emphasized. Employees must know how and when to use personal protective devices. Chemical handling, infectious exposures, waste disposal, and emergency preparedness are requisite subjects. The material should be presented in a manner consistent with the educational level, literacy, and language background of the people to whom it is presented.

Employee right-to-know regulations are exemplified at the federal level in OSHA's Hazard Communication Standard (54). Such regulations typically require the employer to provide the employees access to copies of the regulations and to the facility's safety plans. Chemical hazards may be summarized in the CHP, referencing MSDSs and other suitable sources. Training that accompanies a comprehensive ECP will provide employees with information on infectious hazards, methods to contain those hazards, and emergency medical intervention to be initiated in case of an exposure.

Target Audience

All employees who handle or are exposed to hazardous materials must be included. Training is necessary for newly hired employees, for individuals assigned to new work areas, and periodically to familiarize all employees with technological and policy developments. Training must be implemented when new hazards are introduced into the workplace and when policies or programs are changed. Periodic refresher briefings of existing policies for all employees are strongly recommended.

Teaching Methods

Most safety training will be in lecture format, but consideration of hands-on activities is suggested. The operation of portable fire extinguishers, tight-fitting respirators, and spill containment is best presented in safety workshops. Some of the most effective programs are those that encourage workers to suggest safer alternatives to current practice.

Documentation

The date, title, and a summary of each safety presentation should be recorded. Attendance should be documented in each individual's employment records.

Monitoring and Evaluation

The educational plan should provide assurance that all-important aspects of the safety program are included. Attendance by the individuals in the targeted audience must be recorded to ensure that no one is inadvertently excluded. The effectiveness of the program should be tested in an objective manner. Although this may be provided by a multiple-choice test or other academic tool, the best feedback is provided by assessment of resultant on-the-job behavior.

IMPROVING THE SAFETY PROGRAM

Part of "3 - Laboratory Safety"

There is no clinical laboratory that is so safe that it cannot be made safer. A periodic safety audit should be undertaken by the safety committee to assess the needs of the personnel. A checklist

should be constructed from the facility's safety manual to provide a useful tool for this audit. (For an example, see reference 2.) The audit should provide a description of the physical facility, including housekeeping, warning signs, fire prevention and preparedness, electrical safety, storage of compressed gases and chemicals, and containment of infectious and radioactive hazards. The condition of the equipment and its maintenance should be assessed. It is equally important, although perhaps more difficult, to determine the prevailing attitudes of the employees toward safe work practices and their compliance with existing policies. Required documents of the program should be evaluated for completeness.

The role and effectiveness of the safety committee should be reviewed annually. Minutes of the meetings should convey thoroughness and attention to detail. The review of incident and accident reports, whether required by OSHA or not, should be documented. Formal inquiry and follow-up of serious accidents or repetitive problems are essential. Reports prepared for governmental agencies, such as hazardous waste manifests, should be monitored by the safety committee or the laboratory's administration for adequacy and timeliness.

Reports of external reviewers can be especially helpful in assessing the safety program. Safety audits may be provided by governmental inspectors (e.g., OSHA, Nuclear Regulatory Commission, licensing agencies, fire departments) and by voluntary programs (e.g., College of American Pathologists, Joint Commission on Accreditation of Healthcare Organizations, industrial risk insurers). Each itemized area of concern or recommendation is a potential opportunity to improve the safety program.

The program's action plan should prioritize the identified issues of concern according to the types of problems uncovered, their frequency and severity, and the probability of their recurrence. The effectiveness of all measures taken as a result should be assessed after a suitable period of follow-up.

A successful safety program is the result of the combined efforts of everyone in the laboratory. Each individual in the laboratory should contribute to the total effort. The administration of the facility must make a statement of commitment to the goal of a safe workplace. The authority for implementing the program must be vested in the responsible supervisory personnel. Job descriptions should be concrete with regards to employees' responsibilities to comply with safety policies.

The administration must provide a proper training program for all its employees, both those new to the organization and those whose formal training preceded the introduction of current technology. Employees must bring to the attention of the management any identified but unsafe working condition as well as perceived opportunities to improve general safety conditions. Senior managers should be responsive to suggestions for improving the safety program. The attitudes of laboratory workers toward safety practices should be positive.

Work groups inevitably set their own safety standards; management cannot force employees to work safely (55). An ongoing program of audit for compliance with defined policy is important. This is a function that can (and should) be integrated with other quality assessment and improvement activities within the organization.

REFERENCES

1. Montgomery L. *Health and safety guidelines for the laboratory*. Chicago: ASCP Press, 1995.

2. NCCLS. *Clinical laboratory safety; approved guideline*, NCCLS document GP17-A. Wayne, PA: NCCLS, 1996.

3. Title 21, Code of Federal Regulations, Part 1910, 1910.1450 [abbreviated in subsequent references as 29 CFR 1910.1450].

4. NCCLS. *Procedures and devices for the collection of diagnostic blood specimens by skin puncture; approved standard*, 4th ed., NCCLS document H4-A4. Wayne, PA: NCCLS, 1999.

5. NCCLS. *Procedures for the handling and transport of diagnostic specimens and etiologic agents, approved standard*, 3rd ed., NCCLS document H5-A3. Wayne, PA: NCCLS, 1994.

6. Lymer UB, Schutz AA, Isaksson B. A descriptive study of blood exposure incidents among healthcare workers in a university hospital in Sweden. *J Hosp Infect* 1997;35:223–235.

7. Burnett L, Chesher D. Application of CQI tools to the reduction in risk of needlestick injury. *Infect Control Hosp Epidemiol* 1995;16:503–505.

8. Goldwater PN, Law R, Nixon AD, et al. Impact of a recapping device on venipuncture-related needlestick injury. *Infect Control Hosp Epidemiol* 1989;10:21–25.

9. NCCLS. *Protection of laboratory workers from instrument biohazards and infectious disease transmitted by blood, body fluids, and tissue; approved guideline*, NCCLS document M29-A. Wayne, PA: NCCLS, 1997.

10. Geller SA, Gerger MA. Guidelines for high risk autopsy cases. In: Hutchius G, ed. *Autopsy performance and reporting*. Northfield, IL: College of American Pathologists, 1990:60–66.

11. Koenig AS. *Medical laboratory planning and design*. Northfield, IL: College of American Pathologists, 1992.

12. NCCLS. *Laboratory design; approved guideline GP18-A*. Wayne, PA: NCCLS, 1998.

13. Korniewicz DM, Laughon BE, Butz A, et al. Integrity of vinyl and latex procedure gloves. *Nurs Res* 1989;36:144–146.

14. The National Institute for Occupational Safety And Health. *Alert. Preventing allergic reactions to natural rubber latex in the workplace*, DHHS: National Institute for Occupational Safety and Health, publication 97-135. Washington, DC: 1997.

15. Wieslander G, Norback D, Edling C. Local cold exposure of the hands from cryosectioning work in histopathological and toxicological laboratories: signs and symptoms of peripheral neuropathy and Raynaud's phenomenon. *Occup Environ Med* 1996;53:276–280.

16. 29 CFR 1910.134.

17. NFPA 45. *Standard on fire protection for laboratories using chemicals*. Quincy, MA: National Fire Protection Association, 1996.

18. 20 CFR 1910.145.

19. NFPA 704. *Standard system for the identification of materials for emergency response*. Quincy, MA: National Fire Protection Association, 1996.

20. 49 CFR 172, Subpart F.

21. 40 CFR 261.22.

22. NFPA 30. *Flammable and combustible liquids code*. Quincy, MA: National Fire Protection Association, 1996.

23. 40 CFR 261.21.

24. NFPA 99. Health care facilities. Quincy, MA: National Fire Protection Association, 1999.

25. *TLVs and other occupational exposure values—1997*. Cincinnati: American Conference of Government Industrial Hygienists, 1997.

26. 29 CFR 1910.1048.

27. 29 CFR 1910 Subpart Z.

28. Sweet DV, ed. *Registry of Toxic Effects of Chemical Substances (RTECS)*, DHHS (NIOSH) publication no. 97-119. Cincinnati, OH: USDHHS PHS, CDC, NIOSH, 1997.

29. CGA P-1-1991. *Safe handling of compressed gases in containers*. Arlington, VA: Compressed Gas Association, Inc., 1991.

30. U.S. Centers for Disease Control and Prevention. Guideline for isolation precautions in hospitals, infection control and hospital epidemiology. *MMWR Morb Mortal Wkly Rep* 1987;36(Suppl 2S):2S–18S; 1988;37:377–382, 387–388.

31. 29 CFR 1910.1030.

32. Richmond JY, McKinney RW, eds. *Biosafety in microbiological and biomedical laboratories*, HHS publication no. (CDC) 93-8395. Washington, DC: U.S. Government Printing Office, 1993.

33. Centers for Disease Control. Update: universal precautions for prevention of transmission of human immunodeficiency virus, hepatitis B virus, and other bloodborne pathogens in health-care settings. *MMWR Morb Mortal Wkly Rep* 1988;37:377–82, 387–388.

34. 1988 Agent summary statement for human immunodeficiency virus and report on laboratory-acquired infection with human immunodeficiency virus. *MMWR Morb Mortal Wkly Rep* 1988;37(Suppl S4):1–17.

35. Gordin FM, Gibert C, Hawley HP, et al. Prevalence of human immunodeficiency virus and hepatitis B virus in unselected hospital admissions: implications for mandatory testing and universal precautions. *J Infect Dis* 1990;161:14–17.

36. Resnick L, Veren K, Salahuddin SZ, et al. Stability and inactivation of HTLV-III/LAV under clinical and laboratory environments. *JAMA* 1986;255:1887–1891.

37. Brown P. Guidelines for high risk autopsy cases: special precautions for Creutzfeldt-Jakob disease. In: Hutchins G, ed. *Autopsy performance and reporting*. Northfield, IL: College of American Pathologists, 1990:68–74.

38. Titford M, Bastian FO. Handling Creutzfeldt-Jakob disease tissues in the histology laboratory. *J Histotechnol* 1989;12:214–216.

39. Asher DM, Gibbs CJ, Gajdusek DC. Slow viral infections: safe handling of the agents of subacute spongiform encephalopathies. In: Miller BM, ed. *Laboratory safety: principles and practices*. Washington, DC: American Society for Microbiology, 1986.

40. Hambleton P, Dedonato G. Protecting researchers from instrument biohazards. *Biotechniques* 1992;13:450–453.

41. Centers for Disease Control. Update on adult immunization recommendations by the Immunization Practices Advisory Committee (ACIP). *MMWR Morb Mortal Wkly Rep* 1991;40(RR-12):1-52

42. CDC. Public health service guidelines for the management of health-care worker exposures to HIV and recommendations for postexposure prophylaxis. MMWR 1998;47(RR-7):1–28.

43. 10 CFR 20:101.

44. *ALARA at medical institutions. Regulatory guide 8.18*. Washington, DC: Nuclear Regulatory Commission, 1982.

45. Staiger JW. Techniques for safe handling of radioactive materials. In: Miller BM, ed. *Laboratory safety: principles and practices*. Washington, DC: American Society for Microbiology, 1986.

46. Hoeltge GA, Miller A, Klein BR, et al. Accidental fires in clinical laboratories. *Arch Pathol Lab Med* 1993;117:1200–1204.

47. NFPA 10. *Standards for portable fire extinguishers*. Quincy, MA: National Fire Protection Association, 1998.

48. 29 CFR 1910.334(a).

49. CGA O-12. *Safe handling of cryogenic liquids*. Arlington, VA: Compressed Gas Association, Inc., 1993.

50. Schmid I, Nicholson JK, Giorgi JV, et al. Biosafety guidelines for sorting of unfixed cells. *Cytometry* 1997;28:99–117.

51. Hoeltge GA. Managing hazardous waste in the clinical laboratory. *Clin Lab Med* 1989;9:573–586.

52. NCCLS. Clinical laboratory waste management; approved guideline, NCCLS document GP5-A. Wayne, PA: NCCLS, 1993.

53. Dale JC, Pruett SK, Maker MD. Accidental needlesticks in the phlebotomy service of the department of laboratory medicine and pathology at Mayo Clinic Rochester. *Mayo Clin Proc* 1998;73:611–615.

54. 29 CFR 1910.1200.

55. Songer JR. Laboratory safety program organization. In: Fleming DO, Richardson JH, Tulis JJ, et al., eds. *Laboratory safety: principles and practice*, 2nd ed. Washington, DC: American Society for Microbiology, 1995.

4

The Interpretation of Laboratory Tests

Earle W. Holmes

Laboratory tests and the principles on which they are based are a major focus of the discipline of laboratory medicine. The value of laboratory testing to patient care is evident both by the sheer volume of tests that are ordered and by the high level of confidence that clinicians place on laboratory data. This chapter is intended to address the interpretation of laboratory test results, that is, the act of discerning the significance or meaning of a particular test result in the context of the medical question or hypothesis that the ordering clinician is attempting to address. Interpretation is the penultimate, and arguably the most critical, component of a complex series of events that I refer to as the testing cycle (Fig. 4.1). The testing cycle begins and ends with a patient who presents with a medical problem and expects to receive a diagnosis or a prognosis and some type of intervention or treatment. When laboratory testing is required for the further evaluation of a patient, the clinician expects the testing cycle to generate high-quality laboratory information that will facilitate the process of clinical decision making. Under ideal circumstances, all components of the testing cycle will have been optimized to minimize variability and error, and the interactions between the clinician and the laboratory will be free from any confusion or misunderstanding regarding the medical significance of a particular test result. However, in practice, the testing cycle is influenced by many factors that increase the uncertainty associated with a test result and thereby decrease the accuracy of the information that it provides.

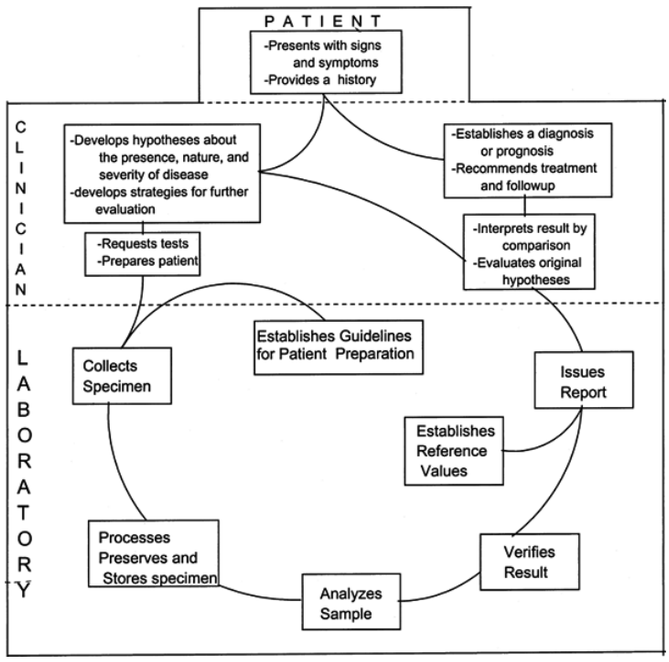


FIGURE 4.1. The testing cycle. The ultimate significance of a laboratory test result depends on a complex series of actions and interactions on the part of the patient, the clinician, and the clinical laboratory. The cycle progresses in a counterclockwise direction.

Factors that have the potential to confound the interpretation of test results occur at all levels of the testing cycle (1). For example, a major source of uncertainty that operates at the patient level is biological variability in the values of laboratory parameters (hereafter referred to as analytes). Variability within the clinical laboratory can enter the cycle during the pre- and postanalytical phases of testing as a result of improper or inconsistent specimen collection or preservation or by the provision of inappropriate or inadequate reference values (RVs) in the laboratory report. Other sources of variability that operate during the pre- and post-analytical phases of the testing cycle and that are largely in the hands of the clinician include the improper or inconsistent preparation of the patient before sample acquisition and the suboptimum selection of laboratory tests relative to the clinical questions that are under evaluation.

Over the past two decades, a great deal of attention has been focused on the analytical variability of laboratory methods. This attention has led to tremendous increases in the quality of test results because of improvements in the precision and accuracy during the analytical phase of the testing cycle. At present, most experts recognize that analytical variation is only a small fraction of the overall variability of most laboratory tests (2). Thus, attention has rightfully shifted to addressing some of the nonanalytical factors that continue to affect the quality and value of laboratory information (3). Specialists in laboratory medicine should become familiar with both the analytical and clinical aspects of laboratory testing so that they are better able to design and implement procedures that reduce variability across the testing cycle and to serve as consultants in the evaluation of unusual or unexpected laboratory results. This chapter deals with three topics of great relevance to the interpretation of laboratory test results: the effects of biological variability on laboratory tests, the establishment and use of RVs, and the evaluation of the diagnostic accuracy of laboratory tests.

BIOLOGICAL VARIATION

Part of "4 - The Interpretation of Laboratory Tests"

Biological variation in laboratory test results is owing to within-subject variation over time and subject-to-subject variation within a population. The principal cause of the former is the existence of endogenous biorhythms for many physiologic parameters, whereas the latter is usually owing to differences among subjects in constitutional factors and/or lifestyle (Table 4.1). Because the principle behind the interpretation of most laboratory tests is to ascribe unusual or extreme test results to the presence of disease, biological variation interferes with interpretation when its magnitude is similar to that observed in diseased subjects. Because small abnormalities in laboratory values are often observed during the early stages of the evolution of disease, biological variation is not only a potential source of false interpretations but can also reduce the diagnostic efficiency of a laboratory test by widening the test's reference interval (RI). Theoretically, the negative effects of biological variation can be minimized through the use of appropriate state-specific RVs; however, the large number of potential sources of biological variation precludes the establishment of RVs that are able to simultaneously account for more than one or two factors. For this reason, it is

crucial that laboratorians and clinicians know how biological variation influences a particular test result and to evaluate the significance of the result in view of this knowledge.

TABLE 4.1. BIOLOGICAL VARIABLES THAT AFFECT LABORATORY TEST RESULTS

Biological rhythms
Circadian
Ultradian
Infradian
Constitutional factors
Gender
Age
Genotype
Extrinsic factors
Posture
Exercise
Diet
Caffeine
Drugs and pharmaceuticals
Oral contraceptives
Alcohol use
Pregnancy
Intercurrent illness

There is a large and complicated literature describing effects of specific biological variables on specific laboratory tests. In many cases, the data for a particular interaction between a variable and a laboratory test are based on the investigation of small samples of subjects. In other cases, conflicting data exist for a particular test–variable interaction. The present review is intended to summarize some of the more important general sources of variability and to focus on those interactions that are well documented. The reader should always be prepared to consult the primary literature and comprehensive reviews on the subject (4,5,6 and 7) when investigating a possible interaction of a specific variable with a specific laboratory test.

Biological Rhythms

Circadian Rhythms

Biological rhythms are frequently observed for many physiologic and laboratory parameters (8,9 and 10). The three types of rhythms that have the most influence on the interpretation of laboratory results are circadian, ultradian, and infradian rhythms. Circadian rhythms (Fig. 4.2a) (11) are responsible for programing the daily sequences of many metabolic and behavioral changes. They have cycle times of approximately 24 hours during which the activity of the rhythmic process (be it laboratory value or physiologic parameter) shows a maximum value (peak or acrophase) and a minimum

value (trough or bathyphase). The amplitude of a circadian rhythm (which by convention is equal to one half of the difference between the peak and trough values) and the clock time at which its peak and trough values occur differs substantially from analyte to analyte (Table 4.2). Circadian rhythms are usually synchronized to the light/dark or sleep/wake cycles and are further modified by social and behavioral cues. A circadian rhythm for a particular variable will generally demonstrate peak and trough values that occur at consistent times in different subjects within the same population. However, a marked subject-specific change in an important entraining cue such as the reversal of the light/dark cycle that occurs in night workers and transcontinental airline travelers can cause a shift in circadian rhythms as they become gradually reentrained to the new cue. Thus, night workers have plasma cortisol rhythms that are reversed relative to their counterparts on the day shift, and an elevated plasma cortisol concentration in a sample drawn at midnight would not be considered to be abnormal. The fact that circadian rhythms are predictable and relatively consistent in a population raises the possibility that rhythm-sensitive RVs can be used to prevent the misinterpretation of test results. Despite this theoretical possibility, rhythm-adjusted RVs are not widely used on a routine basis.

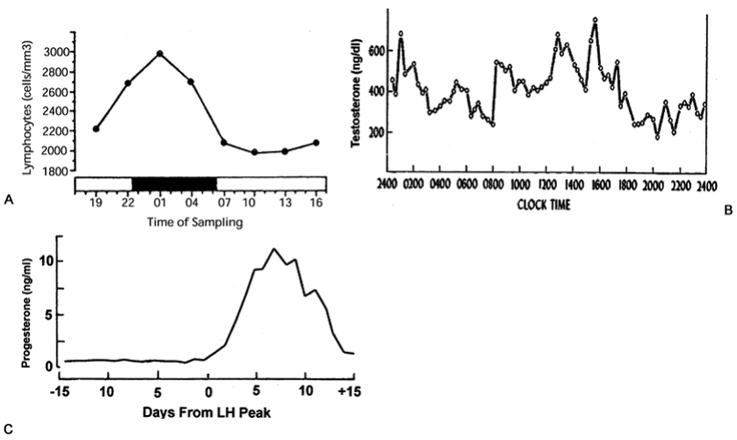


FIGURE 4.2. Biological rhythms in laboratory test results. **A:** The circadian rhythm of the lymphocyte count in whole blood shows a peak (acrophase) at midnight, a trough at 8 AM, and has a period of approximately 24 hours. (From Kanabrocki EL, Sothern RB, Scheving LE, et al. Reference values for circadian rhythms of 98 variables in clinically healthy men in the fifth decade of life. *Chronobiol Int* 1990;7:445, with permission.) **B:** An ultradian rhythm in the serum testosterone concentration is shown in a young adult man who was sampled at 20-minute intervals over the course of 24 hours. (From Griffin JE, Wilson JD. The testis. In: Bondy PK, Rosenberg LE, eds. *Metabolic control and disease*, 8th ed. Philadelphia: WB Saunders, 1980:1535, with permission.) **C:** An infradian rhythm in serum progesterone as demonstrated in a group of cycling women who were sampled daily over the course of their menstrual cycles. (From Thorneycroft IH, Mishell DR, Stone SC, et al. The relation of serum 17-hydroxy progesterone and estradiol-17B levels during the human menstrual cycle. *Am J Obstet Gynecol* 1971;111:947, with permission.)

TABLE 4.2. CIRCADIAN RHYTHMS IN SELECTED ANALYTES

	Analyte[a]	Time of Peak Concentration[b]	% Change[c]
Serum			
	Sodium	1300	2
	Potassium	1100	19
	Glucose	1800	59
	Phosphorus	2200	38
	Urea nitrogen	2300	25
	Cholesterol	2200	11
	Total bilirubin	0700	62
	Total protein	1800	8
	γ-Glutamyltransferase	1000	960
	Thyroid-stimulating hormone	0200	206
	Cortisol	0730	1111
	Melatonin	0300	211
	Iron	1208	32
	Aldosterone	0800	95
Whole blood			
	Total white blood cells	1900	38
	Red blood cells	0430	10
	Lymphocytes	0130	67
	Neutrophils	1700	61
	CD4+ cells	0030	51
Urine			
	Volume	0300	278
	Specific gravity	1600	193
	Calcium	1600	333
	Creatinine	2108	30
	Sodium	2012	54

[a] All analytes demonstrated statistically significant circadian rhythms in healthy middle-aged male subjects (11).
[b] Peak times are given as local clock times.
[c] Percentage of changes in analyte concentration are calculated as [(peak concentration – trough concentration)/daily mean concentration)] × 100.
Data from Kanabrocki EL, Sothern RB, Scheving LE, et al. Reference values for circadian rhythms of 98 variables in clinically health men in the fifth decade of life. *Chronobiol Int* 1990;7:445.

The numbers of different types of blood cells in the circulation and their functional activities show biological rhythms of several frequencies of which the circadian rhythms are the best described (12). Some of these rhythms show large enough amplitudes to be clinically important, especially if consecutive samples from the same patient are to be evaluated. Circadian rhythms in the aggregability and adhesiveness of platelets contribute to a transient hypercoagulability during the morning hours, and this particular rhythm correlates with the peak incidence

of myocardial and cerebral infarctions and sudden cardiac deaths in the morning. One important positive consequence of biological rhythms in the function and activity of the hematopoietic and immune systems is that this rhythmicity provides an opportunity for improving the effectiveness of drug therapies and other treatments by changing the times at which they are delivered.

Ultradian Rhythms

A second type of biological rhythm is ultradian rhythm. This rhythm, which has a period of less than a day, is commonly observed for analytes that are released into the bloodstream intermittently or in pulses. Ultradian rhythms are frequently observed in the blood levels of various glandular secretions. Although the ultradian rhythm for a particular analyte may show some person-to-person uniformity in pulse frequency or duration, the clock times at which pulses occur and the pulse amplitudes can show wide variation from day-to-day within a subject and from subject to subject within a population. The ultradian rhythm in the serum testosterone concentration in a male subject is shown in Fig. 4.2b (13). Considering that the RI for total testosterone is 250 to 900 ng/dl, a single blood sample drawn at, say 0900 would be consistent with a diagnosis of hypogonadism in this subject, whereas a single sample drawn 20 minutes later provides a result that is well within the RI for the test. Ultradian rhythms have the overall effect of widening the RI for all analytes in which they occur because each subject used to establish the test's RVs is likely to have been sampled at a different time relative to his or her own unique pulsatile pattern. Consequently, single test results for analytes that demonstrate ultradian rhythms are often of limited value in the diagnosis and monitoring of disease. In fact, the poor predictive values of single measurements of many blood hormones that are owing in part to their pulsatile patterns of release have led to an emphasis on the use of stimulation and suppression tests in diagnostic endocrine testing. Indeed, the daily rhythms of many blood hormone levels demonstrate both circadian and ultradian components.

Infradian Rhythms

Infradian rhythms have cycle times of greater than 24 hours. The most common example of an infradian rhythm that affects laboratory test results is the menstrual cycle, which is associated with marked rhythmic changes in the levels of many constituents in blood and body fluids including the pituitary gonadotropins and the ovarian hormones (Fig. 4.2c) (14). Substantial clinical and epidemiologic data showing seasonal differences in the incidence of viral disease, depression, malignancy, and suicide suggest the existence of many other infradian rhythms in humans (8). The significance of these other rhythms to the interpretation of laboratory test results remains to be demonstrated. In view of the previous discussion, it is always prudent to consider subject-to-subject differences in biological rhythms that are not accounted for by a test's RVs as a possible cause of an unusual or unexpected result.

Constitutional Factors

Gender (4,6)

Statistically significant gender-related differences in test results have been described for many laboratory parameters; however, most of the male–female differences in test results are not clinically significant and do not require the use of gender-specific RVs. Even for analytes for which gender-related differences eventually develop, test results tend to be similar in males and females until after the age of puberty. The most important general mechanism of male/female differences in test results in adults is sex hormone–dependent physiologic processes. For example, males undergo an androgen-dependent increase in muscle mass compared with females and thus demonstrate increased serum levels of muscle constituents (e.g., creatine kinase) and muscle metabolites (e.g., creatinine). Blood loss during the menses is one basis for gender-related differences in laboratory parameters related to the red cell mass (e.g., red blood cells, hemoglobin, ferritin). The menstrual cycle is also associated with marked changes in the circulating levels of ovarian steroids, which, in turn, influence many physiologic processes. For example, plasma corticosterone concentrations are increased as much as 1.5-fold in the luteal phase compared with the follicular phase. In addition, there are preovulatory increases in plasma renin and aldosterone, and the plasma fibrinogen concentration, as well as the serum concentrations of phosphate and iron, decrease during menstruation.

Age

To consider the effects of age on laboratory test results, it is necessary to divide the population into groups. However, one problem with any categorization scheme based on chronological age is that the biological age of a subject does not necessarily correspond to his or her chronological age. Thus, simple predictions of expected values based on age alone can be problematic. Nevertheless, for the purpose of the following discussion, four groups are considered: infants (younger than 1 year), children (1 to 17 years), adults (18 to 65 years), and the elderly (65 years and older).

The composition of the blood of a newborn varies considerably with the maturity of the infant at birth and changes substantially over the first 24 hours of life (15). Premature infants have lower levels of many tissue enzyme systems, less protein binding, hyperbilirubinemia, and reduced renal function compared with full-term infants. Important changes in the body fat and body water contents occur within the first hours and days after birth. Newborns experience shifts of fluid from the intravascular to the extracellular space followed by a loss of extracellular fluids leading to a decrease in body water content from an initial value in the range of 70% to 75% to the adult value of 50% to 55%. As a result of these fluid shifts, the plasma concentrations of macromolecules, of protein-bound low molecular weight constituents, and of blood cells all increase. The interpretation of laboratory test results in this age group is further complicated by a combination of rapid physiologic changes, the possibility of significant analytical interference by hemolysis and bilirubin, and a general lack of method-specific, age-dependent RVs. Several

excellent compendiums of neonatal RVs are available (16); however, the transferability of published values to analytical methods other than the ones used for their original estimation cannot be taken for granted.

The composition of body fluids changes gradually between infancy and puberty (6). During this interval, the plasma protein concentration increases and the plasma concentrations of enzymes decrease. The onset of sexual development is characterized by progressive increases in plasma levels of gonadal hormones. With the exception of the sex hormones, age-specific RVs usually have little impact on test interpretation in this age group. One significant and rapid change that occurs as children approach the age of puberty is a marked decrease in the activity of the hepatic drug-metabolizing enzyme systems (17). The chronological age at which this change occurs varies from child to child. One consequence of this change is that a previously effective dose of a therapeutic drug (such as phenytoin) may produce toxic drug levels after the rate of drug metabolism abruptly decreases.

Laboratory values remain relatively constant up until menopause in women and middle age in men. The physiologic changes that accompany menopause, which typically occurs between the ages of 40 and 62, are driven by a progressive decline in the ovarian estrogen output, which, in turn, leads to postmenopausal changes in related laboratory parameters including the pituitary gonadotropins and the serum and urine constituents that reflect bone resorption. Although aging per se does not appear to be a major factor in influencing laboratory values in healthy adults, the susceptibility to chronic and debilitating diseases increases during adulthood, and the laboratory values that are found to be characteristic of this age group may not necessarily be consistent with a state of good health. For example, the 95th percentiles for serum cholesterol of 255 mg/dl and 270 mg/dl, which were observed in large samples of healthy middle-aged women and men, respectively, in the Lipid Research Clinics North American Prevalence Study (18) in the early 1980s was far different than the value (200 mg/dl or more) that is now considered to be desirable for all adults. The original sample of healthy subjects apparently included many individuals who were not recognized to be at an increased risk of cardiovascular disease at the time of sampling. Unless the subjects used to establish RVs for adult populations are carefully screened for the presence of subclinical or intercurrent diseases, the RVs may be biased—possibly to the point of being insensitive to the detection of pathological changes in individual subjects.

The elderly constitute the fastest growing age group in our population. These subjects experience a number of physiologic changes that result from aging per se as well as changes that are owing to age-related differences in nutrition and the presence of intercurrent disease. The accurate interpretation of many laboratory tests in this group requires the use of age-specific RVs that have been derived from carefully screened reference subjects (19). A recent study (20) in which investigators measured 47 different laboratory parameters in young adults and healthy individuals in the 60- to 90-year age range and 93 parameters in young adults and healthy centenarians concluded that the RVs for the following analytes showed little or no change with age: urine protein excretion, the serum concentrations of sodium, potassium, chloride, carbon dioxide, pH, total calcium, most serum proteins, liver enzymes, total bilirubin, and magnesium. RVs for the following analytes showed age-related increases both in men and women: serum concentrations of ionized calcium, parathyroid hormone, glucose insulin, C-peptide, luteinizing hormone, follicle-stimulating hormone, γ-glutamyltransferase, urea nitrogen, immunoglobulin (Ig) A, high-density lipoprotein cholesterol, cholesterol, triglycerides, thyroid-stimulating hormone, copper, microsomal antibodies, antithyroglobulin antibodies, lactate dehydrogenase 5, and gastrin. RVs for the following analytes showed age-related decreases: phosphate (men only), IgM, IgD, zinc, T_3, T_4, free and total testosterone (men only), estrone and estradiol (women only), androstenedione, progesterone, hemoglobin, hematocrit, erythrocytes, leukocytes, platelets, and serum iron.

For all age groups, it is important to identify specific analytes for which the accuracy of test interpretation might be improved by the use of age-specific reference limits. However, as described in the section on RVs, it is also important to recognize that such an improvement will only be realized if the interage differences are large enough to result in significant differences in the means and variances of the age-stratified groups.

Genetic Heterogeneity

Genetic heterogeneity within a population leads to person-to-person phenotypic differences that can contribute to the variability in laboratory test results. Racial differences in laboratory values have been described among many populations in various parts of the world. However, in many instances, such differences are not easily verified because it is often difficult to separate effects of race per se from those of other factors. In the U.S. population, African Americans tend to have an increased muscle mass and a more marked skeletal development than their Caucasian counterparts, leading to racial differences in serum levels of creatine kinase and lactate dehydrogenase in adults and in serum alkaline phosphatase in children. African Americans also tend to have higher serum total protein levels and higher serum levels of gamma globulins, α-globulins, ß-globulins, IgG, and IgA than Caucasians (4,6). A more subtle, but perhaps a more significant consequence of genetic heterogeneity are person-to-person differences in the structures of biological macromolecules. Genetic heterogeneity at the molecular level can lead to differences in the reactivity of a subject's DNA, proteins, or cells toward the nucleic acid probes and antibodies that are used as reagents in many diagnostic tests. This type of genetic heterogeneity can result in false-negative findings in molecular biology tests, inaccuracy in immunoassay results, and poor clinical correlation of analyte levels with disease incidence and/or severity. For example, the laboratory determination of mutations within the *BRCA1*gene can identify individuals at risk for a certain heritable form of breast cancer. However, the finding that more than 450 different mutations are associated with this type of cancer makes it practically impossible to exclude the presence of a clinically significant mutation by laboratory testing with currently available techniques (21). Recently, differences with respect to a subject's secretor status and Lewis genotype have been shown to affect the interpretation of serum levels of the tumor

marker CA 19-9 in that Lewis-negative individuals appear to have undetectable CA 19-9 levels even in the presence of substantial tumor burdens (22). A final example is the genetic size polymorphism in apoprotein a, which causes differences in the immunoreactivity of lipoprotein a [Lp(a)] toward analytical assays in which the epitope recognized by the antibody reagent includes a particular variable domain of the apoprotein (23). It now appears that the lack of correlation of Lp(a) levels with cardiovascular risk in some racial groups may be owing, in part, to genotype-dependent inaccuracy in the Lp(a) assay rather than to physiologic differences. As the field of laboratory medicine moves more toward the use of highly specific probes and antibodies in analytical methods, genetic variability will become an increasingly more important consideration in the design and development of test methods and in the interpretation of test results.

Extrinsic Factors

Posture

A subject's posture before and during blood collection affects the concentration of many plasma constituents. A change in posture from supine to standing is accompanied by an increase in the hydrostatic pressure within the vasculature and an efflux of water and small, freely diffusible molecules from the vascular compartment to the interstitial space. This fluid shift leads to a reduction in plasma volume and increases in the concentrations of nondiffusible blood constituents including cells, plasma proteins, enzymes, and protein-bound analytes (e.g., cortisol, T_4, phenytoin, cholesterol). The plasma concentrations of these constituents increase on the order of 4% to 15% within approximately 10 minutes after the posture changes from supine to standing (24). Body water shifts in the opposite direction when the posture changes from upright to supine and the above-named plasma constituents will undergo a corresponding decrease in concentration over the course of approximately 30 minutes. In addition to these nonspecific changes, more marked posture-related differences occur in the concentrations of analytes that are normally involved in the homeostasis of vascular tone. For example, plasma renin, serum aldosterone, and the catecholamines all increase by as much as twofold with a change from supine to erect posture (4,6). Thus, specimen collection procedures for these analytes often specify the posture of the subject during the blood draw as well as the level of physical activity immediately before specimen collection. Posture-related changes should always be considered as a potential cause of day-to-day variability in the plasma or serum concentrations of proteins, cells, and high molecular weight constituents such as lipids and lipoproteins, especially when a particular test result is slightly above or below a reference limit or a clinical decision threshold. The existence of postural effects also provides a rationale for using different criteria for the interpretation of test results in ambulatory and hospitalized subjects.

Exercise (4,25)

Physical exercise can cause numerous alterations in the composition of plasma and urine. Some of the mechanisms of exercise-related changes are shifts in plasma water, decreases in the glomerular filtration rate and urine production, an increased release of hormones that regulate energy metabolism, a generalized stress response, changes in lipoprotein metabolism, and a leakage of macromolecules from cells and tissues. The most marked exercise-related changes occur in the plasma concentrations of hormones such as antidiuretic hormone, catecholamines, growth hormone, glucagon, corticotropin, and cortisol, which can increase by as much as two- to fourfold. Exercise-induced increases in blood hormones typically fall to baseline levels within 1 to 2 hours after the end of the exercise period. However, the changes in aldosterone, renin, and angiotensin can persist for as long as 24 hours. Acute exercise also leads to increases in the plasma levels of pyruvate, lactate, alanine, and free fatty acids as well decreases in arterial pH and PCO_2. Exercise-induced changes in plasma and urine electrolytes and volume usually normalize within 1 hour after exercise. Vigorous exercise can also be associated with a temporary activation of blood coagulation and platelet function and increases in plasma prostaglandin levels. Coagulation factors VIII and Xa also show pronounced increases and platelet factor IV levels decrease. Brief periods of strenuous exercise lead to an increase in the urinary concentration of hemoglobin, and the urine concentration of albumin can increase by as much as 100-fold. Urine protein excretion is greatest during the first 20 to 30 minutes after the exercise period but can remain elevated for as long as 24 hours. Pronounced myoglobinuria can also occur after extreme muscle exertion and persist for as long as 24 to 48 hours. Increases in the plasma concentrations of muscle enzymes like lactate dehydrogenase, creatine kinase, aspartate aminotransferase, and aldolase can range from two- to as high as 40-fold depending on the type of exercise performed. In general, analytes that change with exercise do so to an extent that depends on the degree of strenuousness of the exercise as well as on the level of physical training of the subject. Well-trained athletes are typically more resistant to exercise-related changes than are untrained subjects.

Diet (4,6,26)

The consumption of a standard test meal is followed by increases in the serum concentrations of triglycerides (1.5-fold), aspartate aminotransferase (1.2-fold), bilirubin (1.15-fold), and glucose (1.15-fold). In some cases, meal-related increases are relatively long lasting. For example, ingestion of a protein-rich meal in the evening leads to increases in the blood urea nitrogen, phosphorus, and uric acid that can persist for as long as 12 hours. After a large meal, the serum levels of cholesterol and growth hormone are increased for approximately 1 hour, whereas the postprandial elevation in serum triglycerides can persist for as long as 8 hours. Consumption of diets stressing a particular constituent can cause small changes in some test results. For example, high-fat diets are associated with increases in the serum levels of total cholesterol (1.28-fold) and high-density lipoprotein cholesterol (1.14-fold); high-fiber diets are associated with decreases in serum cholesterol (5%), serum estradiol (15%), and the urinary excretion of albumin (25%); high-protein diets are associated with increases in plasma ammonia, urea nitrogen, and uric acid and an increase in the glomerular filtration rate (1.1-fold).

Prolonged fasting or starvation also affects many test results. For example, the serum concentrations of ß-hydroxybutyrate, acetone, and acetoacetate increase within 14 hours of the start of fasting. After 48 hours of fasting, ß-hydroxybutyrate is increased as much as 30-fold. At this time, free fatty acids and glucagon concentrations are also increased, the serum insulin concentration is decreased, and metabolic acidosis develops. Starvation of 4 weeks' duration is accompanied by 50% decreases in the concentrations of blood urea nitrogen, triglycerides, and γ-glutamyltransferase and 20% to 40% increases in the serum levels of uric acid, creatinine, and aspartate aminotransferase. The recommended duration of fasting before specimen collection is typically 12 hours, and the above data show why deviations from this standard in either direction can interfere with test interpretation.

Caffeine (6)

Of the many dietary constituents that have been shown to affect laboratory values, caffeine is of special importance because of its presence in beverages that are often consumed in large quantities. Caffeine has many physiological actions including the stimulation of the adrenal medulla and the adrenal cortex. Caffeine also has marked effects on lipid metabolism leading to increases in plasma free fatty acids, glycerol, and lipoproteins. Caffeine significantly increases serum gastrin and has powerful diuretic effects that lead to increases in urinary concentrations of erythrocytes, renal tubular cells, calcium, magnesium, sodium, and potassium.

Drugs

It is not uncommon for subjects that are undergoing laboratory tests to be taking one or more prescription drugs. Therapeutic drugs can affect laboratory test results by either interfering with the specific analytical method used for the test (analytical interference) or by causing physiologic changes in the biosynthesis, release, tissue distribution, metabolism, or excretion of the analyte that is being measured. In the former case, the potential for the drug to cause an error in test interpretation will be method specific. In the latter case, an error in interpretation might result from the inability to differentiate a drug-induced physiologic change from a pathologic change. It is the laboratory's responsibility to identify potential sources of analytical and physiologic interference for each test that is offered and to inform clinicians of cases in which relatively common drugs might cause significant errors in interpretation. It is the clinician's responsibility to determine whether and when interfering medications can be withdrawn before specimen collection and to instruct the patient accordingly. Information concerning drug interference with laboratory test results can be found in the manufacturer's package inserts that are supplied with analyzers and test kits, the *Physicians' Desk Reference*, articles from the medical literature, and reference works on the subject of drug interference (7). Drug interference is best avoided by recognizing its potential for occurrence, withdrawing medications before sampling whenever possible, and evaluating any suspicious results in light of a subject's medication history.

Alcohol Use

Alcohol use, both acute and chronic, are important causes of drug-related changes in test values. Some of the more prominent interactions and their magnitudes are summarized in Table 4.3 (4,26).

TABLE 4.3. ACUTE AND CHRONIC EFFECTS OF ALCOHOL ON SELECTED LABORATORY RESULTS

	Test	-Fold Change[a]
Acute (within 2–4 h)		
	Aldosterone	1.5
	Triglycerides	1.3
	Prolactin	0.5
	Cortisol	0.5
	Cholesterol	0.9
	Osmolality	inc
	Glucose	1.0 to 1.5
	Testosterone	dec
	Luteinizing hormone	inc
	Catecholamines	inc
Chronic		
	γ-Glutamyltransferase	10.0
	Aspartate aminotransferase	2.5
	Alanine aminotransferase	1.6
	Estradiol	1.6
	Cortisol	1.6
	Triglycerides	1.3
	Cholesterol	1.1
	Mean cell volume	1.1
	High-density lipoprotein cholesterol	1.2
	Iron	1.8

[a] Alcohol-related changes in serum chemistry results and in the mean cell volume are expressed as -fold changes compared with control subjects. Values <1.0 indicate alcohol-related decreases; values > 1.0 indicate increases.
inc, increase of unspecified magnitude; dec, decrease of unspecified magnitude.
Data from refs. 4, 6, and 26.

Oral Contraceptives (27)

The oral contraceptive preparations that are used clinically contain different combinations of estrogens and progestins or progestins alone. These preparations can influence the results of many laboratory tests. For example, estrogen-containing oral contraceptive preparations affect hepatic protein biosynthesis leading to increases in the plasma concentrations of fibrinogen, and several of the serum proteins that are involved in hormone transport including the binding globulins for cortisol, thyroxine, and sex hormones, and many of the clotting factors including VII, VIII, IX, and X. Many of these drug-related changes have physiologic significance. For example, increased amounts of coumarin derivatives are required to prolong the prothrombin time of subjects taking oral contraceptives, and the platelet aggregation response to catecholamines in these individuals is also increased. The magnitude of the effects of a particular oral contraceptive preparation on particular laboratory tests is often proportional to the estrogen content of the preparation. A sample of the more than 200 interactions of oral contraceptives with laboratory tests that have been reported in Young's reviews (4,7) are presented in Table 4.4.

TABLE 4.4. EFFECTS OF ORAL CONTRACEPTIVES ON SELECTED SERUM ANALYTES

Analyte	-Fold Change
Apoprotein A 1	1.1
Apoprotein B	1.2
Cholesterol	1.1
High-density lipoprotein cholesterol	1.1
DHEA-S	0.7
Luteinizing hormone	0.1-0.2
Prednisolone clearance	0.5
Protein S	0.75
SHBG	2.5-3.0
Testosterone	0.7-0.8
Transcortin	2.0
Triglycerides	1.5
Ferritin	1.6
Follicle-stimulating hormone	0.7
Iron	1.2

Oral-contraceptive-related changes in serum chemistry results and in the mean cell volume are expressed as-fold changes compared with control subjects.
Data from refs. 4 and 7.

Pregnancy

The major physiologic adaptations of normal pregnancy including expansions of plasma volume and erythrocyte mass, an altered distribution of cardiac output, an increase in plasma protein synthesis, hyperventilation, and marked endocrinologic changes lead to significant changes in many laboratory values compared with the nonpregnant state (28). Consequently, the accurate interpretation of test results during pregnancy often requires the use of RVs derived from a population of healthy pregnant women at defined stages of gestation. Pregnancy-related differences in some of the more commonly ordered laboratory tests that were selected from Young's review (4) are shown in Table 4.5. Several sources of pregnancy-specific RVs are presented in ref. 28.

TABLE 4.5. CHANGES IN SELECTED LABORATORY VALUES DURING NORMAL PREGNANCY

	Laboratory Value	-Fold Change
Serum		
	Albumin	0.75
	α_1-Acid glycoprotein	0.66
	Alkaline phosphatase	2.0–3.0
	CA 125	2.0
	Creatinine	0.7
	1,25 dihydroxy vitamin D	2.0
	Erythropoietin	2.0
	Estradiol	10.0
	Free fatty acids	1.7
	Ferritin	0.25
	Iron	0.65
	Sex hormone binding globulin	6.0–10.0
	Thyroxine-binding globulin	2.0–3.0
	Transferrin	1.7
Blood		
	Red blood cells	0.9
	Factors II, V, VII, VIII, IX, X	inc
	Erythrocyte sedimentation rate	inc
	Fibrin split products	inc
	D-dimer	3.5
Urine		
	Glomerular filtration rate	1.5
	Estriol	inc
	Hyaline casts	inc
	Albumin	1.6
	Alanine	2.0

Pregnancy-related changes in serum chemistry results and in the mean cell volume are expressed as-fold changes compared with control subjects.
Data from refs. 4, 24, and 28.

Intercurrent Illness

Specific medical conditions such as shock, malnutrition, fever, thermal burns, trauma, recent transfusions, and component therapies all have distinct effects on the composition of blood and body fluids and can confound the interpretation of disease-related changes in laboratory values when one of these medical conditions is present along with a primary disease. The specific effects of these relatively common conditions on test values are beyond the scope of this chapter. Interested readers can find more information in one of the comprehensive reviews on the subject (5,6).

THE ESTABLISHMENT AND VERIFICATION OF RVs

Part of "4 - The Interpretation of Laboratory Tests"

The diagnosis and management of disease are accomplished through the evaluation of information gained from many sources including the history and physical examination, various imaging and endoscopic techniques, histopathologic examinations, and clinical laboratory tests. Regardless of the diagnostic modality, the medical significance of the result observed in an individual under investigation is usually determined by comparison with some control or RV. The process of interpretation through comparison has led to the emergence of a more objective and scientific basis for the process of medical decision making and is the foundation for future advances in the field of evidence-based medicine. Although the use of RVs in laboratory medicine would appear to be well established and relatively straightforward, several challenges and controversies still remain in this important area of laboratory practice. This section, portions of which have been previously published (29), is intended to provide a brief overview of some of the major philosophical and practical aspects of establishing and using RVs in the clinical pathology laboratory. More detailed treatments of the subjects presented here can be found among the references, which include several authoritative reviews (30,31,32,33 and 34).

Development of Standardized Terminology

In the early days of laboratory medicine, the standard of comparison for a test result was the “normal” value or the “normal” range. In this context, the word normal was intended to indicate that the values were the usual, customary, or typical ones that were associated with good health. Unfortunately, this terminology had several drawbacks. Of relatively minor importance was the fact that the word normal might be taken to imply that the laboratory value under discussion was normally distributed (i.e.,

gaussian) in the population. A more significant limitation of the terminology was that terms like normal range, normal value, and upper limit of normal implied that there was some clear-cut value that could effectively discriminate between health and disease. Although most health care workers will agree that there is rarely a stark dividing line between health and disease, there is still a tendency to interpret laboratory results in a binary fashion as signifying either health or disease using the normal value as the cut point. The previous terminology, which fostered this illogical binary approach to interpretation, also had the negative consequence of causing unnecessary alarm and anxiety to a subject whose laboratory result was found to be abnormal.

In the past decade, there has been increasing support to replace the older terminology with a nomenclature that permits the unambiguous description and discussion of the subject. The new terminology, which is now endorsed by the World Health Organization, the International Federation of Clinical Chemistry, and the National Committee for Clinical Laboratory Standards (NCCLS), is based on the fundamental definition of a RV as: A value (test result) obtained by the observation or measurement of a particular type of quantity in an individual selected on the basis of well-defined criteria.

This definition is part of a larger framework of standardized terms (Fig. 4.3) that apply to all phases of the process by which RVs are established and used (32).

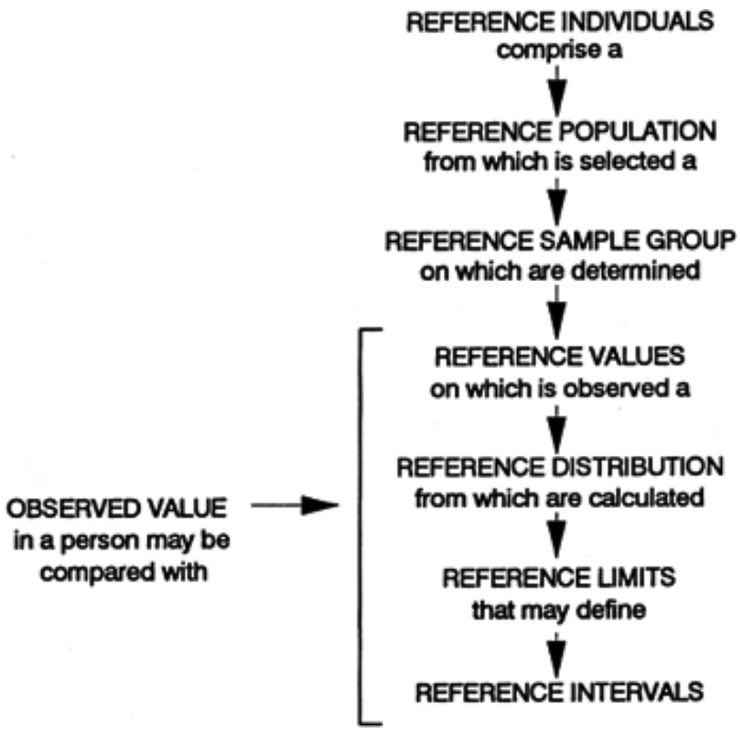


FIGURE 4.3. Standard terminology for the description and discussion of reference values. These terms, which refer to various aspects of the establishment and use of reference values, were proposed by the International Federation of Clinical Chemistry and the International Council for Standardization in Hematology and are endorsed by the World Health Organization and other organizations worldwide. [Redrawn from National Committee for Clinical Laboratory Standards. *How to define and determine reference intervals in the clinical laboratory; approved guideline C28-A*. Villanova, PA: NCCLS, 1995.]

Types of RVs Currently in Use

The RV that is used most frequently in contemporary laboratory medicine is the RI, which is bounded by upper and lower reference limits that are usually chosen to enclose 95% of the values observed in nondiseased subjects. Population-based, health-associated RVs of this type are useful in the process of medical decision making in particular settings, for example, when screening for the presence of a disease in a population in which affected individuals do not present with any obvious symptoms (35). However, only a small percentage of the laboratory tests performed are ordered for this purpose, the majority being used to monitor the progress of a disease process or the effectiveness of a treatment. In the latter situations, other types of RVs are typically more informative. Therefore, the laboratory must consider the possibility that disease-associated RVs, subject-based RVs, multivariate RVs, and clinical decision limits might all have a place among the RVs that it presents to its clients. The ultimate goal is to provide the RV that best supports the medical decision at hand.

Accreditation Requirements

Determining which type of RV is most efficient for a particular test or for a particular clinical setting is the task and responsibility of the laboratory director. Among the guidelines that influence the director's activities are the standards set by the federal government and by voluntary accreditation agencies like the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) and the College of American Pathologists (CAP). The Clinical Laboratory Improvement Act of 1988 (CLIA) regulations contain three standards related to reference values (36). The first (§493.1109) requires that pertinent reference or normal ranges as determined by the laboratory performing the tests must be available to the authorized person who ordered the tests. The second standard (§493.1211) requires that reference ranges or normal values must be included in the laboratory procedure manual. The third standard (§493.1213) requires that before reporting patient results, the laboratory must establish for each method, the specifications for a number of performance characteristics including the reference range(s). Further definitions in this subpart of the standard state that in cases in which a manufacturer's reference range can be used, the laboratory that uses it must verify that the range is appropriate for its patient population. Furthermore, for methods developed in-house, modifications of a manufacturer's test procedure, or a manufacturer's procedure not cleared by the U.S. Food and Drug Administration (FDA) as meeting CLIA requirements, the laboratory must verify or establish reference ranges. Finally, the laboratory is required to have documentation of the verification or establishment of all applicable test performance specifications including the RVs.

The CAP standards related to RVs (37) include a requirement that reference ranges be verified or established and documented as a part of the performance specifications for moderate complexity tests introduced after July 31, 1998, and high complexity tests introduced after July 1, 1992 (01.4208, Phase II). A note appended to the standard specifies that, when reasonable, RVs must be developed for each analyte and specimen type and acknowledges that for many analytes, literature references or information from a manufacturer's package insert may be appropriate. A second standard (01.4211, Phase II) requires that the laboratory periodically evaluates the appropriateness of its RIs and takes corrective action if necessary. A note appended to this

standard specifies that there should be a reevaluation of RVs whenever a laboratory changes an analytical methodology. Finally, a third standard (01.4120, Phase II) requires that RVs be reported with the test result. Under some circumstances, this last standard might be satisfied by distributing a list of RVs to all users and sites where laboratory reports are received.

JCAHO (38) requires that RIs or normal ranges for each test be included in the clinical record as part of the laboratory report or by including a current list of values that are approved by the laboratory director (IM.7.7). This standard emphasizes the possibility of variation in RVs from laboratory to laboratory or with different methodologies or age groups, requires that the units of concentration or activity are included in reports, emphasizes that RVs should be updated as methods change, specifies that RIs be furnished for referred test, requires that documentation is available in the laboratory to show the basis for the reference ranges and is available to the clinical staff on request, and requires that reference ranges are appropriate to the population served by the laboratory.

Generally, these federal regulations and voluntary standards give the laboratory director the flexibility to define the types of RVs that will be used and the procedures by which the values will be defined and determined. However, the various standards do require three major activities on the part of the laboratory: the RVs must be established or verified by the laboratory that uses them, the RVs must be presented to the users of the laboratory along with patient results, and the RVs and the details of their establishment must be documented. The standards do not necessarily require that a laboratory establish its own RVs *de novo*. Instead they imply that the verification and transfer of an existing RV (from a manufacturer or from the literature) is an acceptable alternative. In this case, it is important to consider the types of activities that comprise the process of verification. Fortunately, recently published NCCLS guidelines (32), which are presented later in the chapter, recommend three possible methods for verification. Although there may be some room for debate, it would appear that verification must include the testing of a sample from the laboratory's own reference population to comply with the CAP standards.

Principles for the Establishment of RVs

Specifications

It is important to consider that a RV has little meaning unless it is accompanied by a complete set of specifications that describe precisely how it was derived. A comprehensive definition of the RV serves as documentation for users of the laboratory and for accreditation purposes and also provides the information that another laboratory would require to evaluate the acceptability of a transfer. The following details should be specified to define completely a RV: the characteristics of the reference population and the way in which the reference sample group was selected; the physiologic and environmental conditions under which the specimens were obtained; the techniques and timing of specimen collection, transport, preparation, and storage; a description of the analytical method used including its accuracy, precision, and quality control procedures; the original results obtained for the reference sample; the statistical methods that were applied to the data; and finally, the RVs that were derived (39).

Literature Review

The first step in any reference range study should be a comprehensive review of the literature with the intent of identifying any factors that might affect the observed value of the analyte of interest. Such factors must be known at the very outset of the study to guide the development of exclusion and partition criteria and to develop procedures for patient preparation, specimen collection, and specimen storage. In cases in which an entirely new analyte or analytical method is being addressed, this step might even require the investigation of the potentially significant factors by experimentation.

Exclusion and Partition Criteria

Exclusion (or selection) criteria are used to reduce the heterogeneity of a reference sample by excluding subjects that might have unusual analyte values because of the presence of disease or some unique constitutional or extrinsic factor. The most common criteria are the sources of biological variability listed in Table 4.1. Of course, the criteria used in any given study will depend on the analyte under investigation and the objectives of the study. Partitioning criteria are used to divide the reference population into subclasses that, for one reason or another, are expected to have different and distinct distributions of the analyte of interest. Again, reference populations are often subdivided based on constitutional and extrinsic factors like those presented in Table 4.1. Note that once the decision to partition a reference population is made, each subclass is henceforth treated separately and independently, thus multiplying the resources required to complete the study. Therefore, as emphasized in the first part of the chapter, it is important to partition only if subclass-specific RVs provide a significant improvement in the diagnostic attributes of the test. Although this is often a difficult judgment to make at the outset of a study, it can be assessed after some data are collected by comparing the means and standard deviations of the proposed subclasses. If the ratio of the standard deviations of two proposed subclasses is less than 1.5 (or if an *F* test on their variances is not significant) and if there is no significant difference in the subclass means by the standard normal deviate test (or the *t* test), then a single RV should be determined for the combined group (32,34).

Once appropriate selection and partition criteria have been chosen, a mechanism for applying them to the selection process must be devised. The most effective way to implement the criteria is to evaluate the population by means of a well-designed questionnaire. Guidelines for the preparation and administration of a questionnaire and an example of an effective format are presented in a recent NCCLS document (32). Before evaluating any patients, the laboratory should consult with its institutional review board to determine whether it is necessary to obtain informed consent from the reference subjects. The consent form, the questionnaire, and the reference study itself may require prior approval by this board.

Selecting the Reference Sample

The sampling of individuals from the reference population is one of the most critical aspects of the RV study. The major considerations during this step are the selection of a random sample, whether the subjects will be selected prospectively or retrospectively, and the minimum number of subjects required. Ideally, the reference sample should be randomly drawn from the population. This means that every member of the defined population should have same chance of being included in the reference sample. Unfortunately, it is practically impossible to obtain a completely random sample. Usually, the closest approximation is to randomly select subjects from the reference population and then to isolate from this group those subjects who meet the selection criteria for the study (34). Although it is generally assumed that this compromise is not likely to seriously influence the results, the possibility of obtaining biased RVs increases with more serious violations of the principle of random sampling.

Reference individuals can be selected in either a prospective or a retrospective fashion (33), and when properly performed, both approaches will provide reliable results. The critical aspect of the selection process is that individuals are selected based on defined criteria that are independent of the value of the analyte for which the RVs are being determined. Both of these methods of sampling are considered direct methods because it is the individuals or subjects who are selected and not their laboratory results. In the retrospective mode, direct sampling is possible only when the subjects are selected from a database that contains analytical results and sufficient clinical and demographic data so that appropriate exclusion and partitioning criteria can be used to select a reference sample corresponding to a defined population.

Over the years, there has been considerable debate concerning the use of indirect approaches to the selection of reference samples. In the indirect methods, RVs are established from laboratory results (not subjects) that have been selected from a database. The database typically consists of all the results reported by the laboratory or all the results observed on preadmission tests, and various statistical techniques are used to extract the normal values from the total distribution (40). However elegant the statistical approach, the RVs derived by this method can be used for little more than deciding whether a future patient result is extreme or unusual relative to those of the patient population served by the laboratory. The contribution of such information to the process of medical decision making seems questionable. Indirect sampling is not widely accepted as a valid technique for establishing RVs. However, this method might serve as a quality control monitor for the detection of changes in analytical accuracy over time.

What is the minimum number of reference subjects required to estimate population-based RVs for a new analyte or method? The values recommended in the literature range from 30 to 700 subjects (32,41,42)! This wide range is owing in part to the fact that the precision in estimating population values depends not only on the sample size but also on the interindividual variability of the analyte, the skewness of the distribution of the analyte values in the population and on the statistical method used to estimate the population percentiles (30,34,41). The minimum number of reference subjects required to estimate a 95% RI by the nonparametric technique (see Nonparametric Method section) is 39; however, at this low value, the smallest and largest observations represent the 2.5th and 97.5th percentiles, and it is not possible to calculate confidence intervals (CIs) for the reference limits. The International Federation of Clinical Chemistry and Laboratory Medicine and NCCLS now recommend a minimum sample size of 119 for estimation of RVs by the nonparametric technique (32). This is the smallest value of *N* that permits the estimation of 90% CIs around the upper and lower reference limits. A sample size of 119 has been shown to be reasonable when the distribution of analyte value is symmetrical, but it may be inadequate when the distribution is skewed (41). In the latter case, as many as 700 reference subjects may be required for a precise nonparametric estimate of the 95% RI. Similar calculations for parametric estimates indicate that sample sizes between 50 and 450 are required for precise parametric estimates (41). The practical approach to the issue of sample size is twofold: use the largest sample size possible and always calculate CIs for the reference limits that are estimated. A wide CI indicates poor precision and suggests that the limit may not effectively discriminate an extreme or pathologic test result from one that is just within the RI. Unacceptable precision in an estimated limit can be addressed by increasing the size of the reference sample, partitioning the reference population, or reducing preanalytical and/or analytical variability for the analyte.

Preparation of Subjects, Specimen Collection, and Sample Analysis

The preparation of the reference subjects before sampling should follow defined conditions of prior diet, physical activity, stress, length of fasting, and so forth. The specimen collection procedure must provide for consistent site preparation, posture, and environmental conditions. In addition, specimen handling must be consistent with respect to variables such as the timeliness of separation or processing and the duration and conditions of transport and storage. A consideration of these factors will serve to minimize the biological and preanalytical variability of the analyte. It is important that the conditions that are specified for the RV study are also attainable and practical for routine use in the setting in which future patient samples will be obtained.

The primary concern during the analytical phase of the study is that the test method has been fully characterized with respect to its precision, accuracy, minimum detection limit, linearity, recovery, and interference. There must also be a program in place to assure continuity of these performance characteristics into the future. It is clear that changes in analytical performance characteristics can significantly affect the validity of a previously established RV (43).

Data Analysis

There are two general statistical methods for the estimation of population-based reference limits and their CIs: nonparametric and parametric. The nonparametric method is considered to be the easiest to perform because the desired parameters are obtained by simply ranking the RVs in ascending order and choosing values corresponding to particular rank numbers. The nonparametric

method is applicable to any reference distribution regardless of its shape or symmetry, and it can be used with noncontinuous variables (i.e., titers). One disadvantage of this method is that the reference sample must consist of at least 119 subjects to estimate 90% CIs of reference limits. The nonparametric method is currently recommended by NCCLS for establishing population-based RVs (32).

The parametric method requires that the original reference distribution or some mathematical transformation thereof conforms to a gaussian (i.e., normal) distribution. This method is more complicated than the nonparametric method because the reference distribution must be tested for "normality", transformed if required, and retested for "normality" before the estimation of reference limits. If the data are not normally distributed, the use of the parametric technique will give erroneous results. However, when this condition is satisfied, the procedure for estimating reference limits is simple and straightforward. Major advantages of the parametric method are that 90% CIs can be calculated for samples of all sizes and that the precision of the estimated reference limits is generally better than that of the nonparametric method for any sample size (41). Many of the statistical software packages available for the personal computer contain programs for testing a distribution for normality, as well as the means for applying mathematical transformations to a data set. Thus, the tools required to use the parametric method are available to most laboratorians.

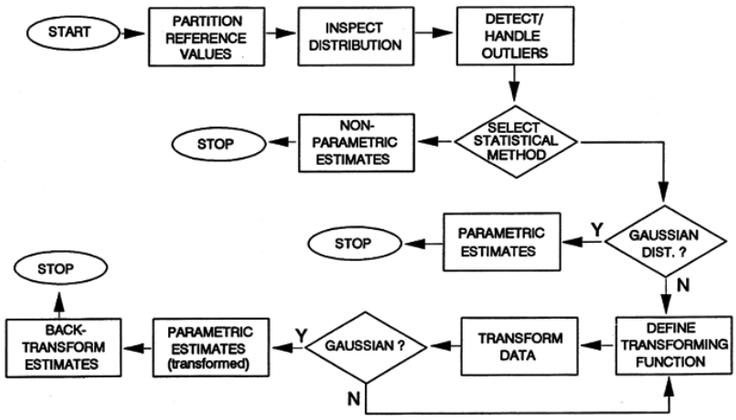


FIGURE 4.4. Statistical treatment of reference values for the estimation of population based reference intervals. (Redrawn from Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. *Tietz textbook of clinical chemistry*. 2nd ed. Philadelphia: WB Saunders, 1995:454.)

Estimating Reference Limits: Some Examples

An outline of a procedure for estimating population-based RIs is presented in Fig. 4.4. As an example of the use of the different techniques, consider a set of values for the serum IgG concentration in a reference sample group of 139 subjects. A partial listing of the data arranged in order of ascending IgG concentration is presented in Table 4.6. The first step in the analysis of any set of observations is the visual inspection of a histogram of the data (Fig. 4.5). This inspection will indicate whether the distribution is unimodal or multimodal, and the shape and symmetry of the distribution will provide the first indication of whether it is normally distributed. Visual inspection will also identify observations that are extreme or atypical compared with the bulk of the data. These extreme results or *outliers* can arise as the result of an analytical error or the misclassification of a reference individual. Because both statistical methods for estimating reference limits are sensitive to the effects of the outliers, is important to have an objective means for detecting them. Among the many methods that are available, a simple test devised by Dixon is the most widely recommended (32). This test involves calculating the ratio D/R, where D is the absolute difference between the highest (or lowest) and the second highest (or lowest) value in the distribution and R is the range of all the values including the extreme

one. If D/R > 0.33, then the extreme value is considered an outlier. Visual inspection of the histogram in Fig. 4.5 suggests that the data are skewed to the right and that the highest IgG value (17.7 g/L) might be an outlier. From the data in Table 4.6, D/R = (17.7 - 15.5)/(17.7 - 5.7) = 0.18.

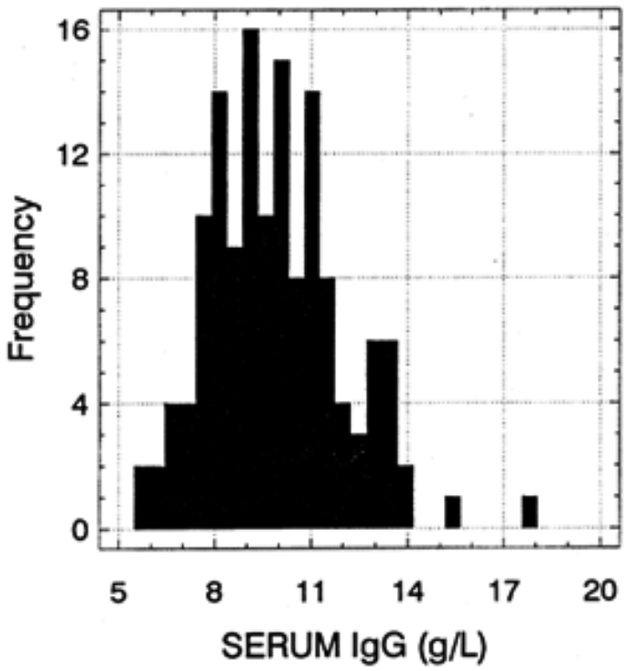


FIGURE 4.5. Reference distribution for serum immunoglobulin (IgG) G concentration. Serum IgG concentrations from 139 reference individuals are displayed in a histogram format. Visual assessment indicates that the data are skewed to the right and may contain several outliers.

TABLE 4.6. PARTIAL LISTING OF SERUM IMMUNOGLOBULIN G RESULTS FROM A REFERENCE SAMPLE OF 139 SUBJECTS

Rank	Immunoglobulin G (g/L)	Rank	Immunoglobulin G (g/L)
1	5.7	130	13.2
2	5.7	131	13.3
3	6.3	132	13.4
4	6.4	133	13.4
5	6.7	134	13.4
6	6.8	135	13.6
7	6.8	136	13.8
8	6.9	137	13.9
9	7.1	138	15.5
10	7.2	139	17.7

The immunoglobulin G results shown in the histogram in Fig. 4.5 were sorted in ascending order and assigned rank numbers of 1 to 139. The results occupying the 10 lowest and the 10 highest ranks are shown here.

Because the extreme value does not violate the one-third rule, it is retained.

Nonparametric Method

For the nonparametric estimates, the observations for the reference sample are arranged in ascending order and assigned rank numbers such that the lowest value is ranked 1 and the largest is ranked n (Table 4.6). The limits that enclose 95% of the reference population are estimated by identifying the rank numbers corresponding to the 2.5th percentile (r1) and the 97.5th percentile (r2) as follows (33):

$$r1 = 0.025(n + 1) = 0.025(140) = 3.5 \qquad (1)$$

$$r2 = 0.975(n + 1) = 0.975(140) = 136.5 \qquad (2)$$

Rank 3.5 corresponds to an IgG value midway between those that occupy ranks 3 and 4. Thus, by interpolation r1 = (6.3 + 6.4)/2 = 6.35. A similar interpolation of rank 136.5 gives a value of r2 = (13.8 + 13.9)/2 = 13.85. Thus, 6.4 and 13.9 are the estimated lower and upper reference limits, respectively. Ninety percent CIs for these limits depend on the sample size and are determined by consulting a table such as that presented in reference 34. For a sample size of 139, the CI for the 2.5th percentile is bounded by values corresponding to ranks 1 and 8 (5.7–6.9 g/L) and that for the 97.5th percentile, the values corresponding to ranks 132 and 139 (13.4–17.7 g/L).

Parametric Method

The first step in the parametric method is to test the assumption that the distribution of the RVs is normal. Simple tests for a normal distribution include the assessment of distribution shape; the comparison of the mean, median and mode of the distribution; and plotting of the data on normal probability paper (Fig. 4.6). More objective assessments include the evaluation of skewness and kurtosis and application of the chi-square test, the Kolmogorov-Smirnov test, or the Anderson-Darling test. These tests, which are available in many personal computer–based statistical software packages, are described in detail elsewhere (33,34). When the original data are normally distributed, the percentiles can be calculated directly as described below. If this is not the case, then the original data must be transformed. Transformation of many biological variables (which are often skewed

to the right) to normality can often be accomplished by taking the logarithm (y = lnx) or the square root [$y = (x)^{1/2}$] of the original results. If these simple transformations fail, more complicated single or sequential transforming functions can be applied (30,33,34). The success of any transformation must be verified by retesting using one of the graphical or statistical tests described above. The 2.5th and 97.5th percentiles are calculated from either the original or the transformed data as follows (33):

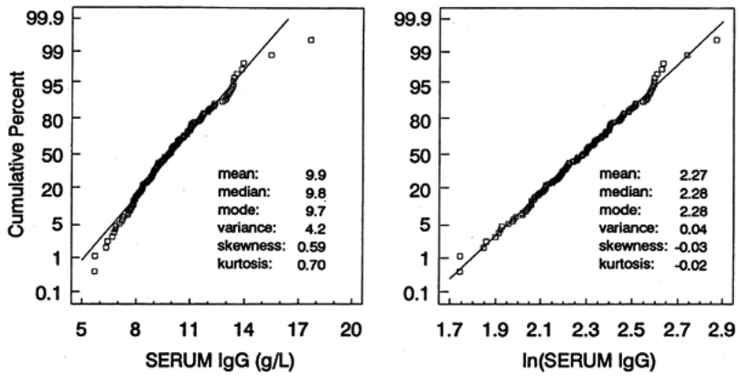


FIGURE 4.6. Evaluation of reference distributions of original and transformed data. **A:** The original reference values from Fig. 4.5 were analyzed by calculating descriptive statistics and by plotting on normal probability paper. The nonequivalence of the mean, median, and mode; the nonzero values for skewness and kurtosis; and the nonlinearity of the normal probability plot indicate a nonnormal distribution for the original data. **B:** The evaluation of the same criteria show that the reference distribution was transformed to "gaussianity" using a logarithmic transformation (y = lnx).

$$2.5\text{th percentile} = \bar{x} - 1.960(s) \qquad (3)$$

$$97.5\text{th percentile} = \bar{x} + 1.960(s) \qquad (4)$$

where [x with bar above] is the mean of the sample, s is the standard deviation, and 1.96 is the multiple for a standard normal deviate. (When n is less than 30, the t value for the 0.05 probability level at n degrees of freedom should be substituted for 1.960.) The 90% CIs for each percentile are estimated as:

$$\text{percentile} \pm 2.81[(s/(n)^{1/2}] \qquad (5)$$

In the IgG example presented above, the distribution of IgG values was nonnormal by both subjective and statistical criteria; however, the natural logs of the original data were normally distributed (Fig. 4.6). Using the mean and standard deviation of the transformed data, the percentiles and their 90% CIs are calculated as:

2.5th = 2.27 - 1.96 (0.20) = 1.87

CI = 1.87 ± 2.81 (0.20/11.78)

97.5th = 2.27 + 1.96 (0.20) = 2.67

CI = 2.67 ± 2.81 (0.20/11.78)

Back transformation by taking the exponents of each parameter gives a lower limit of 6.5 g/L (CI, 6.2 to 6.8) and an upper limit of 14.3 g/L (CI, 13.8 to 15.1).

Transference of RVs

The *de novo* establishment of RVs can be both difficult and costly, hence many laboratories have sought to adopt or transfer RVs from the literature or from the manufacturer's of instruments and reagent systems. In fact, it has been estimated that this practice may well account for the majority of RVs in use in clinical laboratories today (32). The act of transference of a RV from its originator to a receiving laboratory appears to be synonymous with the act of verification, as stated in the CLIA regulations and the CAP standards. One prerequisite for a valid transfer is that the original RV has been completely specified and that the original RV study was correctly performed. This being the case, the validity of the transfer hinges on two factors: the comparability of the analytical systems used in the original and the receiving laboratories and the similarity of the patient or client populations served by the two laboratories. If the analytical systems in the two laboratories are not identical, then the first issue should be addressed using studies and acceptance criteria such as those recommended in NCCLS document EP9-A. The second issue should be addressed by evaluating the match between the specifications of the original study and the conditions and the characteristics that prevail in the receiving laboratory and its patient population.

The minimum requirement for the verification (validation) of a transferred range is documentation showing that the pertinent factors of the original RV study have been reviewed and found to be applicable in the receiving laboratory (32). Depending on the circumstances, the receiving laboratory may either wish to or be required to further verify applicability of a transferred range by testing a small sample of reference individuals from its own population. Obviously, the small sample should be selected according to the principles described previously and its characteristics should match those of the original reference population in every significant way. Two options for sampling and statistical analysis have been recommended by NCCLS. The first involves the testing of 20 reference individuals from the receiving laboratory's subject population. These test results are examined for homogeneity, and outliers are excluded and replaced with new subject values to keep the total number of observations at 20. The transfer is considered to be verified if no more than two of the 20 values fall outside the 95% reference limits established in the original study. This critical value for rejection is derived from binomial probability theory, which holds that when the 95% RIs of the two laboratories are identical, the probability that more than two of 20 results will be outside the original RI is approximately 0.075 (30). If three or more values fall outside the limits, a second group of 20 subjects is evaluated by the same criterion. If the verification fails a second time, then the receiving laboratory should establish its own RVs in a full-scale study. The second option involves the testing of 60 reference individuals from the receiving laboratory's subject population. The criteria for verification—no significant difference in the means or variances of the original reference sample and the test sample—are the same as those recommended for evaluating the need to partition a reference sample into subclasses (44).

Recently, a novel statistical method involving Monte Carlo sampling was proposed for use in the verification of the transfer of RVs. This nonparametric method can be used to simultaneously evaluate multiple statistical parameters (e.g., means, medians, percentiles) and thus can provide for a more extensive comparison of an original and a test distribution. The theory behind this technique and examples of its application are presented in reference 45. The recently published NCCLS guidelines and the Monte Carlo method now provide several possible alternatives for evaluating objectively the acceptability of RV transfers. There is still a need for more research in this area to refine the current approaches and to develop techniques for the evaluation of more difficult cases including the transfer of RVs between different methods and different analytical systems.

Multivariate RIs

The comparison of a single laboratory result with its population-based RI is called a univariate comparison. When results from many different tests are each compared in this fashion, multiple univariate comparisons are made. Indeed, this is the way that panels of test results are usually evaluated. A major problem with this practice is that when 95% RIs are used, we expect 5% of the reference population to fall outside the interval. Therefore, when n multiple univariate comparisons are made, the probability that one or more of the results will fall outside of its univariate RI is:

$$p = (1 - 0.95^{n}) \qquad (6)$$

Thus, for five and 10 test panels, this probability would be 0.33 and 0.40, respectively. One way to solve this problem is to develop multivariate RVs that are based on the joint distribution of two or more variables in the reference population. In the bivariate case, the RI consists of a two-dimensional reference surface. In the trivariate case, the RI is a three-dimensional ellipsoid body. In the n-variate case, the RI is an n-dimensional body in hyperspace (46).

Multivariate RIs are established according to the same guidelines as univariate ones; however, the statistical analysis of the data as well as the subsequent use of the intervals requires computer programs designed for this purpose. The methods and conventions required for the establishment of multivariate RIs are discussed elsewhere (30,46). There are only a few instances in which bivariate RIs are routinely used in laboratory medicine. Our knowledge of the operating characteristics of tri- and n-variate RIs is even more limited. In general, the experience with multivariate RIs indicates that they do reduce the rate of chance positive findings among the reference population compared with the multiple univariate approach. A test battery can be highly unusual in the multivariate sense even though each of the individual test results is within its univariate RI. Conversely, the multivariate result can be within its RI even though a result for one of the variates is extremely unusual. Furthermore, multivariate RIs, which are typically population based, show the same insensitivity to significant changes in individual subjects as do univariate RIs (see Subject-Based RV's). The potential to use multivariate RIs effectively will surely increase with the data-processing capabilities of laboratory and hospital information systems; however, each potential application will require a thorough evaluation to determine whether the approach actually improves the diagnostic efficiency over that of the multiple univariate mode.

Subject-Based RVs

Individuality of Biological Variables

Although clinical laboratories devote considerable effort to the establishment and verification of population-based RVs, this type of reference system often lacks the sensitivity to detect clinically significant changes in an individual subject. This is owing to the fact that for many biological variables, the average amount of intraindividual variability over time is much less than the subject-to-subject variability (47). An example of this is shown in Fig. 4.7, in which the distribution of the means and ranges of serially measured IgG concentrations from single subjects are plotted (48). Considering that the population-based RI for this analyte is 600 to 1,200 mg/dl, it is apparent that a subject could have a serum IgG concentration that would be considered extreme relative to his or her own personal mean and yet still fall well within the population-based RI.

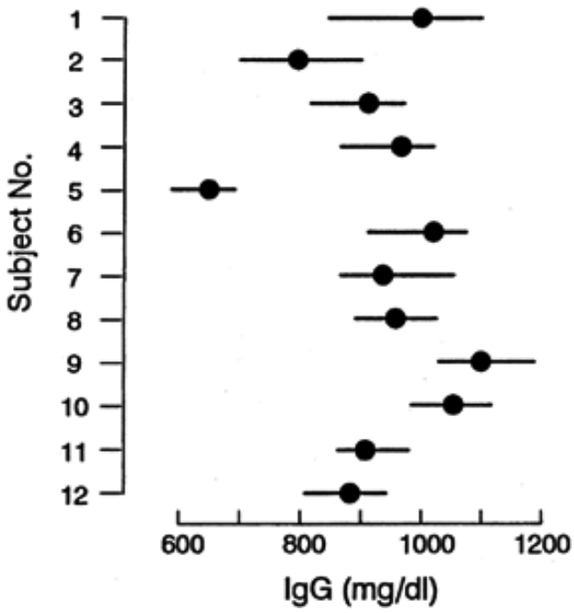


FIGURE 4.7. Intraindividual variability of serum immunoglobulin (IgG) concentration in 12 subjects. The means and ranges of serum IgG results from 12 subjects sampled on 10 occasions over a 3- to 5-month period are plotted in relation to the 95% population-based reference intervals of 600 to 1,200 mg/dl. Note that the range of values for any one individual encompasses only part of the conventional reference interval. (Redrawn from Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* 1989;27:409.)

One way to describe the degree of personal uniqueness or the individuality of a biological variable is by performing multiple determinations in individual subjects over time, calculating the different components of variance, and expressing the result as an index of individuality (II):

$$II = (CV_I^2 + CV_A^2)^{1/2}/CV_G \qquad (7)$$

where CV_I is the intraindividual coefficient of variation, CV_A is the analytical coefficient of variation, and CV_G is the within-group (subject-to-subject) coefficient of variation (31,48). Analytes with low indices demonstrate a high degree of individuality and vice versa. Harris (47) showed that when II is less than 0.6, the population-based RI will be insensitive to significant fluctuations within a subject. In contrast, when the index is more than 1.4, the population-based RI will be a more reliable reference (at least for a person whose mean and CV_I are about average). The within- and between-subject variations for many analytes are available in the literature (49,50 and 51), thus permitting the IIs for many of laboratory tests to be readily calculated. Index values for some common tests are presented in Table 4.7. One general conclusion from the literature on this subject is that IIs more than 1.4 tend to be the exception rather than the rule! This observation has several important implications for the field of laboratory medicine: the population-based RI is often a rather weak diagnostic tool, efforts to decrease CV_G (for example, by partitioning or exclusion) will improve the utility of the RIs that are derived, and the practice of using a patient's previous result (as determined in some defined state of health) as a RV should dramatically improve the diagnostic value of many laboratory tests.

Reference Change

The effective use of subject-based RVs requires the establishment of criteria for evaluating the significance of a difference between two successive test results. Empirical criteria for some routine laboratory tests have been established based on the results of surveys in which clinicians were asked to indicate how much a result would have to change before they would take action (52). Techniques for developing more objective statistical criteria have also been presented. One such criterion, the reference change

(RC), is defined as the difference between two successive measurements that would be statistically significant ($p < 0.05$) in 95% of the subjects in the reference sample (53). The development of a RC value requires the determination of the intraindividual variation under steady-state conditions (CV_I), the distribution of CV_I across the reference sample, and the incorporation of a model that describes the statistical relationship between serial measurements within an individual (30,52). If it is assumed that the intraindividual variation is constant across the reference population and that serial test results are statistically uncorrelated, the RC, i.e., the percentage of change that is considered to be significant, can be calculated as (51):

$$RC = 2.77\ (CV_A^2 + CV_I^2)^{1/2} \qquad (8)$$

This simplified model, which has been widely used, has been criticized as being too conservative, thus leading to a high proportion of positive findings that are neither medically nor clinically significant (30,51). Using what is claimed to be a more realistic model, RC can be calculated as:

$$RC = 2.77\ [\sigma_p^2\ (1 - r)]^{1/2} \qquad (9)$$

where σ_p is the 90th percentile of the true within-subject variance and r is the average serial correlation coefficient (54). A thorough discussion of the equations for calculating RC and the relative merits of the different models can be found in the references previously cited. An increase in the development and application of subject-based RVs is likely to occur in the field of laboratory medicine over the next decade. Such a change should greatly increase the value of laboratory test results to the process of medical decision making.

Clinical Decision Limits

These RVs consist of thresholds for an observed test result above (or below) which a clinician should (would) act to make some change in the management of a patient (55). These limits are developed by first defining the various reference populations relevant to the clinical question at hand. For example, in the simplest case we might consider two populations: one healthy and one affected with a particular disease. Next, the RVs for the populations are evaluated along with other relevant scientific or medical knowledge (i.e., the consequences or costs of misclassification, the prognostic implications of a high or low result, and the natural history of the pathologic process) to establish threshold values for the interpretation of the test result. The decision limits, which may be quite distinct from the reference limits for the test, can have either a broad or a narrow focus depending on the characteristics of the reference populations considered. Clinical decision limits are particularly useful for integrating laboratory test results into algorithms and critical pathways designed for the diagnosis and the management of different clinical problems. One well-known application of this type of RV is the establishment of plasma cholesterol concentrations of 200 and 240 mg/dl as clinical decision limits for the evaluation of hyper-cholesterolemia.

Presentation of RVs

Several different conventions for presenting quantitative results and their accompanying RVs have been proposed (32,34). The typical practice of presenting an observed result along with an upper and lower limit of the RI and flagging it as either high or low has been criticized for fostering the two-verdict (pathologic/nonpathologic) approach to interpretation and for being wasteful of the information contained in the reference distribution (34). Some laboratories have addressed this issue by reporting results in terms of standard deviation intervals of the mean of the reference distribution, multiples of the median, or as fractiles

TABLE 4.7. BIOLOGICAL VARIATION OF SELECTED ANALYTES IN HEALTHY POPULATIONS

		Biological Variability (%)		
	Analyte	CV_I	CV_G	Individuality Index
Serum				
	Sodium	1.07	0.54	5.9
	Potassium	1.2	1.6	2.0
	Calcium	2.4	3.8	3.9
	Osmolality	1.89	1.40	1.4
	Cholesterol	4.9	17.3	0.33
	Fasting glucose	4.75	6.08	5.2
	Creatine kinase	31.7	37.7	0.84
	Albumin	2.7	3.4	1.19+
	Haptoglobin	8.8	70.5	0.13+
	α_1-Antitrypsin	2.9	15.7	0.32+
	Complement C_3	3.8	19.7	0.34+
	Complement C_4	5.9	35.6	0.23+
	Immunoglobulin A	5.0	35.0	0.15+
	Immunoglobulin G	4.4	13.0	0.36+
	Immunoglobulin M	5.9	48.5	0.13+
	Kappa/lambda ratio	0.7	12.1	0.12+
	Rheumatoid factor	8.5	24.5	0.42+
	CA 125	36.0	59.3	0.60
	Carcinoembryonic antigen	9.1	55.6	0.17
	Prostate-specific antigen	18.1	72.4	0.25
	Thyroglobulin	4.4	12.6	0.84+
	Ferritin	17.8	13.5	1.34
	Lipoprotein(a)	8.6		0.15+
	Homocysteine	9.4	23.9	0.42+
	Vitamin E	11.9	19.9	0.59
	Cortisol	15.2	38.1	0.41
	Thyroxine	7.4	12.1	0.65
	Thyroid-stimulating hormone	19.3	19.7	0.99
Blood				
	Erythrocyte sedimentation rate	25		0.37+
	Reticulocyte Count	11	33	0.30+
	Red cell magnesium	18.1	20.3	0.9
Urine				
	Potassium (24 h)	25.8	23.7	1.09
	Creatinine (random)	41	13	3.16
	Calcium (24 h)	26.2	24.3	1.08
	Albumin (24 h)	36	58	0.62

Within-individual (CV_I) and within-group (CV_G) of variation are given as percentages. The values for individuality index (II) that are marked with + were calculated using the observed total analytical variability for the analyte in the laboratory where the study was performed. In all other cases, II was calculated assuming a value of 3.0% for CV_A. In practice, the actual value for II will be less than the value given when CV_A is >3% and greater than the value given when CV_A is 3%. II at any value of CV_A can be readily calculated using formula (7) as given in the text.
Data from refs. 48,49,50 and 51.

or percentiles of the reference distribution. The most common example of this type of presentation is the use of multiples of the median in reporting results for maternal serum α-fetoprotein, human chorionic gonadotropin, and unconjugated estriol. Although the rationale for these alternate methods of presentation is well founded, such formats are not widely used in clinical laboratories in the United States.

For RVs to accomplish their intended purpose, they must be reported to the ordering physician along with a patient's test results. Ideally, the laboratory report should contain all RVs pertaining to the interpretation of a particular result. This could entail the inclusion of an assortment of RVs for healthy and diseased populations, a patient's previous test result and the RC, appropriate clinical decision limits, and so forth. Because the laboratory is usually unaware of the precise reason that a test was ordered and of the presence of preanalytical variables that might be unique to a particular patient, this comprehensive approach would permit the clinician to choose the most appropriate RV as a standard comparison. Conversely, the laboratory must strike a balance between supplying the most useful RV(s) and obscuring the decision-making process with superfluous data. One solution to this dilemma would be to supply the single most relevant RV and to make available via a computer terminal or a workstation other RVs of potential interest to a physician or client. With the future prospect of a totally electronic medical record and the great rate of advancement in the field of medical informatics, it is not unreasonable to expect that the application of multiple RV systems to a single analytical result will soon become a routine practice.

CLINICAL PERFORMANCE CHARACTERISTICS

Part of "4 - The Interpretation of Laboratory Tests"

The ability of a laboratory test to distinguish among the various possible alternative states or conditions of a subject is one of its most basic performance characteristics. Indeed, this characteristic, which is called the clinical accuracy of the test, is the most important property with respect to a test's utility or value in the process of medical decision making (56,57).

The clinical laboratory of a large hospital is likely to perform or refer hundreds of different tests that have been ordered to evaluate or monitor a variety of different medical conditions and disease entities. These considerations raise a number of important questions. What methods are used to evaluate the clinical accuracy of laboratory tests? When multiple tests reflect the functional status of an organ system or the activity of a disease, which test will be the most informative? How can the results of laboratory tests be combined with information derived from the clinical evaluation and other diagnostic procedures? How can data from multiple sites be combined to derive the best available estimate of the clinical accuracy of a particular test?

Such questions are critical not only from the standpoint of promptly achieving the best outcome for the patient but also from the standpoint of efficiency and cost-effectiveness. Once established, clinical performance characteristics help clinicians to select the most effective tests and testing strategies and to translate laboratory results into a probability statement regarding the likelihood of a particular disease or outcome in their patient. Clinical performance characteristics help the laboratory director to focus the resources of the laboratory on tests with the highest clinical relevance. This section is intended to review the theory and methodology related to the characterization of the medical usefulness of laboratory tests.

Establishing Clinical Performance Characteristics

The data used to establish the clinical performance of a laboratory test are obtained during the phase II and phase III clinical trials that are required for a test to be approved by the FDA as a medical device. Studies of a similar design are also often performed in the field after a test has been released. In a phase II trial, test results from selected patients having the disease for which the test is indicated are compared with results from control subjects who are disease free. In a phase III trial, the clinical accuracy of the test is investigated in the target population in which the test is intended to be used in practice. Four crucial design elements are essential to the validity of the parameters derived from clinical trials (58). (i) The clinical status of the subjects in each group should be established using a reference (gold standard) method. (ii) The trial should evaluate a broad spectrum of healthy and diseased subjects, and the diseased group should include patients with a wide range of clinical presentations. (iii) CIs should be calculated for all the parameters that are estimated. (iv) When the diagnostic accuracy of a new test is being compared with that of a previously established test, a large sample of patients should be studied, the exact number being based on the sensitivity and specificity of the new test and the prevalence of the disease in the population in which the new test will be utilized. A more comprehensive discussion of the philosophical and statistical rationale pertaining to these design elements is presented elsewhere (56,58).

Sensitivity and Specificity

Two of the most widely used parameters for describing the clinical performance of a laboratory test are the clinical sensitivity (Se) and clinical specificity (Sp). (These parameters should not be confused with the *analytical* sensitivity and the *analytical* specificity of a test—both of which are properties of test methods.) Figure 4.8 shows the distribution of test results from a phase I clinical trial of a hypothetical test in samples from healthy and diseased subjects and several diagnostic thresholds or cutoff values (A–D) that might be used to discriminate between the two groups. The Se of the test is defined as the proportion of the diseased subjects that have a test result that is higher than a particular cutoff. The results for these subjects are interpreted as being positive for the presence of the disease. The Se is equivalent to the true positive (TP) rate for the test. The Sp of the test is, by definition, the proportion of disease-free subjects who have test results that are lower than a particular cutoff value. These subjects are considered negative for the presence of the disease. The Sp is equivalent to the true negative (TN) rate for the test. If there was no overlap between the distributions of results in the healthy and diseased groups, a cutoff value could be chosen to classify correctly all subjects as either healthy or diseased. The Se and Sp of the test at this particular cutoff would both be equal to 1.0 and

the clinical accuracy of the test would be perfect. In practice, however, the distributions of the test results in the healthy and diseased groups usually overlap. Thus, no matter where the cutoff value is set in the area of overlap, there will be false-positive (FP) and false-negative (FN) test results. FP results will occur in nondiseased subjects whose results exceed the cutoff, and FN results will occur in diseased subjects whose results are lower than the cutoff. It is possible, and often desirable, to use different cutoffs for a particular test in different clinical situations, but in all cases, cutoffs that improve Se will diminish Sp and vice versa.

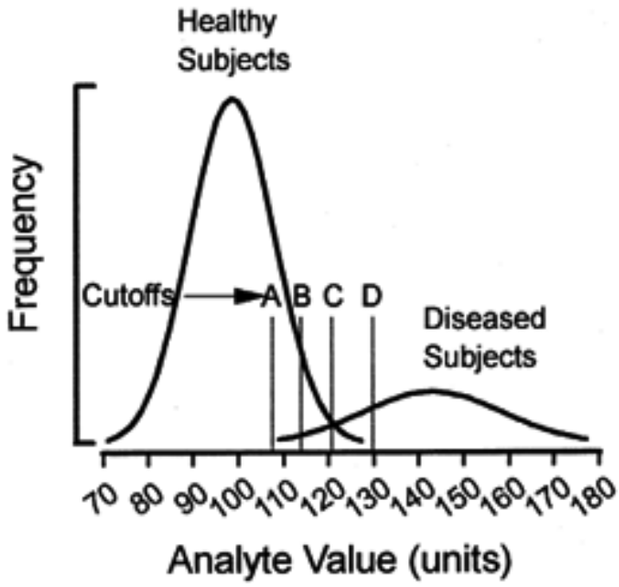


FIGURE 4.8. Frequency distributions of analyte values in healthy and diseased subjects. The values for the concentration of a hypothetical analyte were determined in a group of 200 healthy subjects and in a group of 50 diseased subjects. The raw data for each group were fitted to gaussian distributions. A through D represent possible cutoff values (or discrimination thresholds) that could be used to classify subjects based on their analyte values. See the text for details.

The calculation of Se and Sp with respect to any cutoff value is conveniently performed by arranging the test results for the healthy and diseased subjects in a two-by-two table based on the classification of each result as a TP, a FP, a TN, or a FN (Fig. 4.9) relative to the correct clinical state as determined by the gold standard method. Once the table is laid out, the parameters of interest can be quickly estimated using simple formulas.

Calculating Se and Sp

As an example of the calculation of Se and Sp, consider the hypothetical test used to determine the results presented in Fig. 4.8. The distributions of test results for 200 healthy subjects and 50 diseased subjects were obtained by gaussian fits of the original data. A tabulation of the original results relative to a cutoff value of 116 units is presented in a two-by-two format in Table 4.8. Using the formulas from Fig. 4.9, we calculate Se and Sp as:

Se = 48/(48 + 2) = 0.96

Sp = 194/(194 + 6) = 0.97

TABLE 4.8. A 2 × 2 TABULATION OF TEST RESULTS FOR 250 SUBJECTS

	Test Results		
Disease Status	**Positive**	**Negative**	**Totals**
Diseased	48	2	50
Healthy	6	194	200
Totals	54	196	250

Predictive Value and Bayes Theorem

Se and Sp are widely used for categorizing the clinical performance of a laboratory test. However, these parameters only express the probability that a diseased subject will have an abnormal result and the probability that a nondiseased patient will have a result within the RI. In practice, the clinician is usually most interested in evaluating the probability that a subject with a positive test result truly has the disease against the alternative hypothesis that the subject's test result is falsely positive (59). The proportion of total positive test results that are truly positive is called the predictive value of a positive test (PV+). Similarly, the proportion of total negative test results that are truly negative is the predictive value of a negative test (PV-). Again, calculation of these parameters can be quickly performed using the data in the two-by-two table along with the appropriate formulas. For the example presented above:

$$\begin{aligned} PV+ &= TP/(TP + FP) \\ &= 48/(48 + 6) = 0.88 \end{aligned} \qquad (10)$$

and

$$\begin{aligned} PV- &= TN/(TN + FN) \\ &= 194/(194 + 2) = 0.99 \end{aligned} \qquad (11)$$

It is important to remember that PVs determined in a clinical study are only accurate when the test is applied to populations with a disease prevalence that is similar to that of the study sample.

The dependence of PV+ on disease prevalence is expressed by Bayes theorem (60), which holds that:

$$\begin{aligned} PV+ = (Se)(prevalence)/[(Se)(prevalence) \\ + (1 - prevalence)(1 - Sp)] \end{aligned} \qquad (12)$$

When equation 12 is used to calculate PV+ in the above example in which the disease prevalence was 0.2 (i.e., 50 of the 250 subjects studied had the disease), the value of PV+ is identical to the value calculated using equation 10. However, equation 12 provides for the calculation of the expected value of PV+ at any disease prevalence. For example, if we were to use our test to screen for the presence of the disease in a population in which the prevalence was only 0.05, we can calculate that PV+ would decrease to 0.62. Thus, in this population, more than 40% of the positive results

obtained would be FPs. At a disease prevalence of 0.01, we calculate that 74% of the total positive results would be FPs. These exercises demonstrate why many a laboratory test that has looked promising when initially evaluated in a high-prevalence study population has later proved to be of limited value when applied to the general population with a lower disease prevalence.

Receiver Operating Characteristic Curves (61)

The receiver operating characteristic (ROC) curve is a graphical approach for determining the most efficient decision threshold or cutoff value for a laboratory test. The ROC curve can also be used to compare the relative clinical accuracies of different laboratory tests in a particular clinical setting. An ROC analysis is carried out by classifying the test result for each subject in a clinical trial as either positive or negative using a wide range of different cutoff values. The Se and Sp of the test at each cutoff are then calculated as described above, and the variables (1 - Sp) and Se are plotted on linear scales on the *x* and *y* axes, respectively.

The mythical “perfect” test in which Se and Sp are both 1.0 will graph as a rectangular hyperbola with its apex at the upper left-hand corner of the ROC plot (Fig. 4.10, left panel) (62). A test that does no better than a coin flip in differentiating healthy from diseased subjects will graph as a diagonal line with a slope of 1 and an intercept of zero (Fig. 4.10, right panel). Useful laboratory tests typically give plots that lie in between these two extremes (Fig. 4.10, center panel). In cases in which a single laboratory test is being evaluated, the single most efficient cutoff, i.e., the one that will maximize the number of correct classifications, is the one that plots nearest to the upper left corner of the curve.

DISEASE STATUS	TEST RESULT pos	neg	Total
DISEASED	TP	FN	TP+FN
HEALTHY	FP	TN	FP+TN
Total	TP+FP	FN+TN	TP+FP+FN+TN

FIGURE 4.9. A two-by-two table for calculating the sensitivity (Se) and specificity (Sp) of a laboratory test. A subject's test result is categorized into one of four groups based on whether it is above or below the cutoff value and the subject's true disease status as determined by independent criteria. Se and Sp are readily calculated using simple formulas. Sensitivity (Se) = TP/(TP + FN); Specificity (Sp) = TN/(TN + FP). TP, true positive test results; FP, false positive test results; TN, true negative test results; FN, false negative test results.

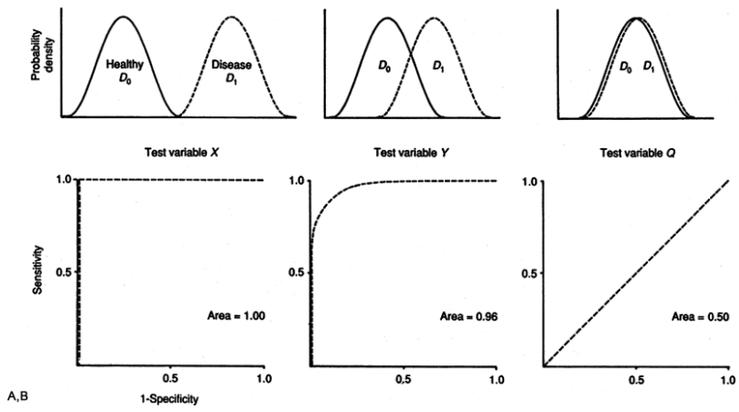


FIGURE 4.10. Receiver operating characteristic (ROC) curves for three tests of differing discriminatory power. The distribution of analyte values in healthy and diseased groups and the corresponding ROC curves are shown for a perfect test (i.e., with 100% clinical accuracy) **(a)**, a typical test that has adequate discriminatory power but will give some false positive and false negative results **(b)**, a test that does no better than chance at discriminating healthy and diseased subjects **(c)**. (Redrawn from Strike PW. *Measurement in laboratory medicine*. Oxford, UK: Butterworth-Heinemann, 1996.)

Constructing a ROC Plot

A ROC analysis of the test described in Fig. 4.8 was performed by selecting multiple cutoff points, calculating Se and Sp at each cutoff, and plotting (1 - Sp) versus Se as shown in Fig. 4.11. It is recommended that when possible, at least 10 cutoffs be evaluated in an analysis. These results show that the most efficient diagnostic threshold (i.e., the one that provides the most correct classifications of the study population) is 120 units. Keep in mind that, although this cutoff may be the most efficient overall, it might be desirable to use different cutoffs in different clinical situations. For example, if it was especially important that the test perform with 100% Se (and thus to identify correctly all diseased subjects) a cutoff value of 108 units would be the most effective. If achieving a Sp of 100% is crucial (thus avoiding the misclassification of a healthy subject), a cutoff value of 124 units might well be selected. When the relative performances of several tests are being compared, their ROC curves are plotted on the same axes (Fig. 4.11B) (58). The curve that most closely approaches the upper left corner of the plot is declared the more efficient test at the cutoff value corresponding to the upper left most point on its curve. This figure shows the comparison of serum parathyroid hormone (PTH), serum phosphate, and

serum chloride concentrations for the differentiation of patients with primary hyperthyroidism from control subjects. Serum PTH can be seen to have a higher Se and Sp than serum phosphate at all decision thresholds. Serum chloride has poor discriminatory power as demonstrated by the observation that the TP (Se) and FP rates (1 - Sp) are approximately equal at every decision threshold. Such a plot can be used to quickly determine the relative Se of the three tests at any desired level of Sp or their relative Sp at any desired level of Se.

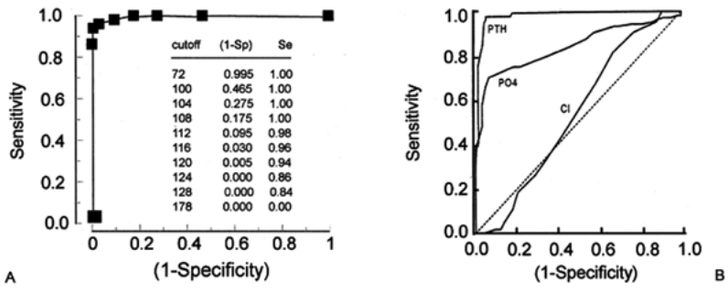


FIGURE 4.11. Construction and analysis of receiver operating characteristic (ROC) curves. **A:** A ROC curve for the hypothetical test shown in Fig. 4.8 was drawn by calculating sensitivity (Se) and specificity (1 - Sp) with respect to 10 cutoff values and plotting the results on a coordinate axis in which *X* = (1 - Sp) and *Y* = Se. **B:** A ROC analysis of the discriminatory power of serum parathyroid hormone, serum phosphorus, and serum chloride for the diagnosis of primary hyperparathyroidism. See the text for details. (Redrawn from Boyd JC. Mathematical tools for demonstrating the clinical usefulness of biochemical markers. *Scand J Clin Lab Invest* 1997;57 (suppl 277):46.)

The area under a ROC curve (AUC) can be used to express the overall clinical accuracy of a test as a single number (Fig. 4.10). By convention, the area is always in the range of 0.5 to 1.0, where 1.0 indicates that the test discriminates perfectly between the two groups, and a value of 0.5 indicates that the test does no better than a coin toss in discriminating one from the other. Unlike a point estimate, the AUC does not depend on a single region of the plot, nor does it indicate how well the test will discriminate at its single best cutoff. Rather, the AUC is a global measure of test accuracy. When the relative overall accuracies of multiple tests are compared by estimating the areas under their ROC curves, the test with the largest area is declared the most accurate. If the estimation of the areas is carried out using a computer program that calculates the CIs of the estimates, areas can be compared by statistical methods to determine the significance levels of any observed differences. For the comparison in Fig. 4.11B, the AUC for PTH, phosphate, and chloride were 0.983, 0.843, and 0.537, respectively, and the AUC for PTH was found to be significantly greater than that for the serum phosphorus test. The strength of AUC as an overall measure of performance can also be construed as a weakness because of the information that is lost when the curve is reduced to a single number. Two ROC curves with similar areas can have different shapes and if the curves cross one another, one of the tests will have a better Sp for a given Se, whereas the other will have a better Se at a given Sp (58). When the clinical accuracies of multiple tests are to be compared, it is always best if the clinical performance data for all the tests have been obtained from the same group of subjects. If this is not the case, then the statistical approach to data analysis will have to be adjusted to avoid bias in the results (58). Several commercial and public domain software programs for ROC analysis are available. These programs generally contain routines that will calculate Se and Sp, generate ROC plots, calculate AUCs, generate standard deviations and CIs for estimated parameters, and compare AUCs for multiple tests. A list of some of the more popular programs is presented in reference 61.

Summary ROC Analysis

In the emerging new field of evidence-based laboratory medicine, a composite version of the ROC curve called the summary ROC curve (sROC) is becoming more widely used for the assessment of diagnostic accuracy through metaanalysis (63,64). This approach involves extracting from the published literature clinical accuracy data from multiple independent studies and combining the results using mathematical models that account for study-to-study differences in diagnostic thresholds and sample sizes (65). By this technique, the Se and Sp determined for a laboratory test at multiple sites and involving a wide range of study subjects and clinical presentations can be combined. The composite sROC curve so derived represents, at least in theory, the best available estimate of the clinical accuracy of a diagnostic test.

Figure 4.12 shows a sROC curve for the metaanalysis of the Papanicolaou test based on the results of 62 studies involving a total of more than 20,000 subjects (65). The Se and Sp at the most effective discrimination threshold observed in each of the individual studies are plotted on the figure along with the sROC curve that was fitted from the combined data. Note that the estimates of Se for the individual studies ranged from 11% to 99% and the observed values for Sp ranged from 14% to 97%. This example shows the marked differences that can occur when clinical performance parameters are estimated at different sites and surely raises questions as to how much confidence should be placed in the results of any one published study. The sROC curve showed that the overall clinical accuracy of this test was not

high in studies performed before 1992 and that there was no cutoff point at which the test could simultaneously provide adequate Se and Sp. From the fitted curve, we can see that decision thresholds that provide a Sp in the 0.90 to 0.95 range (1 - Sp = 0.05 to 0.10) are associated with Se in the range of only 0.2 to 0.35. Metaanalysis is expected to have a significant positive impact on laboratory medicine by helping the field to establish reasonable expectations for clinical accuracy and to determine which tests provide the best information for clinical decision making.

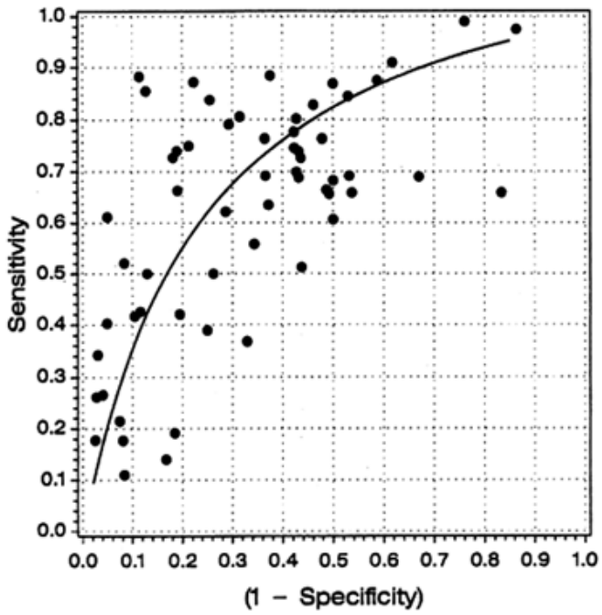


FIGURE 4.12. Estimation of clinical accuracy by metaanalysis. This summary receiver operating characteristic (sROC) curve for the Papanicolaou test is based on metaanalysis of 62 clinical studies. The sensitivity and (1 - specificity) values at the most effective discrimination threshold for each study are plotted as individual points. The combined data were used to derive a composite sROC curve. See text for details. (From Fahey MT, Irwig L, Macaskill P. Meta-analysis of pap test accuracy. *Am J Epidemiol* 1995;141:680, with permission.)

Likelihood Ratios

Another way of characterizing the clinical accuracy of laboratory tests is through the calculation of likelihood ratios (LHR) (58,59 and 60). A LHR expresses the probability that a particular outcome (e.g., test result, clinical finding) will occur in a diseased subject divided by the probability that the same outcome will occur in a nondiseased subject. For tests with only two outcomes or tests in which the results are interpreted using cutoff values that define two classes (positive and negative) of results, the LHR of a positive test (LHR+) is:

$$\begin{aligned} \text{LHR+} &= \text{Se}/(1 - \text{Sp}) \\ &= \text{TP rate}/\text{FP rate} \end{aligned} \qquad (13)$$

The LHR of a negative test (LHR-) is:

$$\begin{aligned} \text{LHR}- &= (1 - \text{Se})/\text{Sp} \\ &= \text{FN rate}/\text{TN rate} \end{aligned} \qquad (14)$$

For a continuous variable that is measured in two or more groups of subjects, the LHR+ at any cutoff value is equal to the ratio of the heights of each group's frequency distribution at the cutoff point. As is the case for Se and Sp, LHRs can be calculated for a test at a single or at multiple-decision thresholds or cutoffs. In addition, for continuous variables, LHRs can be established for different ranges of test results. In the latter case, the magnitude of a test result provides additional discriminatory power compared with the situation in which the result is merely categorized as positive or negative with regard to a cutoff.

In theory, LHRs for laboratory tests vary from 0 to infinity. A test with an LHR+ of 1 is considered to be of little or no value because in this case, an abnormal test result will be equally likely to be found in diseased and nondiseased subjects. Note that the LHR+ has the elements of a cost-benefit ratio in that the numerator in equation 13 (the TP rate) represents the benefit of testing and the denominator (the FP rate) represents the costs associated with testing (i.e., the FP results may lead to unnecessary further testing, inappropriate treatment, and anxiety among the subjects tested) (66). The discriminatory power of a test increases as its LHR+ increases above a value of 1.0. The LHR- expresses how many times less likely a normal test result is expected to occur in a diseased subject than in a healthy subject. The values of LHR- of clinical interest lie between 0 and 1, with the smallest values being associated with tests of the highest negative predictive value. The LHR- can also be interpreted in terms of cost and benefit with the numerator of equation 14 (the FN rate) representing the cost of testing and the denominator (the TN rate) representing the benefit of testing.

The greatest advantage of using LHRs to evaluate the clinical performance of laboratory tests is that when the probability of disease before testing is known (or can be estimated), the LHR+ and LHR- allow the direct estimation of the posttest probability of disease in view of the new information contributed by the test result (67). Furthermore, when serial testing strategies are used, the posttest probability after the first test is performed can be used as the pretest probability for the next test in the series. Thus, the overall impact of applying of a series of tests can be calculated. Finally, LHRs can be estimated for clinical signs and symptoms as well as for nonlaboratory diagnostic procedures (e.g., electrocardiographic findings, radiographic results), thus allowing the information from different modalities to be combined in serial diagnostic schemes or “critical pathways” for which overall predictive values can be calculated (68).

Working with LHRs

The most convenient way to work with LHRs is to convert probabilities of disease into the odds of disease and to use a simplified version of Bayes theorem:

$$\text{Pretest odds} \times \text{LHR} = \text{Post test odds} \qquad (15)$$

Odds and probabilities are interconverted using the following two formulas:

$$\text{Odds} = \frac{\text{Probability}}{(1 - \text{Probability})} \qquad (16)$$

$$\text{Probability} = \frac{\text{Odds}}{(1 + \text{Odds})} \qquad (17)$$

As an example of how to calculate and use LHRs, consider a patient who is being evaluated for the possible presence of a condition that has a prevalence of 20% in the population. The pretest odds of disease, as calculated from equation 16 are

$$\frac{(0.2)}{(1-0.2)} = \frac{0.2}{0.8} = 0.25$$

If during the physical examination the clinician were to identify a clinical sign or symptom that had previously been shown to have an LHR+ of 3.5 for the presence of the disease, the post-examination odds of disease would be raised to:

0.25 × 3.5 = 0.875

(disease probability = 0.875/1.875 or 0.47)

If this clinical assessment were then followed by the performance of a laboratory test with a LHR+ of 30, the posttest odds of disease in the presence of a positive test result would be:

0.875 × 30 = 26.25

(disease probability = 26.25/27.25 = 0.96)

Thus, the serial application of the clinical assessment and the laboratory test increased the probability of disease in this patient from 0.20 to 0.96. The LHRs for many clinical criteria, laboratory tests, and nonlaboratory diagnostic tests are now available in the literature. A sample of the range of LHR+ and LHR- values for selected clinical and laboratory evaluations is presented in Table 4.9 (59,67,68,69 and 70). Note that there are many cases in which clinical assessments have a discriminatory power that meets or exceeds that of a laboratory test that might be ordered to support a diagnosis. Also note that when multiple LHRs are established for different ranges of values for continuous variables (e.g., fasting plasma glucose), the magnitude of a disease-related change adds further discriminatory power to the test.

TABLE 4.9. LIKELIHOOD RATIOS FOR SELECTED LABORATORY AND CLINICAL EVALUATIONS

Disease	Evaluation	LHR+	LHR-
Sinusitis	Clinical	39	0.12
Deep vein thrombosis	Clinical	17	0.15
	Ultrasound	32	0.03
Streptococcal pharyngitis	Clinical	6.3	0.18
	Rapid antigen test	7.0	0.33
Bacterial infection, newborn	Band cell count > 1,000/mm^3	5.4	0.55
	White cell count ≥ 20,000 or < 5,000	1.8	0.81
	C-reactive protein > 10 mg/L	6.4	0.51
	Ear swab culture	7.5	0.33
Intermittent bacteremia	Blood culture, 1 set	129	0.26
	Blood culture, 2 sets	76	0.025
	Blood culture, 3 sets	53	0.015
Tuberculosis	Sputum smear	31	0.79
Myocardial infarction	Electrocardiogram	2.6	0.27
	CK-MB (single test)	2.8	0.75
	CK-MB (serial tests)	50.0	0
Cushing's syndrome	Central obesity	3.1	0.14
	Skull radiograph	10.6	0.49
	Plasma cortisol, 0800 h	2.5	0.16
	Plasma cortisol, 2400 h	24.0	0.04
Non-insulin-dependent diabetes mellitus	Fasting plasma glucose		
	>80 mg/dl	3.1	0.68
	>100 mg/dl	3.4	0.22
	>110 mg/dl	9.6	0.37
	>120 mg/dl	29.8	0.47
	>130 mg/dl	211	0.58
	>140 mg/dl	∞	0.69
Pheochromocytoma	Computed tomography scan	5.6	0.10
	Urinary catecholamines	16.4	0.19
Primary hypothyroidism	Serum thyroxine, decreased	6.0	0.12
	Thyroid-stimulating hormone, elevated	99	0.01
Primary hyperthyroidism	Serum thyroxine, increased	9.0	0.11
	Thyroid-stimulating hormone, suppressed	99.0	0.01
Ankylosing spondylitis	Clinical	6.3	0.18
	Human lymphocyte antigen B27 testing	11.5	0.09

Data from refs. 56,67,68,69 and 70.

One caveat that must be observed when combining LHRs for individual tests or criteria in a serial evaluation scheme is that each test or criterion in the series must be independent of all the others. In the absence of independence, the value calculated for the postseries probability of disease may be grossly erroneous. A second caveat is that LHRs, as estimates based on samples from a population, have CIs that indicate the reliability of the estimates. As was the case for reference limits, LHRs with narrow CIs are expected to give more accurate predictions of posttest probability than LHRs with wide CIs.

Nomogram for Estimating Posttest Probabilities

A posttest probability can also be directly determined from a pretest probability and a LHR using the nomogram (Fig. 4.13) originally published by Fagan (71). In addition to being convenient, this nomogram is instructive in that it demonstrates graphically the value of tests with very high or very low likelihood ratios for ruling in or excluding disease. For example, if a patient with a pretest probability of disease of 10% tests positive in a test with an LHR+ of 50, the posttest probability of disease is raised to 85%. If the same patient were to test negative on a test with an LHR- of 0.005, the posttest probability of disease would be decreased to less than 1%. Tests with LHRs near 1 are rarely effective in adding information to a clinical decision. This nomogram permits the potential information content of a test result to be rapidly evaluated.

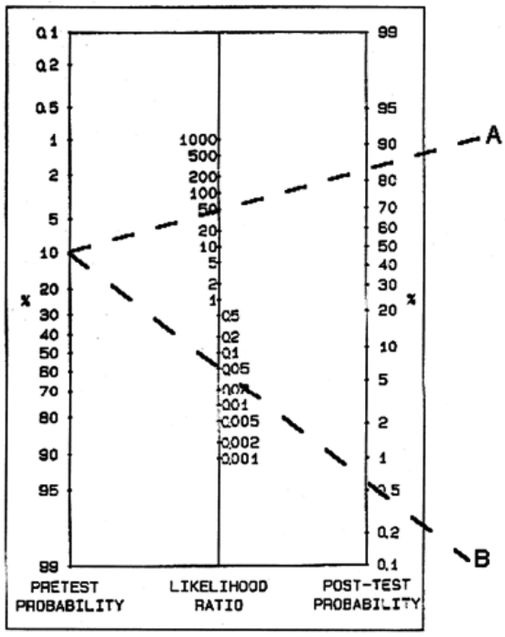


FIGURE 4.13. Nomogram for applying likelihood ratios (LHRs). This nomogram is used by drawing a line from the pretest probability on the left axis through the likelihood ratio of the test on the center axis and extending the line until it intersects the right axis. The posttest probability after application of the test is indicated at the point of intersection. The dashed lines show how positive findings for tests with LHRs of 50 (line A) or 0.005 (line B) markedly change the probability of disease from the pretest value of 10%. (Redrawn from Fagan TJ. Nomogram for Bayer theorem. *N Eng J Med* 1975;293:257.)

Estimating LHR from ROC Curves (72)

LHRs for a particular test can also be obtained directly from a ROC curve by evaluating the slopes of the curve in different regions (Fig. 4.14). If a test generates results on a continuous scale, then the slope of a tangent drawn to the ROC curve at any point is equal to the LHR for a single test result corresponding to that point on the curve. If a test generates dichotomous results (i.e., positive or negative relative to a cutoff), the slope of the curve between the origin and a point on the curve corresponding to any cutoff is equal to the LHR of a positive test when the results are classified as positive or negative according to that cutoff. If the test generates results that are classified into more than two intervals, the slope between the two points corresponding to the upper and lower bounds of the interval is equal to LHR for a test result that lies within the interval. These relationships provide a convenient means for the estimation of LHRs from ROC and sROC curves.

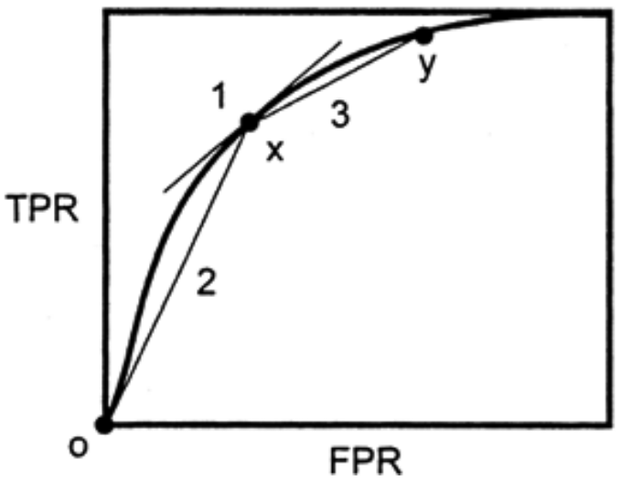


FIGURE 4.14. Estimating likelihood ratios from a receiver operating characteristic (ROC) curve. Three different slopes of an ROC curve: 1, the slope of a tangent at any point; 2, the slope of a line from the origin to any point on the curve and 3, the slope of a line connecting any two points on the curve correspond to different likelihood ratios for the test. See the text for details. (From Choi BCK. Slopes of a receiver operating characteristic curve and likelihood ratios for a diagnostic test. *Am J Epidemiol* 1998;148:1127, with permission.)

As stated previously, information obtained from the history and physical examination and from nonlaboratory diagnostic modalities is often sufficient to raise the pretest probability of disease to such a high level that additional testing may be superfluous. The clinician must ultimately decide when and when not to test based on the level of certainty that is required before applying various treatments or therapies. For example, a lesser degree of certainty might be acceptable in cases in which intervention is inexpensive and innocuous and the consequences of nonintervention are serious. Although such decisions are not always straightforward, careful consideration of the cost-benefit relationships using clinical performance characteristics can have a great impact on the economics of laboratory testing by minimizing the utilization of tests that do not make a significant contribution to diagnosis and management.

The three main topics of this chapter—biological variation, the establishment and use of RVs, and the clinical accuracy of laboratory tests—are highly interrelated, and the full appreciation of the significance of any test result requires a considerable understanding of the theoretical and practical aspects of each subject. As laboratorians, our efforts to minimize sources of error, variability, and misunderstanding in these three areas of practice will increase the value of laboratory services to our clinicians and their patients.

ACKNOWLEDGMENTS

Part of "4 - The Interpretation of Laboratory Tests"

The author thanks Ms. Debra Holmes and Ms. Decarlo Sykes for their help in preparing the manuscript. The author is especially

grateful to Edward W. Bermes, Jr., Ph.D., Professor Emeritus of Pathology, Loyola University of Chicago, for serving as a mentor, a colleague, and a friend throughout the author's career in clinical chemistry. The author also thanks his many colleagues in Pathology including the faculty, technologists, residents and administrators who were part of the dynamic and challenging practice environment from which this effort emerged.

REFERENCES

1. Watts NB. Clinical relevance of laboratory tests. *Arch Pathol Lab Med* 1988;112:379.

2. Doumas BT. The evolution and limitations of accuracy and precision standards. *Clin Chem Acta* 1997;260:145.

3. Witte DL. The complex connections between test properties and relevant outcomes: widening the perspective. *Clin Chem Acta* 1997;260:117.

4. Young DS. *Effects of preanalytical variables on clinical laboratory tests*, 2nd ed. Washington, DC: AACC Press, 1997.

5. Friedman RB, Young DS. *Effects of disease on clinical laboratory tests*, 3rd ed. Washington, DC: AACC Press, 1997.

6. Young DS, Bermes EW. *Specimen collection and processing: sources of biological variation*. In Burtis CA, Ashwood ER, eds. *Tietz textbook of clinical chemistry*, 3rd ed. Philadelphia: WB Saunders, 1999:42.

7. Young DS. *Effects of drugs on clinical laboratory tests*, 4th ed. Washington, DC: AACC Press, 1995.

8. Arendt J. Biological rhythms: the science of chronobiology. *J R Coll Phys Lond* 1998;32:27.

9. Scheving LE, Fevers R, Cope FD, et al. General principles of chronobiology. *Lab Med* 1994;25:306.

10. Ticher A, Sackett-Lundeen E, Ashkenazi IE, et al. Human circadian time structure in subjects of different gender and age. *Chronobiol Int* 1994;11:349.

11. Kanabrocki EL, Sothern RB, Scheving LE, et al. Reference values for circadian rhythms of 98 variables in clinically healthy men in the fifth decade of life. *Chronobiol Int* 1990;7:445.

12. Haus E. Biological rhythms in hematology. *Pathol Biol* 1996;44:618.

13. Griffin JE, Wilson JD. The testis. In: Bondy PK, Rosenberg LE, eds. *Metabolic control and disease*, 8th ed. Philadelphia: WB Saunders, 1980:1535.

14. Thorneycroft IH, Mishell DR, Stone SC, et al. The relation of serum 17-hydroxy progesterone and estradiol-17B levels during the human menstrual cycle. *Am J Obstet Gynecol* 1971;111:947.

15. Goldsmith B. Clinical chemistry of the newborn and infant. *Lab Med* 1997;28:659.

16. Meites S. *Pediatric clinical chemistry, reference (normal) values*, 3rd ed. Washington, DC: AACC Press, 1989.

17. Hicks JM, D'Angelo L. The laboratory's role in serving adolescent patients. *Lab Med* 1997;28:581.

18. Rifkind BM, Segal P. Lipid Research Clinics program reference values for hyperlipidemia and hypolipidemia. *JAMA* 1983;250:1869.

19. Faulkner WK, Meites S. *Geriatric clinical chemistry: reference values*. Washington, DC: AACC Press, 1993.

20. Tietz NW, Shuey DF, Wekstein DR. Laboratory values in fit aging individuals—sexagenarians through centenarians. *Clin Chem* 1992;38:1167.

21. Farkas DH. *DNA simplified II*. Washington DC: AACC Press, 1999.

22. Vestergaard EM, Ole Hein H, Meyer H, et al. Reference values and biological variation for tumor marker CA 19-9 in serum for different Lewis and secretor genotypes and evaluation of secretor and Lewis genotyping in a Caucasian population. *Clin Chem* 1999;45:54.

23. Lippi G, Guidi G. Standardization and clinical management of lipoprotein (a) measurements. *Clin Chem Lab Med* 1998;36:5.

24. Annesley T. Analytical test variables. In: McClatchey K, ed. *Clinical laboratory medicine*. Philadelphia: Williams & Wilkins, 19:77.

25. Ward KM. Recognizing effects of acute exercise on plasma and urine biochemistries. In: American Society of Clinical Pathologists Check Sample Clinical Chemistry no. CC 95-4. ASCP; 1995:47.

26. Irjala KM, Grönroos PE. Preanalytical and analytical factors affecting laboratory tests. *Ann Med* 1998;30:267.

27. Goldfein A, Monroe SE. Ovaries. In: Greenspan FS, ed. *Basic and clinical endocrinology*, 3rd ed. Norwalk, CT: Appleton & Lange, 1991:442.

28. Lockitch G. Clinical biochemistry of pregnancy. *Crit Rev Clin Lab Sci* 1997;34:67.

29. Holmes EW. The establishment and verification of reference values. In: Rose NR, Conway de Macario E, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:1158.

30. Harris EK, Boyd JC. *Statistical basis of reference values in laboratory medicine*. New York: Marcel Dekker, 1995.

31. Harris EK. Statistical aspects of reference values in clinical pathology. *Prog Clin Pathol* 1981;8:45.

32. National Committee for Clinical Laboratory Standards. *How to define and determine reference intervals in the clinical laboratory; approved guideline C28-A*. Villanova, PA: NCCLS, 1995.

33. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, eds. *Tietz textbook of clinical chemistry*, 2nd ed. Philadelphia: WB Saunders, 1995:454.

34. Solberg HE, Grasbeck R. Reference values. *Adv Clin Chem* 1989;27:1.

35. Young DS. Determination and validation of reference intervals. *Arch Pathol Lab Med* 1992;116:704.

36. Federal Register. Clinical Laboratory Improvement Amendments of 1988; Final Rule. *Fed Regist* 1992;57:7002.

37. College of American Pathologists. *Commission on Laboratory Accreditation Inspection Checklist. Laboratory General*. Northfield, IL: CAP, 1998.

38. Joint Commission on Accreditation of Healthcare Organizations. *1998-99 Comprehensive Accreditation Manual for Pathology and Clinical Laboratory Services*. Oakbrook Terrace, IL: JACHO, 1998.

39. Sunderman FW Jr. Current concepts of “normal values,” “reference values” and “discrimination values” in clinical chemistry. *Clin Chem* 1975;21:1873.

40. Solberg HE. Using a hospitalized population to establish reference intervals: pros and cons. *Clin Chem* 1994; 40:2205.

41. Linnet K. Two stage transformation systems for normalization of reference distributions evaluated. *Clin Chem* 1987;33:381.

42. Lott JA, Mitchel LC, Moeschberger ML, et al. Estimation of reference ranges: How many subjects are needed? *Clin Chem* 1992;38:648.

43. Frolich A, Nielsen BF, Conradsen K, et al. Do local hospitals meet the analytical goals for the use of common reference intervals? *Scand J Clin Lab Invest* 1994;54:169.

44. Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. *Clin Chem* 1990;36:265.

45. Holmes EW, Kahn SE, Molnar PA, et al. Verification of reference ranges by using a Monte Carlo sampling technique. *Clin Chem* 1994;40:2216.

46. Naus AJ, Borst A, Kuppens PS. Determination of n-dimensional reference ellipsoids using patient data. *J Clin Chem Clin Biochem* 1982;20:75.

47. Harris EK. Effects of intra-and interindividual variation on the appropriate use of normal ranges. *Clin Chem* 1974;20:1535.

48. Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* 1989;27:409.

49. Fraser CG. The application of theoretical goals based on biological variation data in proficiency testing. *Arch Pathol Lab Med* 1988;116:916.

50. Fraser CG. Biological variation in clinical chemistry, an update: collated data, 1988–1991. *Arch Pathol Lab Med* 1991;116:916.

51. Fraser CG. Data on biological variation: essential prerequisites for introducing new procedures? *Clin Chem* 1994;40:1671.

52. Elion-Gerritzen WE. Analytical precision in clinical chemistry and medical decisions. *Am J Clin Pathol* 1980;73:183.

53. Harris EK, Yasaka T. On the calculation of a reference change for comparing two consecutive measurements. *Clin Chem* 1983;29:25.

54. Queralto JM, Boyd JC, Harris EK. On the calculation of reference change values, with examples from a long term study. *Clin Chem* 1993;39:1398.

55. Statland BE, Winkel P. Reference values: are they useful? *Clin Lab Med* 1984;4:61.

56. Sox HC. The evaluation of diagnostic tests: principles, problems and new developments. *Annu Rev Med* 1996; 47:463.

57. Zweig MH. Evaluation of the clinical accuracy of laboratory tests. *Arch Pathol Lab Med* 1988;112:383.

58. Boyd JC. Mathematical tools for demonstrating the clinical usefulness of biochemical markers. *Scand J Clin Lab Invest* 1997;57(suppl 277):46.

59. Barry HC, Ebell MH. Test characteristics and decision rules. *Endocrinol Metab Clin North Am* 1997;26:45.

60. Shultz EK. Selection and interpretation of laboratory procedures. In: Burtis CA, Ashwood ER, eds. *Tietz textbook of clinical chemistry*, 3rd ed. Philadelphia: WB Saunders, 1999:310.

61. National Committee for Clinical Laboratory Standards. *Assessment of the clinical accuracy of laboratory tests using receiver operating characteristic (ROC) plots; approved guideline GP10-A*. Villanova, PA: NCCLS, 1995.

62. Strike PW. *Measurement in laboratory medicine*. Oxford, UK: Butterworth-Heinemann, 1996.

63. Moore RA. Concepts and principles of evidence-based laboratory medicine. *Am Clin Lab* 1999;Jan/Feb:24.

64. Irwig L, Tosteson ANA, Gatsonis C, et al. Guidelines for meta analyses evaluating diagnostic tests. *Ann Intern Med* 1994;120:667.

65. Fahey MT, Irwig L, Macaskill P. Meta-analysis of pap test accuracy. *Am J Epidemiol* 1995;141:680.

66. Dujardin B, Van den Ende J, VanGompel A, et al. Likelihood ratios: a real improvement for clinical decision making? *Eur J Epidemiol* 1994;10:29.

67. Anonymous. Interpretation of diagnostic data: 5. How to do it with simple maths. *Can Med Assoc J* 1983;129:947.

68. Halkan A, Reichman J, Schwaber M, et al. Likelihood ratios: getting diagnostic testing into perspective. *Q J Med* 1998;91:247.

69. Aronson MD, Bor DH. Blood cultures. *Ann Intern Med* 1987;106:246.

70. Fowlie PW, Schmidt B. Diagnostic tests for bacterial infection from birth to 90 days—a systematic review. *Arch Dis Child* 1998;78:F92.

71. Fagan TJ. Nomogram for Bayes theorem. *N Engl J Med* 1975;293:257.

72. Choi BCK. Slopes of a receiver operating characteristic curve and likelihood ratios for a diagnostic test. *Am J Epidemiol* 1998;148:1127.

5
Laboratory Information Systems

Kenneth D. Mcclatchey

Marisa Saint Martin

Information systems in pathology provide opportunities for pathologists and clinical laboratory scientists to have an impact both clinical care and modern research agendas. Not many technologies have had the implications and critical contributions to humanity as have the information technologies.

Radical changes in the health care system are creating a new role for the pathologist. This role has to do more with the management of information than with reading a slide and making the correct diagnosis. This is a new opportunity for pathologists that needs to be seen and captured in time. Business skills including marketing and management are becoming as important to success as clinical expertise. Information technology is essential as a link to all segments of integrated medical delivery systems: a tool to convert data into useful information and a source of knowledge for determining clinical efficacy versus cost. Clear communication between the pathologist and the clinician is vital for quality health care, and even though the primary mode of communication has been the pathology report, we may have to rethink the way we report in an attempt to provide the clinician (and possibly the patient) with integrated useful information for each particular patient instead of just data. The pathologist with interest in and appropriate knowledge of information systems is in a unique position to become an expert in information management.

HISTORY

Part of "5 - Laboratory Information Systems"

Digital computers became available for general use in the late 1950s. Early computers provided few user-oriented features and required considerable knowledge and skill to operate. They ran batch-oriented programs and supported single-tasking only. FORTRAN (formula-translation) and COBOL (common business-oriented language) were the only high-level languages available. Typically, mainframe computer systems were encased in rooms with glass walls and were tended by computer specialists who spoke their own language and were often far removed from normal hospital activities. This specialization and remoteness often resulted in the development of applications that were ill-suited to the workday departmental information processing needs of hospitals.

Early in the development of the field, most major computer manufacturers such as IBM, Burroughs, Honeywell, and NCR, understanding the potential for sales in the health care market, were active in their support of hospital information systems. Corporations such as Lockheed and McDonnell-Douglas, which were experienced in using computers to manage complex management tasks, also decided to offer health care computing as business ventures. For example, Lockheed supported the early development of the Technicon Hospital Information System.

The Technicon system, begun by Lockheed in the 1960s, installed at the El Camino Hospital in Mountain View, CA, became the best known and most successful patient management system of its time. The ultimate success of the system at El Camino led to the spread of this and other systems to other hospitals. The Technicon system is still being marketed today, which is unusual because many of the early software and hardware vendors have gone out of business or are no longer active in the hospital market.

In 1987, Hammond (1) listed five factors that have driven the development of hospital information systems:

1. Hardware and software technological factors: During the development period, computers evolved from single-tasking “untouchable” and unfriendly mainframe computers to highly interactive and multiuser minicomputers
2. Increasing technological expertise on the part of system developers and users
3. Economic factors: As the cost of delivering health care increased, computers seemed to offer at least the opportunity to operate in a more efficient manner
4. An increase in the amount of laboratory data generated, which necessitated more automated procedures
5. Increased influence of external organizations: Information requirements of federal and state governments and third-party payers forced hospitals to turn to automated systems for financial and patient management applications

PERSONAL COMPUTERS: HOW DO THEY WORK?

Part of "5 - Laboratory Information Systems"

A typical personal computer will have several basic components. Whatever the machine, it will have a keyboard or any other means of entering information (input device), a video monitor, printer or other means of getting information back out to the user (output device), and a method of making permanent

records of running additional software (disk drive). All these components are arranged in the system board (Fig. 5.1). The system board includes the central processing unit (CPU), which is the microprocessor that directs all the computer activities; an auxiliary CPU, a crystal quartz clock that coordinates the responses of the computer's many circuits, the ports for connecting input and output attachments, and two very important microchips for internal memory: the ROM and the RAM (Fig. 5.2).

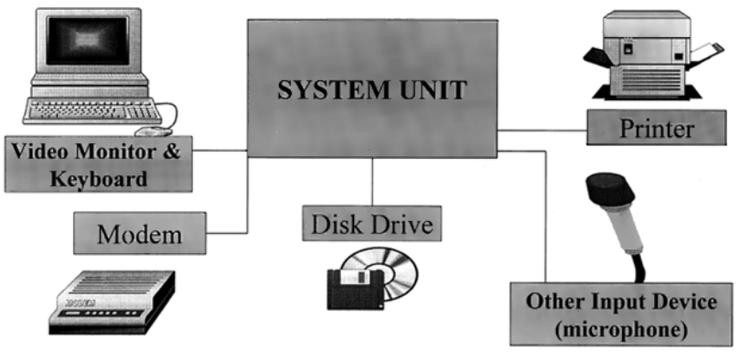


FIGURE 5.1. Components of a personal computer.

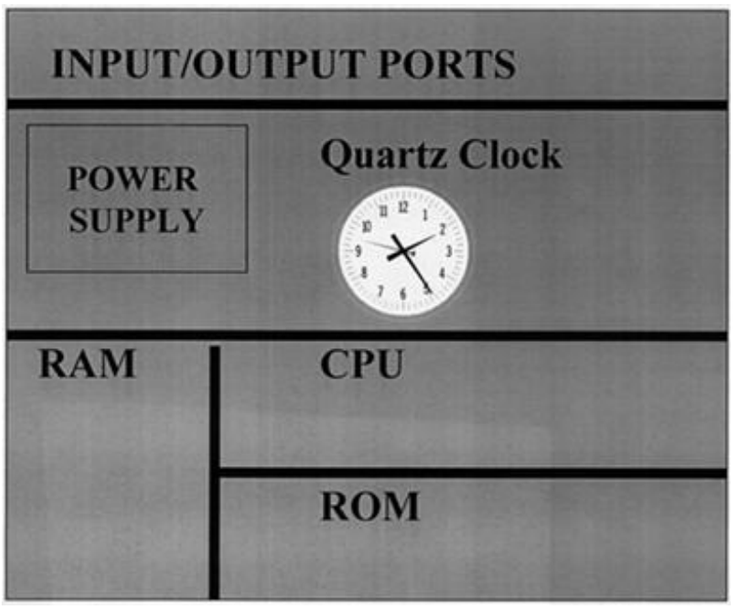


FIGURE 5.2. Components of a computer.

The ROM (read only memory) holds instructions that cannot be altered. The RAM (read and write memory or random access memory) stores programs and information only when the computer is operating. It can be changed by the user.

Computers only read or understand binary digits. Binary digits have only two components or possible states, such as 1 and 0 or true and false. Computers use a combination of 1s and 0s to create numbers and words. A bit (short for binary digit) is the smallest unit of information. A nibble is four bits, and a byte is a sequence of eight bits, treated as a storage unit. For example,

bit: 0 or 1
nibble: 0010
byte: 00100010 → this sequence equals the number 34

A megabyte (MB) is the hard disk storage capacity and represents 1,000,000 bytes. A gigabyte (GB) is 1,000 megabytes and a terabyte (TB), 10,000 megabytes.

The executed instructions of a computer are in the form of high-level machine languages such as Lobol (used for numbers), Pascal, Fortran, Mumps (for medical terms), Basic, and Java (for networks).

Computers today usually handle memory using a 32- to 64-bit (or better) architecture; they have 90 to 550 MHz (speed of replacing one image with the next one), and 28 dot per inch screen (defines sharpness). The chip in the motherboard for speed of the run of the computer has several upgrades from 286 to 386 to 486 to Pentium, Pentium II, and Pentium III.

A modem is a device that allows connection of the personal computer through telephone or fiber optic lines. It can be internal or external. The modem speed unit is called the baud.

In general, the hardware constitutes the physical parts of a computer, and software represents all the instructions that tell the computer what to do, how to act, and so on. Computer programs are examples of software. Program files are the commercial ones that can be purchased; data files are the ones the operator creates. The operating system is a combination of programs that coordinates the computer duties and manages the storage of other programs and other information. It acts as the “stage manager.” The application programs are the “stars” (e.g., word processor, games). Translator programs convert statements of higher level languages into machine language. Utilities programs perform sorting/merging of files. When a computer is turned on, a series of self-checks of internal components and functions takes place so rapidly as to seem almost simultaneous. Only seconds after, a message appears on the monitor to indicate that the computer is set for action.

DIVISION OF PATHOLOGY INFORMATICS

Part of "5 - Laboratory Information Systems"

Seventy to 80% of important medical decisions involves a laboratory test, and the formation of a pathology informatics division is imperative in this new era of health care delivery. The division would be responsible for improvement and growth of the laboratory and anatomic pathology and its interconnection with other hospital information systems. It should have an administrative component and personal computer support and should also be responsible for the academic computing of the department (Fig. 5.3). An intense collaborative process must exist between the hospital information system (HIS) and the laboratory

information system (LIS). The LIS relates to the HIS, the billing system, ancillary reporting and record keeping, and electronic record and custom reporting. When an order is placed on the floor (HIS), the information is transferred to the LIS. The LIS prints the labels and the labeled specimen is delivered to the laboratory. After the results are obtained, the information is entered into the LIS and the data are sent to HIS upon inquiry. The reports are prepared on a LIS (Fig. 5.4).

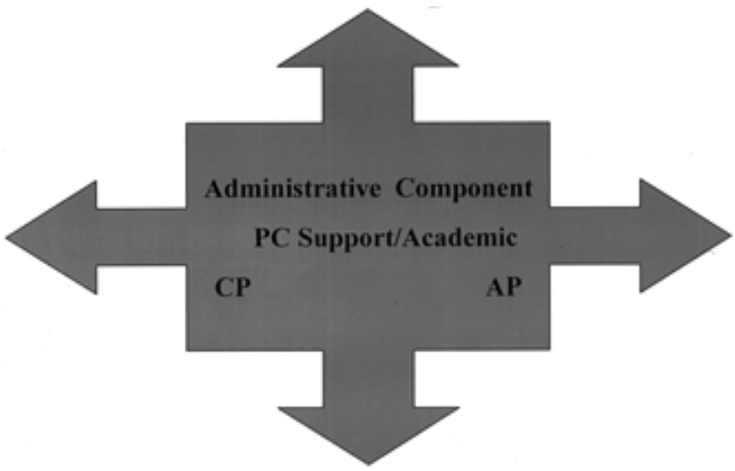


FIGURE 5.3. Division of pathology informatics.

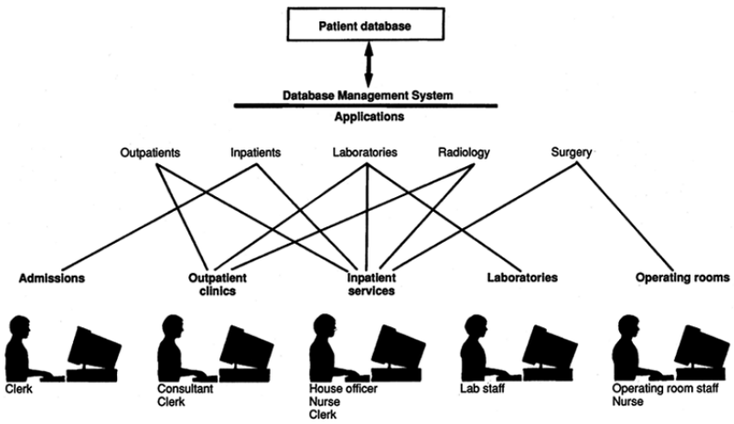


FIGURE 5.4. Applications sharing an integrated database.

Several types of interfaces are required for this process to run: ADT (admission/discharge/transfer), orders, results, billing, reference labs. HL7, (a protocol for formatting and transmitting data in the health care environment) is an example of commonly used interfaces. Transferring of the data is only one of the functions of the interfaces. Others include code translation, conversion of protocol and date, and error handling.

ANATOMIC AND CLINICAL PATHOLOGY: THE NEW ERA

Part of "5 - Laboratory Information Systems"

Indisputably, with the many economic forces influencing the medical field, the demands on the anatomic and clinical pathology laboratories continue to escalate with the increasing volume of specimens and expectations for a shorter turnaround time. Not only the laboratory, but radiology, radiation oncology, nuclear medicine, pharmacy, and home health care, among others, are part of the scope of the clinical, information systems. This clinical integration of information is developing new career opportunities for many people. Seminars, conferences, training programs for residents, students, pathologists, and fellowships in informatics are only the beginning of this fast-growing segment of medicine, which is becoming a new medical specialty.

THE LABORATORY INFORMATION SYSTEM

Part of "5 - Laboratory Information Systems"

The LIS provides different services for the laboratory itself as well as for the clinical services. Order and result entry, inquiry, inventory control, and time management are the most common laboratory services obtained from the LIS. For the clinical services, reporting and patient institutional billing are the most used services. The LIS is also extended into decision support through internal or external rule engines, hospital mergers, and outreach

to remote laboratories or physician practices. The pathologist as an information manager has become a vital link between the patient, the clinician, researchers, and other pathologists (Fig. 5.5). Considering that approximately 70% to 80% of the information regarding patient care will be generated in pathology laboratories, we have an obligation to communicate that information in a clear, timely, available, and complete manner.

FIGURE 5.5. Anatomic pathology impact. Making the diagnosis is only half of the struggle. We are only as good as the clarity, timeliness, availability, and completeness of our reports.

Reporting

Pathologist–clinician communication is vital; however, it is difficult to find the time to discuss across a microscope or a conference table. Providing a clear and complete report with integrated information including all significant diagnostic and prognostic tests will improve the clinical service that patients receive (Fig. 5.6). There is an enormous educational value in accurate/educational reporting, not only for the clinicians but also for the interested patient, who may already come to see his or her doctor with pages of information obtained from the Internet.

Integrated report example

Patient demographics	Case worker
Gross pictures	Gross description
Microscopic picture	Microscopic description or comment
Molecular or Flow cytometry studies and special procedures.	
Final Diagnosis (with integrated comment)	

FIGURE 5.6. Integrated report example.

Work Flow

With a good computer system, the work flow in the pathology lab can be improved to the limit of the imagination of the medical director or the information division chief and the system capabilities. Cytology quality assurance, tissue coding, online stain ordering, resident signout/review, and electronic sign out can all be done faster, in a more reliable manner with a state-of-the-art anatomic pathology system, for example.

Productivity

Productivity is enhanced by the use of e-mail, personal computers allowing multiple functions, transcription reengineering, training of residents/fellows, and studying material for examinations, marketing, information. Informatics can be the gateway for pathology to develop links to the community we serve.

System Selection: Purchase of a LIS

For those clinical laboratories that are already computerized, the rationale for upgrading an existing LIS is relatively simple: it would be unthinkable to return to a manual information processing system. The need to develop a rationale for purchasing a system is much more appropriate for laboratories not currently running a LIS. Despite the usual visceral reaction that modern organizations ought to be computerized, it is a useful exercise to list formally the advantages and disadvantages of a LIS. The development of a list may avoid the naive expectation that a LIS will solve all the problems of the laboratory.

Strengths

- Ability to perform the same task as many times as needed in the same way
- Speed
- Perfect memory

Weaknesses

- Stupidity (computers can do only what they are programmed to do)
- Inability to make inference
- Inability to deal with exceptions
- Inability to deal with broad constructs

Benefits with Proper Human–Machine Interface

- Laboratory professionals acquire “perfect memory”
- The computer’s ability to perform repetitive tasks relieves the laboratory staff of onerous aspects of work
- Laboratory professionals deal with exceptions
- Laboratory professionals increase their capabilities and shift their emphasis toward the tasks for which they are trained

The realization of these benefits assumes that the human–machine interface is properly designed. The basic thrust of Korpman's (2,3) reasoning is that a LIS will shift much of the drudgery of routine repetitive rote tasks to the computer, allowing laboratory personnel to exercise their intellect by dealing primarily with exceptions. This is a very praiseworthy objective.

Although Korpman is undoubtedly correct in his analysis, the pathologist seeking to justify the purchase of a LIS to hospital administrators may sometimes require a more tangible, exact list of reasons for the purchase of such an expensive item.

A number of data-handling activities must be performed manually if there is no LIS support for the clinical laboratory activities. They include the following:

1. Creation of phlebotomy drawing lists
2. Distribution of completed manual test requisitions to the individual laboratories
3. Creation of intralaboratory worksheets
4. Distribution of hard copy test results to patient care units
5. Individual response by laboratory to telephone queries for test results

To develop a cost-benefit analysis for a LIS to total costs of personnel to perform these and other functions must then be compared with the purchase and support cost of LIS hardware and software, as well as the personnel costs for the direct support of the computer. It is important not to overlook that some of the personnel costs associated with a LIS are related to individuals not directly assigned to the LIS unit, such as technologists in other clinical laboratories.

Hendricks and Langhofer (4) in 1982 evaluated the impact of the acquisition of a LIS on a community hospital. The acquisition produced a decrease in inflation-adjusted labor costs per unit of work, a decrease in inflation adjusted direct cost per unit of work, a decline in the ratio of clerical-to-technical labor force, and an increase in overall laboratory productivity. Comparison of the average College of American Pathologists' (5) *Manual for Laboratory Workload Recording Method* units per laboratory full-time equivalent for the 3 years before and 4 years after the LIS installation showed a 12% increase in productivity compared with a 5% increase in automated tests. The authors concluded that productivity in their laboratory rose by approximately 7% (overall productivity increase minus the increase in automation) as the result of computerization.

Hendricks and Langhofer (5) stress that there are several nonquantifiable effects of computerization that must be factored into any cost-benefit analysis. An example is the development of *ad hoc* reports by extraction of information from the laboratory database for quality assurance, utilization review, or risk management. With manual systems, such reports are the result of many man-hours of work.

Justifying the cost of a LIS is most interesting from the perspective of increasing the efficiency of physicians and providing a better patient outcome. Few would deny that the creating of a hard copy interim test report, organized by laboratory with a moving 4-day window of current and previous results, makes physicians more efficient. They are also made more efficient by remote stat delivery of test results to critical care units immediately after test verification. Access to a long-term on-line archive of test results becomes critical when handling and transporting the paper record is difficult.

Ultimately, the necessity to computerize the clinical laboratories in a hospital is driven by two factors: (a) high labor costs in a hospital makes the automation of manual processes attractive from a financial point of view and (b) reports and quality assurance activities, which simply cannot be provided by manual systems, are being demanded with increasing frequency in all hospitals regardless of the resource implications.

There are 13 major steps in the selection and installation of a LIS:

1. Creating a LIS project team
2. Gathering information
3. Reviewing and analyzing the information collected
4. Developing and distributing a request for a proposal
5. Evaluation vendor responses
6. Selecting a vendor
7. Negotiating with a vendor and developing a contract
8. Training and site preparation
9. Installing the hardware
10. Making parallel runs
11. Going live
12. Fully operating
13. Evaluating and monitoring

THE LIS PROJECT TEAM

Part of "5 - Laboratory Information Systems"

Once the decision has been made to purchase a LIS or to upgrade an existing system, it is necessary to assign a task force of hospital personnel to coordinate the project from its inception until the system is turned over to personnel in the LIS unit after the go-live date. The working core of the project team *must be* clinical laboratory personnel. They are the natural leaders of the project team because a successful LIS implementation is critical for them. Conversely, the participation of personnel from outside the clinical laboratories can also be critical to the success of the project, and their participation should be encouraged. For example, a clinician who can speak to the needs of the physicians in terms of test reporting formats and locations of terminals for data retrieval should be represented on the project team. Gunton describes the design process using a diagram. He explains the initial process as trying to close the information gap. Ultimately, maintenance programmers and users will attempt to close the maintenance gap throughout the life of the system (Fig. 5.7).

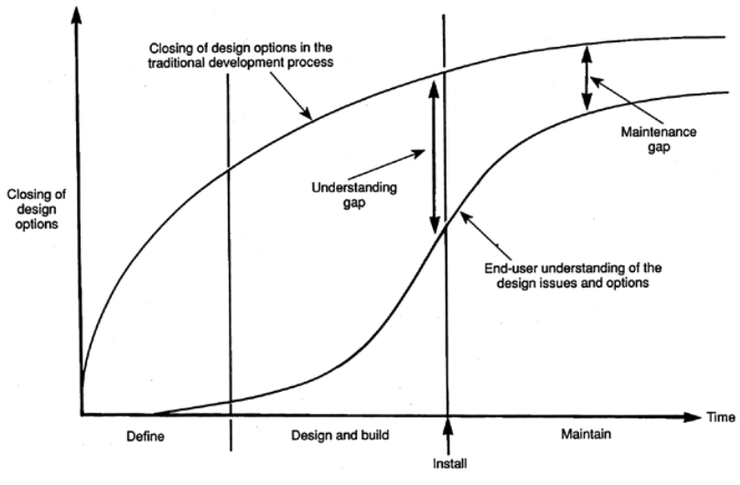


FIGURE 5.7. Laboratory information system design process and end-user learning curve.

INFORMATION GATHERING

Part of "5 - Laboratory Information Systems"

There are four general sources of information about laboratory information systems: causal conversation among professional colleagues, written publications by professional societies, vendor-generated print materials, and vendor displays at national and regional meetings. Even with written material from multiple sources at hand, it may be difficult to isolate the essential pieces of information (e.g., what language is the system written in, are all laboratory modules up and running, how many sites were installed in the past calendar year?).

In most instances, it is possible to reduce the large number of possible LIS vendors to a half dozen or fewer on the basis of the initial needs analysis of the clinical laboratories and information about the various systems gleaned from both informal discussions and formal sources. Generally, hospital personnel should not feel the need for hands-on demonstrations of systems until they have pared their list down to a small number of vendors. Under such conditions, the competing vendors will be delighted to visit the hospital and set up a formal demonstration of the system under controlled conditions.

DEVELOPING A REQUEST FOR A PROPOSAL

Part of "5 - Laboratory Information Systems"

A request for proposal (RFP) is a document developed by a LIS project team in which the information processing needs of the clinical laboratories and the hospital are delineated. This document is distributed to those LIS vendors with a reasonable chance of being successful bidders in the competition, some of whom will choose to issue a written response to the request for proposal (RFP). A successful vendor will then be chosen from the responders.

The key elements of a RFP are

- Hospital and laboratory profile
- Functional system requirements
- Technical systems requirements
- Training requirements
- Implementation requirements
- Financial considerations
- Vendor profile
- Conditions of bidding
- Evaluation criteria
- Decision timetable

The key goals during the development of a request for proposal are to:

1. Gather information about all potential LIS vendors and then develop a basis for comparison among those responding
2. Analyze current laboratory operations to assist in vendor selection and to ensure successful system installation
3. Create a long-term strategic plan for the automation of the clinical laboratories, including a prioritization among all available computer applications if funds are limited
4. Develop a justification for the purchase and configuration of the LIS to ensure current and continuing resource support
5. Build a consensus among the members of the LIS project team
6. Develop a formal plan for staffing the LIS unit and daily operations after implementation

Analysis of Current Laboratory Operations

An important first step in the development of an RFP is a systems analysis of the various clinical laboratories. There are three primary objectives for a systems analysis in the clinical laboratories as part of the RFP process: (a) to evaluate the information-handling

systems currently operating within the laboratories, (b) to document specific responsibility for “loop closure” in laboratory testing requests and recording procedures, and (c) to analyze the laboratories in terms of input, output, processing, and control. Loop closure means addressing the fundamental cause of a recurring problem rather than continuing to treat it symptomatically.

Krieg and colleagues (6) use several techniques for performing systems analysis including logic flow charts, data flow diagrams, and wall chart displays. Logic flow charts, also called system flow charts, are the most useful for planning and documenting laboratory operations, whereas data flow diagrams are especially useful for planning and documenting computer programs. The wall chart display is used to visualize forms and documents used in a particular data-processing system.

The reason for proceeding with the systems analysis step is to take a hard look at current laboratory operations and information flow. An attempt should be made to correct any problems encountered before the installation of a LIS.

The installation of a LIS often requires that certain operational changes be made in laboratory operations. Although not always the case, such changes frequently serve to enhance the efficiency of information processing within the laboratories. Veteran laboratory personnel may balk at the implementation of changes mandated by a LIS installation, preferring to continue with the “tried and true.” Such employees may be uncomfortable with change or may have instituted procedures scheduled to be discarded and therefore may have a vested interest in their continuation. Friedman (7) calls the persistence to use familiar procedures that appear to be more efficient than they actually are the “competency trap.”

Having complete a systems analysis, it becomes much easier to make informed judgments about the efficiency of current laboratory operations. Systems analysis will ultimately provide an inside look into current operations and suggest ways in which the LIS can be customized to meet the information needs of local users.

A fairly typical question that arises during the time that (a) a systems analysis is being performed on laboratory operations and (b) vendor offerings are being reviewed in a preliminary way is whether all laboratory modules should be brought up simultaneously on the “go-live” day.

The classic process has been to bring up the “numerical production” laboratories such as chemical pathology and hematology at one time, and then other laboratories with more complex applications such as microbiology and blood bank at a later time. With the sophistication of today's systems, the strategy is “all laboratories go live at the same time.” Such a method hastens the return to normality from what is perceived to be an awkward transition state.

Purchasing a LIS with Sufficient Capacity

The proper sizing of a LIS for a hospital is a complex issue with major implications for the hospital. It is important to have a sufficiently powerful central processor and disk storage with the response time of the system and the number of months of test results on-line being satisfactory to users.

To design a system of sufficient size, vendor personnel start with hospital volume statistics provided in the RFP including the (a) current laboratory workload, (b) the number of yearly inpatient discharges, (c) the number of yearly ambulatory care and emergency department visits, (d) workload measurement, and (e) the anticipated growth in laboratory activity.

Using these data plus the anticipated software configuration of the LIS, including the number of applications running and the requirements for on-line order entry and test results, the vendor will make recommendations about system size. The cost of the hardware will then be rolled up in the quotation for the system.

Vendors know that they are competing with other vendors on the basis of cost as well as system functionality. On the one hand, they will design a system with adequate capacity for near-term requirements. On the other, they will not build in excessive system capacity for the long term. LIS performance for the future will depend on the increase in the laboratory workload and increasing use of the system. Moreover, the vendor will undoubtedly release software upgrades with new features that may increase the system functionality but will also require more computer power.

In the final analysis, the sizing of a LIS depends on the following:

1. The recommendation of the vendor in the RFP
2. The hardware configuration of comparable hospitals in terms of complexity and workload, which will be analyzed during site visits
3. Some “fudge factor,” taking into consideration local hospital resources and the interest and willingness of the hospital to invest in a state-of-the-art system

Whenever possible, it is wise to buy more system capacity than will be necessary to run the LIS software on day 1 and then grow into the system.

Several questions pertaining to hardware upgrades should be addressed after vendors submit recommendations for a particular system configuration. For example, what is the relationship of the recommended system capacity to the maximum capacity of that configuration? What are the increments of growth when the recommended hardware configuration is deemed inadequate? What is the recommended hardware (and software) upgrade path from the recommended configuration, and is it smooth or is it complicated?

VENDOR SELECTION

Part of "5 - Laboratory Information Systems"

Extreme care should always be taken when selecting a vendor. The importance of this decision rests in part with the large amount of money required for the initial system purchase, as well as for continuing software and hardware support. Equally important is the fact that the installation of a LIS requires a major investment in training. Having made such an investment, laboratory professionals are generally loath to switch to another vendor at some later point. The status quo is often the path of least resistance, even when the current vendor may be providing an inferior product and service. LIS vendors understand that the

switching costs for their current clients are high and so will try to maximize them to the extent feasible. For example, one major issue in switching from one LIS vendor to another is file conversion from the old to the new system, particularly for anatomic pathology and blood bank files. The LIS vendor being abandoned will do little to facilitate this process. The new LIS vendor may also be wary of actively participating in the file conversion for fear of charges of copyright infringement by the previous vendor. The scenario goes on and on.

Criteria for Vendor Selection

Some general criteria to use in the search process are

1. Is the system designed to meet the needs in multiple settings, such as a multihospital group with laboratories in geographically distributed sites?
2. Is the company adequately capitalized and financially healthy?
3. Is the system well documented, well designed, and written in a language such that software support could be obtained in the event of vendor failure?
4. Will the vendor place and maintain, on site or in escrow, a copy of the source code and complete documentation?
5. Is the vendor investing sufficient resources in product development, as judged by the number of assigned personnel in the area and the frequency of major new product announcements?
6. Are contracted or purchased software modules clearly identified with their origin acknowledged, and do they integrate well into the total system?
7. Has the development cycle for the entire system been completed, and has it been running in test or prototype sites for at least a year?
8. Are all planned interfaces (instrument, hospital mainframe computer) running successfully at other sites?
9. Is the system, with all the desired laboratory modules and applications, running successfully in another site of equal complexity?
10. Has the vendor shown continued and steady growth, as measured by the number of clients installed per year, or is the number of installations declining, indicating decreasing acceptance in the market?
11. Does the vendor's support and training staff meet local needs, and does the staff appear to be adequate for anticipated installations?
12. Does the hardware vendor have a major position in the market and a close collaborative relationship with the LIS vendor?

The installation of a LIS commits the clinical laboratories and the LIS vendor to a long-term and close relationship. It is extremely important that the “chemistry” is right between the two organizations to foster a positive and productive relationship.

Connelly et al. (8) provide an informative discussion of both formal and informal methods for sorting through the enormous amount of data about vendors generated through the RFP process. The subjective viewpoint holds that the large number of qualitative factors inherent in the LIS selection process made rigid and elaborate procedures for quantifying differences between favored vendor impractical. Such thinking favors either the *educated guess method*, wherein in overall impression of the system is garnered by the decision makers, the *cost only method*, or the *easy way out method*, whereby most prominent vendor or the current vendor is selected.

According to Connelly and colleagues (8), there are two formal or structured approaches to vendor selection. In the first, the *elimination by aspect method*, key features that serve as mandatory requirements are identified. Systems not meeting these mandatory requirements are dropped from consideration. The approach that Connelly and colleagues favor is the *multiattribute utility model (weighted scoring model)*. This approach consists of nothing more than developing a set of major and minor attributes for the desired system and values for each. Evaluators from the LIS project team score each minor attribute for each vendor. Some simple calculations provide a performance score for each vendor. The leader of the LIS project team should review the scores of the anonymous evaluators carefully to make sure that no one is “gaming” the process by assigning highest scores to the favored vendor and zeros to all others. This could seriously skew the mean scores for the whole set of evaluators.

The Site Visit in Vendor Selection

The site visit is often omitted or given short shrift, perhaps because of the expense, coordination, and effort in transporting a large number of personnel to multiple hospitals. For either the first-time buyer or the department-switching vendors, it is necessary to view LISs in a production environment before final selection. Be sure to choose hospitals for site visits of equal size and complexity to the home institution, even though this may increase the travel costs. Vendors have a small cadre of “show” sites for visits where personnel are generally happy with the product. It may be worthwhile to attempt to schedule through the vendor a visit to a hospital other than a standard show site. Not all clients of a particular vendor will be equally satisfied with the product, so do not be surprised if some criticism is aired during site visits.

Variables that enter into the pricing equation for a laboratory information system are

- Workload
- Complexity of the system (e.g., multihospital cluster, large outreach programs)
- Number of laboratory modules installed (e.g., blood bank, blood bank donor, anatomic pathology, microbiology)
- Discretionary products installed such as *ad hoc* report generator decision-support software, or even voice dictation
- Number of instrument interfaces
- Complexity and function of the interface to the hospital mainframe computer (e.g., ADT, order entry result reporting, test status)
- Processor power
- Disk storage and archive storage requirements
- System redundancy
- Scheduled down-time requirements

Some possible reasons for a low bid from a vendor in response to a RFP are

1. A reliable company with a good product, but anxious for the hospital as a client (e.g., prestigious hospital, first hospital in a region)
2. A relatively new company anxious to develop a client list
3. A very new company that wants to use the hospital as an alpha test site
4. An established company that has already amortized the cost of its current software product and is not investing resources in research and development at present
5. An underpowered or undersized system
6. The bid is for a microcomputer-based LIS, which will always be cheaper by a substantial margin than a minicomputer-based system
7. The HIS vendor already has the hospital as a client and will bid low on the LIS software to maintain its position in the hospital

It is important to write language into the contract guaranteeing system response time under specified system load conditions. Some vendors will balk at such contract language, calling it unrealistic or unnecessary, whereas others may actually insist on the inclusion of performance criteria in the contract to protect themselves. One of the most important reasons to include specifications about system performance in the contract is to discourage a vendor from undersizing a system recommendation to underbid competitors by keeping hardware costs low.

The advantages and disadvantages (9,10) of the use of an outside consultant in the selection of a LIS are

Advantages

1. Saves work for the laboratory director and staff
2. Provides special expertise and a potentially large repertoire of solutions
3. Identifies problems that may be easier for an outsider to recognize
4. Allows the arbitration of disputes between the various hospital factions
5. Lends an air of credibility to the process

Disadvantages

1. Consulting services have cost, which may be high or perceived as high
2. Selecting an effective consultant may be difficult
3. Some organizations do not readily accept outside advice
4. Consultants are not forced to live with their bad decisions
5. Consultants may be tempted to apply a previous successful solution to all future problems
6. Frequent use of consultants may stifle the development of local expertise of local buy-in to a system
7. Consultants may be unwilling to share the decision-making process with hospital personnel

The major reasons for the failure of a LIS are

1. Poor fit between the LIS and the laboratory/hospital environment
2. Poor design of the LIS such that the replaced manual systems were actually superior
3. Lack of acceptance or sabotage of the newly installed LIS by various user groups such as medical technologists and hospital physicians
4. Unacceptable system response time
5. Inadequate leadership during a politically complex and technically challenging process
6. Inadequate training of the various user groups
7. Inadequate system support from the vendor
8. Belief that purchasing a turnkey system can create unrealistically low expectations about resources necessary to support the system
9. Because it is difficult to hire skilled personnel to run a newly installed LIS, inexperienced personnel must be rapidly trained in addition to other duties
10. Overambitious implementation schedule

MANAGING A LIS

Part of "5 - Laboratory Information Systems"

One of the most important elements for the management of a high-quality LIS is a test system, which consists of a copy of the production software and a patient test result database that can be used to simulate live runs. The basic idea of the test system is that new releases of software can be tested in the local environment before being brought up live on the production system. This is important because no software, even from the most reputable vendor, is error free. Moreover, no vendor can exactly replicate the various computing environments of all their hospital clients when writing new software.

Even though no test system can exactly mimic the production environment of the hospital, it is always preferable to run new software on the test system rather than immediately installing on the production system. If a hospital does not have a test system and new software must be immediately installed on the production system, the only alternative is to fall back to an audit approach in which the system output is monitored closely for a period of time for errors.

Routine Operation

Each LIS unit needs to develop a written set of guidelines pertaining to some important aspects of routine operations. The following issues should be addressed:

- Physical site security
- Backup and recovery
- Software and database security
- Power outages and fluctuation of current

Because the LIS plays such a critical role in laboratory operations and because it is relatively vulnerable to both vandalism and adverse changes in its operating environments, care must be taken to protect the hardware and software from harm. Many of

the systems for protecting the LIS will be put in place when the system is installed such as door locks, fire protection system, and an adequate air conditioning unit, and an uninterruptible power supply. Limiting access to the computer room and surveillance for the presence of unauthorized personnel is the best means for guarding against vandalism of the system.

Backup and Recovery

Backup and recovery are the means to restore a LIS to perfect working order after the system goes down, either on a scheduled or unscheduled basis. Of course, coming up after a scheduled downtime should present no problems in terms of restoring the systems because the system should be taken down on an orderly, regular basis. Unscheduled downs can occur on the basis of either a hardware of software problem. It is possible to build hardware redundancy into a system to minimize the risk of a hardware failure through the use of parallel processors or use of hardware architecture with more than one CPU in a cluster. Such hardware redundancy, of course, does not guard against software defects creeping into the system and corrupting the database. The insidious aspect of software failure is that the system may continue to function on an apparently normal basis. Even if shadowing is in place to protect the database with simultaneous writing of data to two drives, both copies of the files may be corrupted by a software defect.

After an unscheduled downtime, the patient database must be reconstituted assuming that an up-to-date copy of system software is readily available. There are three components of the patient database: active on-line patient files, historical on-line patient files, and purge tapes of patient data. Although backup strategies will vary from vendor to vendor, a typical approach is to write test results to both purge tapes and to the on-line historical file when they are purged from the active files. Therefore, reconstituting lost data after an unscheduled downtime will usually involve asking the individual laboratories to reenter lost test results from active files plus recovering historical patient data from purge tapes.

Reentry of results from the individual laboratories may be simple if data are maintained for a short period of time in the memory of automated laboratory instruments. Manual data entry may also be required on occasion, if results are not maintained in electronic form in the laboratories. The extent to which data are maintained in an electronic format in the laboratories in the event of an unscheduled downtime is a trade-off question, comparing the resources necessary to maintain such files versus the cost of manual data entry at the time of system failure.

System Security

For the most part, turnkey LISs will come equipped with adequate security systems. The lowest level of security, commonly assigned to hospital physicians, is the ability to read test results only. Higher levels of security involving the ability to write to files will be assigned to bench-level medical technologists. Higher levels of security would be assigned to shift supervisors or chief technologists and involves the ability to modify test methodologies and generate management reports. High-level LIS managers in larger hospitals, of course, are assigned the highest security levels.

Most LISs are run as true turnkey operations after the system is installed and stable. This means that a small number of clinical laboratory personnel are trained to supervise and troubleshoot the system, but that it frequently runs unattended. The LIS vendors respond to software problems via dial-in modems. Hardware problems are handled in two different ways. If the equipment has been purchased directly from the hardware vendor, LIS personnel will summon field engineers under contract to attend to hardware problems. If both software and hardware have been purchased from the LIS vendor, apparent hardware problems may also be routed through the LIS vendor. Larger and more complex hospitals may not be content to run LIS operations on a true turnkey basis and will establish a management structure within the clinical laboratories to provide leadership to the clinical LIS operation.

Imaging and Voice Recognition

System automation integration and implementation requires full commitment by the department leadership, fully dedicated informatics technicians, and administrative support. Modernizing the laboratory may find resistance from the conservative pathologist fearing the “change,” administrative concern about employee layoffs, and economic justification for such expense, but if the efficiency of the new system can be demonstrated and the long-term savings can be projected, success should follow. Images are the essence of anatomical pathology. They can be incorporated in the report, be published on the Internet, or be sent to outside centers for special consultation. Gross pictures can be electronically taken and printed as integral parts of the report. Long-distance education and consultation can be achieved by developing an Internet-based imaging strategy. There are some limitations to digital imaging in pathology, cost being one of the most important ones; however, new and better technologies providing faster images with better resolution and more storage capacity are being developed at a lower cost.

Telepathology can be defined as the practice of pathology through the viewing of digital images for primary diagnosis, consultation, distance education, proficiency testing, and so forth. The telepathology system can be set up utilizing a fully robotic microscope that can be controlled at a distance or using a regular microscope with accessories to control the position of the stage controlled by an operator/technician.

Static images can be stored and e-mailed or reported on web sites for consultation, mentoring, or teaching purposes.

In summary, telepathology can be classified according to the type of images (static or dynamic) and the control system of the microscope. It can also be divided according to the type of information that the images contain (general pathology versus specialty), the type of communication (Internet/Intranet), or telephone service.

Another opportunity for laboratory automation and modern departmental reengineering is provided by the acquisition and utilization of automated speech/voice recognition. From the electric typewriter to the first autospeech recognition system instituted in 1990, a tremendous evolution in providing clinicians,

other pathologists, and patients with clear and timely information has been achieved.

The generic operations in anatomic and clinical pathology demonstrate a clear stepwise progression. In anatomic pathology, the steps include accession of the specimen, dictation of the gross description, production of a working draft by the transcriptionist, pathologist review of gross dictation, review of the slides, and preparation of microscopic description (final diagnosis, transcription of final diagnosis and correction of gross description, final corrections by pathologist, and sign out). With the installation of a personal computer voice recognition system, the cycle is simplified to four or five steps including accession, pathologist gross dictation, and real time editing of the description, printing or electronic saving of working draft, issuing of microscopic description/full diagnosis with real-time editing, and signed-out report. The demand for decreased turnaround time and budget reductions for the pathology laboratory have brought up new alternatives in providing quality information. Automated voice recognition has been more easily accepted by other industries and businesses with a large number of voice transactions like telephone companies, stock transaction groups, airlines.

Speech recognition is based on the conversion of analog to digital voice signals by specialized computer hardware. Kurzweil, MedSpeak, and Broca are examples of automated voice recognition for pathologists. The software can be speaker dependent/speaker adaptive, or speaker independent.

Voice recognition implementation decreases the turnaround time, improves productivity, and gives pathologists more independence, providing the entire department with opportunities for increasing the number of specimens diagnosed without adding staff, increasing outreach programs to improve revenue, providing a better competitive outreach advantage.

Different companies are working in providing speech recognition systems with increased speed, a larger vocabulary, an error rate less then 5%, and natural language capability with total speaker independence.

Internet Reporting

We are witnessing a major communication transformation since the invention of the printing press. The Internet has changed the way we think of communication by providing a new, efficient, and effective method of interactive multidirectional transferring and obtaining of information.

The Internet is basically a “network of networks,” a network that links two or more computers. It was created as Aparnet by the U.S. Department of Defense in the 1960s to connect different networks at different sites.

The worldwide web has grown from 130 sites in June 1993 to 4,389,131 as of March 1999. On-line physician activity increased by 87% since 1997, as reported by the Healtheon Corporation (11).

The Internet functions as a virtual endless library. Distance education, telepathology, and Internet reporting are few of the opportunities provided by this technology that is affecting our lives on a scale as significant as the telephone or the television. Ideas, data, and opinions are shared all over the world by people regardless of their age, sex, race, creeds, social background, or country of origin. Strategies for publishing reports, receiving and sending images, consulting colleagues at a distant site, or diagnostic options for remote laboratories are being developed daily among different pathology departments.

The reports published on the Internet can be customized to fill the needs of other pathologists, primary care physicians, oncologists, and patients. Patients are increasingly driving the decision-making process as they become more informed. The pathologist can utilize the Internet to provide other pathologists, doctors, patients, and the general public with integrated information, enhancing the care and well-being of the community.

REFERENCES

1. Hammond WE. *Patient management systems; the early years, The Association for Computing Machinery Conferences on the history of medical informatics*. New York: The Association for Computing Machinery, 1987:153–164.

2. Korpman RA. Laboratory computerization. A new analysis of workflow and reporting. *Clin Lab Med* 1983:3:79–100.

3. Korpman RA. Using the computer to optimize human performance in health care delivery. *Arch Pathol Lab Med* 1987;111:637–645.

4. Hendricks EJ, Langhofer LA. A community hospital laboratory computer system: an eight-year longitudinal study of economic impact. *Am J Clin Pathol* 1982;77:297–304.

5. College of American Pathologists. *Manual for laboratory workload recording method*. Northfield, IL: College of American Pathologists, 1990.

6. Krieg A, Marquardt VC, Lundberg GD, et al. Systems analysis and planning. *Pathologist* 1982;36:30–35.

7. Friedman BA. Laboratory information systems and the competency trap. In: Greenes RA, ed. *Proceedings of the twelfth annual symposium on computer applications in medical care*. New York: IEEE Computer Society Press, 1988:659–662.

8. Connelly DP, Glaser JP, Chou D. A structured approach to evaluating and selecting clinical laboratory information systems. Pathologist 1984;38:714–720.

9. Detmer DE. Improving health and reducing costs through information technology. *Acad Med* 1998;73:286–287.

10. Dimitrov I. The questions of medical informatics (review). *Methods Inform Med* 1996;35:8–11.

11. Growing numbers of physicians are accessing the Internet. *AABB News* 1999;21:9.

Additional Readings

Aller RD, Elevitch FR. *Clinics in laboratory medicine, computers in the clinical laboratory*. Philadelphia: WB Saunders 1983:1.

Aller RD. The pathologist's workstation. *Clin Lab Med* 1997;17:201–227.

Aller RD, Weilert M. Beyond flashy features. *CAP Today* 1998;12:33–34.

Bader SA, Braude RM. “Patient informatics”: creating new partnerships in medical decision making. *Academic Medicine* 1998;73:408–411.

Bergeron BP. Usable voice-recognition technology. *Postgrad Med* 1997;102:39–44.

Blum BI, Duncan K. *A history of medical informatics: Association for Computing Machinery Inc. (ACM)*. New York: Addison-Wesley, 1990:267–277.

Bull BS, Korpman RA. The clinical laboratory computer system: who is it for? *Arch Pathol Lab Med* 1980;104:449–451.

Chapman B. A modern “merger.” *CAP Today* 1999;56:60–70.

Chueh H, Barnett GO. “Just-in-time” clinical information. *Acad Med* 1997;72:511; comment in *Acad Med* 1998;73:1.

Ciotti V. The request for proposal: is it just a paper chase? *Healthcare Finan Manag* 1988;42:48–50.

Elevitch FP, Aller RD. *The ABCs of LIS, computerizing your laboratory information system*. Chicago: American Society of Clinical Pathologists, 1989.

Friedman BA. The challenge of managing laboratory information in a managed care environment. *Am J Clin Pathol* 1996;105(Suppl 1):S3–S9.

Gunton T. *Inside information technology: a practical guide to management issues*. Cambridge: Prentice-Hall International, 1990.

Hardwick DF, Morrison JI, Cassidy PA. Clinical laboratory past, present and future: an opinion. *Hum Pathol* 1985;16:206–211.

Hollander SM, Lanier D. Orientation to the Internet for primary care health professionals. *Bull Med Library Assoc* 1995;83:96–98.

Hospital Computer System Planning. *Preparation of the request for proposal*. Chicago: American Hospital Publishing, 1984.

Jussbaum B, Mickler T, Roby R, et al. Economic impact of a computer-based centralized organization in a clinical laboratory. *Am J Clin Pathol* 1977;67:149–158.

Leslie KO, Rosai J. Standardization of the surgical pathology report: formats, templates, and synoptic reports. *Semin Diagn Pathol* 1994;11:253–257.

Levinson SE. Speech recognition technology: a critique. *Proc Natl Acad SCI U S A* 1995;92:9953–9955.

Mandell SF. The request for proposal (RFP). The key document in successful systems acquisition. *J Med Syst* 1986;10:31–39.

Marquardt VD, Wertz RK, Wertman BD, et al. Relating to consultants. *Pathologist* 1982;36:89–91.

Masys DR. Advances in information technology. Implications for medical education (review). *West J Med* 1998;168:341–347.

McDonald JM, Brossette S, Moser SA. Pathology information systems: data mining leads to knowledge discovery. *Arch Pathol Lab Med* 1998;122:409–411.

Meijer GA, Baak JPA. Reporting by digital speech recognition. *Hum Pathol* 1995;26:813–815.

Rosenthal Al, Chew FS, Dupuy DE, et al. Computers in radiology. Computer-based speech recognition as a replacement for medical transcription. *AJR Am J Roentgenol* 1998;170:23–25.

Teplitz C, Cipriani M, Dicostanzo D, et al. Automated speech-recognition anatomic pathology (ASAP) reporting. *Semin Diagn Pathol* 1994;11:245—252.

Walker K. Selecting computer applications using the RFP process. ***Dimens Health Serv*** 1987;64:18–20.

Weinberg DX. Digital imaging as a teaching tool for pathologists. *Clin Lab Med* 1997;17:229–244.

Wong ET, Abendroth TW. Reaping the benefits of medical information systems (review). *Acad Med* 71996;1:353–357.

Wright B, Bhugra D, Booth SJ. Computers, communication and confidentiality: tales of Baron Munchausen. *J Accid Emerg Med* 1996;13:18–20.

6
Ethics in Laboratory Medicine

Brian J. Linehan

Mohamed M. El-Nageh

Stephen Cordner

Alan Richter

Ethical practice is an integral part of quality laboratory services, as well as the practice of medicine in general. For example, in clinical laboratory medicine, an invalid or wrongly interpreted result is not only of no benefit but may also lead to actual harm for the patient, and in forensic pathology, an incorrect observation or an invalid conclusion may lead to injustice. These are illustrative of the obvious links between quality and ethics.

Pathologists and the staff of medical laboratories are bound by the values and ethics of the medical profession, the ethical codes of their national medical association, and the ethical codes of other relevant scientific and professional bodies. Those with responsibility for the services provided by medical laboratories must accept that, in common with other health professionals, they have obligations of service over and above the minimum required by law. These obligations may vary somewhat from country to country and laboratory management will need to determine what is appropriate for their own situation. The details should be incorporated in a “Laboratory Quality Manual.”

Laboratories must not engage in practices forbidden by law (except in unusual situations in which the law is contrary to generally accepted ethical practice or contrary to the dictates of individual conscience).

All staff working in a medical laboratory should endeavor to protect the human rights of patients and seek to prevent unlawful discrimination and thereby uphold the reputation of their professions.

DEFINITIONS

Part of "6 - Ethics in Laboratory Medicine"

Doctor

Doctor is used in the general sense of a registered medical practitioner. In some parts of the world the term *physician* would normally be used in this way, whereas in other places, the term physician is restricted to a specialist in internal medicine.

Medical Laboratory

Terminology varies from country to country, but the name is intended to be inclusive and cover all institutions providing a pathology service or performing tests on specimens from human beings.

A typical medical laboratory is staffed by pathologists and medical scientists as well as a variety of other staff often including nurses, administrative staff, information technology staff, and accountants. Clearly, however, the details vary with the size of the facility and local custom.

Ethical Practice

Ethical practice can be regarded as appropriate technical practice accompanied by integrity in attitudes and behavior. In deciding what has integrity, reference is often made to moral values voluntarily adhered to within the community and to standards espoused in various codes of professional practice.

WHY DO WE NEED ETHICS?

Part of "6 - Ethics in Laboratory Medicine"

Laboratory medicine, including the whole range of clinical and anatomic pathology, is an integral part of modern medical practice. It would be impossible to practice medicine at a standard considered acceptable in most countries without access to laboratory services. Even in less developed countries, access to some laboratory testing is expected. Medical laboratory services are an integral part of medical services, and there is a public expectation that the same standards and ethical principles that govern the delivery of clinical services will also apply to the delivery of laboratory services.

It may be argued that regardless of how we look at the traditional doctor--patient relationship, the relationship between a patient and the laboratory has little ethical content: the patient has an advocate in the form of the doctor and all that is required from the laboratory is a technically competent service subject to normal commercial rules. In some situations, this would appear to be how laboratory services are supplied, but the authors believe that this is not ideal and in some cases not even acceptable. It is true that the further removed the service is from direct contact with the patient, the less the requirement for the special demands arising from the traditional doctor–patient relationship. Even so, the requirements are still present. It cannot be assumed

that the patient's own doctor will always be able to select and interpret the tests to the maximum benefit of the patient. In some cases, fortunately rare, the medical laboratory may even have to act in the patients' interests to protect them from unethical exploitation by their own clinicians. With the frequency of third-party payment systems, the opportunities for unethical collusion between the treating doctor and the laboratory are common, and it therefore requires an ethical approach on the part of the laboratory, as well as the doctor, to ensure that this does not occur.

Even in reference laboratories, where direct contact with the patient is even further removed, it is still important that the laboratory operates within an overall ethical framework that establishes the patient's interests as always being the first consideration.

Many of the requirements of medical ethics are common to other professions such as accountancy and law; that is, these are common ethical issues not unique to medicine or pathology. Confidentiality is a prime example of this, and in some countries, this requirement is supported by law. However, the law cannot cover all situations. The adoption of a code of medical ethics, as applicable to medical laboratories, increases community confidence that medical laboratories are part of medicine generally and that they will operate to the high standards of ethical behavior required of the medical profession. Medical ethics exist for the protection of patients, and this protection must not be compromised.

Medical laboratories (and increasingly the hospitals of which they may be part) are frequently organized as corporate structures. Many are owned by large national or international companies. Consequently, they often act like corporations, and focus on business-oriented goals. Usually these are profit centered and technology driven, and the law in many countries requires company directors to act in such a way as to maximize wealth for their shareholders.

Unless there is a clearly defined and accepted ethical framework within which medical laboratories operate, it is likely that they will become increasingly commercially oriented with traditional medical values and the public good becoming subordinate to commercial goals.

Many would argue that the delivery of medical laboratory services (or, indeed, all health services) is the same as the delivery of any other commercial service and should be governed by the same commercial rules. This implies the concept of *caveat emptor* (“let the buyer beware”) and assumes the patient or customer is able to determine the value of the services provided. In medicine, this is generally not the case. For market forces to apply in a pure capitalist model, we need unlimited buyers, unlimited sellers, and perfect information. Although it can be argued that these conditions seldom, if ever, exist, even in the normal commercial world, they apply to a much lesser extent in medical interactions. The main reasons why medical services, including laboratory-based medical services, are different from ordinary commercial transactions are

1. *The asymmetry of information between doctor and patient*. Usually the patient is seldom able to evaluate in any meaningful way the quality of the medical information or service supplied. Although increasingly patients will question the accuracy of a medical opinion or the adequacy of a particular outcome, they are usually dependent on the opinion of another doctor in coming to their conclusion. Patients can assess some aspects of a medical service such as whether they have been kept waiting for an excessive period or whether there has been unauthorized disclosure of information. However, the ability of the patient to meaningfully evaluate the medical service is weak.
2. *The patient is often dependent on the doctor and in a weak position psychologically to assess the value of the medical service or consider alternatives*. Effectively the “seller” has a monopoly. This is obvious if the patient is, for example, unconscious, but the same factors apply to a variable extent in most medical encounters. Patients often feel that the doctor or health professional is in a position of power and many feel intimidated. Often they feel particularly dependent on the doctor, and patients recently advised of a serious diagnosis may be in a state of shock and unable to assess a medical service as they might another type of service.
3. *There is often a third-party payer such as the state*. This distorts the normal market signals and alters the normal critical evaluation a customer might make before deciding on a purchase. Again, there is an effective monopoly because the “purchaser” has no motivation (as far as price is concerned) to question the seller, and the system tends to function as if they had no choice.

For these reasons, and perhaps others, the provision of a medical service, including a laboratory medical service, is different from an ordinary commercial service, and there is a special obligation on the person providing that service to ensure that it is appropriate and reasonable for the recipient under the circumstances. This obligation exceeds those that normally apply to a seller of goods or services and is embodied in the concept of ethics.

This does not mean that health professionals or laboratories should not act in an efficient and business-like manner; in fact, they have an obligation to make optimal use of the resources available to them—but they must ensure that their overall purpose is the welfare of the patient. This is achieved by adopting ethical principles that serve as a guide for the operation and governance of the medical laboratory. If their principles are the same as those adopted by the medical profession, then the patient can have the same confidence in the laboratory as he or she has in the doctor.

PRINCIPLES OF ETHICS

Part of "6 - Ethics in Laboratory Medicine"

The essential purpose of medical ethics is to ensure that the patient's trust, and the public's confidence, in the medical profession is deserved. This is achieved by protecting patients and ensuring that they are able to obtain the maximum benefits available from medicine. At the same time, medical ethics aim to protect patients from the abuse that can occur when one person is in a position of power (in this case, based on superior medical knowledge and often status) vis-à-vis another.

Medical ethics are generally considered to be derived from the teachings of the Greek physician Hippocrates (460–377 B.C.), commonly known as the Father of Medicine. He was a man of good character and high moral standards. The ethical principles he taught survive today in the form of an oath (the Hippocratic Oath) traditionally (if not actually) taken by those entering medical practice. Although the exact wording has changed to reflect more modern thinking and practice, the essential principle remains the same: the patient's interests are paramount. The best known modern version is the Oath of Geneva, adopted by the World Medical Association (WMA) in 1948 and subsequently amended in Sydney in 1968 (Appendix 1). The International Code of Medical Ethics of the World Medical Association-1949 was adopted by the WMA at London in October 1949 and has been used as the basis for various codes of ethical practice adopted by different national medical associations (Appendix 2).

The ethical standards of those working in medical laboratories are derived from medical ethics and incorporate the same principles. Public expectations of medical laboratories will reflect, with regard to ethical standards, those expected of the medical profession generally.

For these and other reasons, it is desirable that medical laboratories should be under the overall direction and control of a medically qualified person, usually a pathologist, although it is appreciated that this is not always possible or customary. Nevertheless, any person appointed to direct or control a medical laboratory must accept that the principles of medical ethics will serve to guide the operations of the laboratory. Goals based solely on financial or technical considerations are not acceptable.

It is the responsibility of the professionals working in those institutions to ensure that these principles are practiced and that they are worthy of the same level of trust that the medical profession has come to enjoy.

In recent times, as an aid to decision making in medicine and as a starting point for discussions in medical ethics, four principles have been generally agreed as important. They are derived from the Declaration of Geneva, which in turn is derived from the Hippocratic Oath, and reflect the values and responsibilities implicit in the Oath. These are

1. Autonomy: The right of patients to make decisions on their own behalf. This right presupposes informed consent and the right to knowledge about their condition.
2. Beneficence: The duty or obligation to act in the best interests of the patient.
3. Nonmaleficence: The duty or obligation to avoid harm to the patient.
4. Justice: This embodies concepts of fairness and giving what is rightfully due. It applies not only to the individual but also in the wider medical context and incorporates notions of equity and fair distribution. This is important when medical services are distributed, as they usually are, in an environment of limited resources.

The patient's right to privacy and the correlating medical duty of confidentiality are also important to the practice of laboratory medicine. They can be regarded as subsets of autonomy and beneficence and nonmaleficence. More generally, information is gathered and records are created about the patient for his or her benefit. The patient is usually in the best position to judge whether a particular disclosure will be a benefit. This is the basis of the general rule, which is subject to exceptions, that information should not be disclosed to a third party without the patient's consent.

An ethical decision-making tool is hereby proposed to assist in methodically or systematically tackling an ethical dilemma.

Step 1: Do a stakeholder analysis. This uncovers all who have a stake in the issue and could include, among others, the patient, the patient's family, doctors, laboratory technicians, and medical service organizations.

Step 2: Gather the facts and ensure data integrity. This involves questioning all assumptions and being clear about the facts in the case.

Step 3: Apply the four principles (autonomy, beneficence, nonmaleficence, and justice) to the case. This involves testing that these principles have been upheld. Implicit in the four principles are the concepts of integrity, respect, honesty, and confidentiality. It is also imperative that the case be viewed through the lens of culture, should this be appropriate, because respect, for example, can be expressed very differently across cultures, and these differences must be considered within the boundaries of ethics.

Step 4: Once a decision has been reached, apply the newspaper (or mother/parent) test before acting. This involves conjecturing how a proposed action would be reported in a respected newspaper or what your mother or parent might say to you when told of your action. It is a simple “gut test” to see whether what you plan to do passes your conscience.

GENERAL APPLICATION OF ETHICAL PRINCIPLES TO LABORATORY MEDICINE

Part of "6 - Ethics in Laboratory Medicine"

Responsibilities

Based on a stakeholder analysis, medical laboratories have responsibilities to three main groups:

1. *The patient*. Medical laboratory professionals are accountable for the quality and integrity of the services they provide. This obligation includes maintaining individual competence and endeavoring to protect the patient from incompetent or illegal practices by others.
2. *Colleagues and the profession*. Medical laboratory professionals should strive to uphold the dignity and respect of their professions and maintain a reputation for honesty, integrity, and reliability. They should aim to contribute to the advancement of the profession by improving the body of scientific knowledge, promoting high standards of education and practice, and collaborating with colleagues and other health professionals where practicable.
3. *Society*. Professionals working in a medical laboratory also have a responsibility to contribute to the general well-being of society. This may be within their sphere of professional competence or simply as members of the community.

Medical professionals should comply with relevant laws and regulations pertaining to their professional activities. The medical

profession is committed to a high standard of care and practice, and professionals should endeavor, to a reasonable extent, to influence those that do not meet this standard.

Collection of Information

Laboratories must collect sufficient information to identify adequately patients and specimens. They also should collect sufficient information for other legitimate purposes, but unnecessary information should not be collected. If possible, there should be sufficient clinical information to enable the test to be correctly interpreted. Legitimate concerns may also involve information relevant to the safety of other patients and staff as well as information required for billing purposes and resource management including utilization reviews. The patient should be aware of the information collected and the purpose for which it is collected.

Collection of Specimens

All procedures carried out on competent patients require their informed consent. Where the patient is incompetent by reason, for example, of age or mental state, consent may be given by the parents or a properly authorized person. In exceptional circumstances when this is not possible, necessity may justify the procedure when it is clearly in the best interests of the patient that the procedure be performed. For most routine laboratory procedures, consent can be inferred when the patient presents at a laboratory and willingly submits to the usual collecting procedures, e.g., venepuncture. However, certain procedures, especially the more invasive procedures (e.g., bone marrow aspiration) will require a more detailed explanation of their risks before consent being given. Some tests, e.g., certain genetic testing will require special pretest counseling to ensure that the patient fully understands the implications of the test result.

Adequate privacy for the patient must be made available. It should be appropriate for the type of specimens (or information) being collected and be sensitive to the cultural expectations of the patient and the resources available.

Performance of the Test

All tests will be carried out to an appropriate standard that should be determined in detail by professional organizations or regulatory authorities. Accreditation programs designed to promote standards and ensure compliance are to be encouraged. Where no such guidance is available, the patient's interests will prevail. In some situations, this may mean that a laboratory should refuse to attempt a test rather than produce an unreliable result that could result in harm being done to the patient. All laboratory work will be carried out with the high level of skill and competence expected of the medical, scientific, and allied health professions.

Reporting of Results

Test results are confidential unless disclosure is authorized. They will normally be reported to the clinician who requested the tests and may be reported to other parties with the patient's consent or as required by law. Decisions concerning implied consent for the reporting of results to other involved practitioners (e.g., consultant practitioners to whom the patient has been referred) should be made carefully, taking into account local customs. The laboratory should have written procedures detailing how various requests are handled, and this information should be made available to patients on request. The laboratory is also responsible for taking all reasonable precautions to ensure that the method of transmitting results to requesting clinicians or other authorized persons is secure and reliable. This applies whether the method used is by courier, public post, or electronic means. The laboratory also has a responsibility to ensure that the turnaround time for results is reasonable taking into account the type of test and the patient's condition. There should be the facility to report urgent results as soon as they are available.

In addition to the accurate and timely reporting of test results, the laboratory has an additional responsibility to ensure that, as far as possible, the results are correctly interpreted and applied in the patient's best interests. Care must be given to the construction and format of the report to facilitate correct interpretation and diagnosis. When appropriate, a pathologist or some other competent professional should be available to discuss results. Consultation with regard to the selection and interpretation of tests is part of a medical laboratory service.

Storage and Retention of Medical Records

The laboratory must ensure that information is stored so that there are reasonable safeguards against loss, unauthorized access, tampering, or other misuse. Test results must never be altered or corrected, except by properly authorized persons in accordance with established procedures. The retention of medical records may be defined by various statutory requirements and these need to be considered together with any guidelines issued by relevant professional bodies. Laboratories should develop their own protocols indicating the length of time different results, specimens, and slides will be kept. Test results should be physically readily available for authorized access. When the time comes for medical records to be destroyed, this should be carried out in a way that minimizes the risk of unintentional disclosure.

Access to Medical Records

Access to medical laboratory records should normally be available to the following:

1. the clinician requesting the test,
2. the patient,
3. laboratory and hospital staff if required for the management of the patient,
4. other authorized individuals.

Incompetent patients such as children and intellectually impaired individuals have the same right of access as competent adults, although this right may often be expressed through their parents or authorized agent. Parents, on the other hand, do not always have automatic right of access to their children's medical information, and different countries have different laws and customs in this respect. The laboratory should develop protocols on how to handle different requests with consideration of local laws

and customs. In exceptional circumstances, the withholding of health information from individuals normally authorized to receive it may be justified. An example of such a circumstance is when disclosure may be contrary to the patient's best interests.

When a request is made for access to test results by an authorized person, the laboratory must first satisfy itself as to the identity of the person making the request. The way in which this is done and the degree of certainty associated with the process will vary with different situations.

Different methods may exist in the same laboratory for different tests. For example, a degree of certainty associated with the identity of an authorized person seeking an human immunodeficiency virus result may be much greater than that required of one asking for the results of a routine hemoglobin. Laboratories need to establish appropriate procedures for each situation.

Financial Arrangements and Organizational Matters

Irrespective of ownership, medical laboratories should be under the operational control of a suitably qualified person (normally a pathologist) who takes professional responsibility for the operation of the laboratory, including ethical matters, and the quality of the test results.

Medical laboratories must be able to function with professional independence to avoid conflicts of interest, which is another key ethical principle that medicine has in common with all other professional services. They should not be subject to nonmedical control where this has the potential to interfere with their ability to act freely in the best interests of the patient. They may not enter into financial arrangements with referring practitioners or funding agencies where that arrangement acts as an inducement or an impediment for the referral of tests or patients or interferes with the doctor's independent assessment of what is best for the patient. This assessment, however, will usually be made in an environment of limited resources and so excessive application of these resources to any one individual may not be acceptable, particularly if it results in a failure to deliver a fair share of required services to another individual.

Private laboratory collecting rooms should be completely separate and independent from the referring practitioner's rooms, but when this is not practicable, financial arrangements must be of a normal commercial nature and not include any element of inducement.

Laboratories should also be aware of situations that could give rise to conflicts of interest and take particular care. Such situations may arise when pathologists in private practice can self-refer work. Any such self-referred work must be justifiable.

The medical laboratory has a difficult ethical responsibility when operating in an environment of limited resources provided by a third party such as the state. On the one hand, there is the obligation to ensure that patients receive all the necessary services to which they are entitled, but, on the other hand, there is the obligation to see that resources are not wasted so that other patients are consequentially deprived of their fair share nor the taxpayer (or other funding agent) unreasonably burdened. The practical implications of this will vary in different situations and particularly from country to country. There will also be different pressures on a laboratory depending on whether funding is on a budget or a fee-for-service basis and the extent to which those resources are under the control of the requesting clinician rather than the laboratory. Nevertheless, there is a responsibility on the laboratory to be involved to the extent that is reasonable and practicable in the equitable allocation of resources.

This requirement is based on the principle of justice.

SOME SPECIAL APPLICATIONS

Part of "6 - Ethics in Laboratory Medicine"

Autopsies

Generally speaking, there are two types of autopsy: the hospital autopsy, which requires the voluntary consent of a properly authorized person (usually the senior next of kin) and the forensic autopsy performed at the request or direction of a coroner or other authority to meet statutory death investigation requirements. This section deals with the former.

An autopsy is the postmortem examination of a body to provide information of medical or scientific use including the cause of death or for other relevant purposes such as the resolution of legal issues. The autopsy is an investigation that can have sig-nificant public and private consequences. The pathologist is obligated to ensure the examination is sufficiently thorough to make sure that all the potential benefits are realized.

It can be difficult for families to cope with issues related to autopsies at the time of a bereavement. Hospitals and forensic pathology institutions should have adequate facilities to advise, counsel, and support bereaved relatives. Sudden or traumatic circumstances leading to death are recognized as leading to much psychological stress, and the pathologist, and other staff, should not add to this by insensitivity. The body of the deceased person must at all times be handled with respect and the relatives must be able to rely on this occurring.

Consent for Autopsies

It is recognized that many religions and cultures do not accept the need for, or the desirability of, autopsies, and this must be accepted. Procedures relating to consent for autopsies will usually be governed by law, and these must be followed meticulously. In most countries, the nonforensic or hospital autopsy requires prior consent from the next of kin. This means that the nature and outcomes of the autopsy have been properly and sensitively explained. This will include an explanation of any necessity to retain tissues for the purpose of the autopsy and allow next of kin the opportunity to consider donating tissues that may be used for research or teaching purposes. In some cultures, the removal of the brain and/or the heart, especially if they are retained after the rest of the body is released for burial, is particularly sensitive. The interaction with the next of kin should be conducted in a manner that promotes discussion and encourages them to ask questions. The laboratory should have a clear understanding of the procedures for authorizing an autopsy in the absence of any next of kin or if they cannot be contacted.

Histopathology and Cytology

In the course of a histopathologic or cytologic examination, particular observations may be made, e.g., the presence of spermatozoa, which are not related to the purpose of the examination. Careful thought should be given as to whether such observations should be reported as there could be significant social implications. As in clinical pathology, the seriousness of tissue diagnosis is such that clear information must be provided to assist the treating doctor in properly advising the patient about the diagnosis and its consequences. The pathologist must consider the patient's best interests at all times.

Reproductive Technology

Issues raised by discussions of reproductive technology touch on deeply held convictions and religious beliefs, as well as perceptions about what constitutes a human being or a person, about identity, about the family, and about the sense of one's own characteristics living on in some form after death. Just as these convictions and perceptions vary, so does the way in which different societies, cultures, and religions treat these issues. Consequently, it is not possible to arrive at a universally accepted ethical view on this subject. The different and strongly held views on abortion, artificial insemination by donor, in vitro fertilization, gamete intrafallopian transfer, and other procedures and techniques are examples of this concern. Detailed discussion of these important areas is beyond the scope of this chapter. A laboratory will need to develop its own policies based on local laws and customs.

Clinical Pathology (Clinical Chemistry, Hematology, Microbiology, Immunology)

As with histopathology and cytology, the results of tests in these areas can have a life-altering impact on patients. Information provided about the results, and the manner of its provision, must assist the treating doctor in properly advising the patient about the diagnosis and its consequences.

Transfusion Medicine

The Code of Ethics for Blood Donation and Transfusion adopted by the International Society of Blood Transfusion is given in Appendix 3. The League of Red Cross and Red Crescent Societies adopted the Statement on the Ethics of Voluntary, Non-Remunerated Blood Donation, which is given in Appendix 4.

Voluntary Donation

Blood donors should give their blood voluntarily and without expectation of payment. No pressure should be exerted on a potential donor. Volunteer blood donors give blood of their own free will and without coercion. This is in line with the right to self-determination and rights to protection of physical integrity and privacy.

Protection of the Donor

No coercion or pressure should be exerted on potential donors, who should be provided with adequate information about the process to properly consent to the donation. Blood should be collected under the overall supervision of a physician. Confidentiality concerning all personal donor details, including laboratory results, should be ensured. The donor has the right to know about results of tests on their blood and the use to which their blood is put.

Protection of Recipient

The patient in need of a blood transfusion should, where clinically possible, be provided with reliable information of the risks, benefits, and any available alternatives to blood transfusion.

A proper application of the principle of autonomy means that the patient needing blood (provided she/he retains the capacity to understand and assess the information provided) is free to accept or refuse blood transfusion.

Measures to ensure quality are paramount throughout all the stages of blood transfusion starting with the detailed criteria for donor selection and deferral. This also includes the complete range of management and operational systems needed to ensure the safety of blood and blood components or products to prevent adverse reactions and transfusion-transmitted infections.

Acknowledgment

This chapter is based substantially on El-Nageh M, Linehan B, Cordnev S, et al. *Ethical practice in laboratory medicine and forensic pathology*. Alexandria, Egypt: World Health Organization Regional Office for the Eastern Mediterranean, 1999.

APPENDICES

APPENDIX 1 - Oath of Geneva

Adopted by the General Assembly of the WMA at Geneva in 1948 and amended by the 22nd World Medical Assembly at Sydney in 1968, the Declaration of Geneva was one of the first and most important actions of the Association. It is a declaration of physicians' dedication to the humanitarian goals of medicine, a declaration that was particularly important in view of the medical crimes that had recently been committed in Nazi Germany. The Declaration of Geneva was intended to update the Hippocratic Oath, which was no longer applied to modern conditions.

At the time of being admitted as a member of the medical profession:

I solemnly pledge myself to consecrate my life to the service of humanity;

I will give to my teachers the respect and gratitude which is their due;

I will practice my profession with conscience and dignity;

The health of my patient will be my first consideration;

I will respect the secrets which are confided in me, even after the patient has died

I will maintain by all the means in my power, the honor and the noble traditions of the medical profession;

My colleagues will be my brothers;

I will not permit considerations of religion, nationality, race, party politics or social standing to intervene between my duty and my patient;

I will maintain the utmost respect for human life from the

time of conception; even under threat, I will not use my medical knowledge contrary to the laws of humanity.

I make these promises solemnly, freely and upon my honor.

APPENDIX 2 - International Code of Medical Ethics of the WMA–1949

Adopted by the Third General Assembly of the WMA in London in October 1949.

Duties of Doctors in General

A doctor must always maintain the highest standards of professional conduct.

A doctor must practice his profession uninfluenced by motives of profit.

The following practices are deemed unethical:

1. Any self advertisement except such as is expressly authorized by the national code of medical ethics.
2. Collaborate in any form of medical service in which the doctor does not have professional independence.
3. Receiving any money in connection with services rendered to a patient other than a proper professional fee, even with the knowledge of the patient.

Any act, or advice which could weaken physical or mental resistance of a human being may be used only in his interest.

A doctor is advised to use great caution in divulging discoveries or new techniques of treatment.

A doctor should certify or testify only to that which he has personally verified.

Duties of Doctors to the Sick

A doctor must always bear in mind the obligation of preserving human life from conception. Therapeutic abortion may only be performed if the conscience of the doctors and the national laws permit.

A doctor owes to his patient complete loyalty and all the resources of his science. Whenever an examination or treatment is beyond his capacity he should summon another doctor who has the necessary ability.

A doctor shall preserve absolute secrecy on all he knows about his patient because of the confidence entrusted in him.

A doctor must give emergency care as a humanitarian duty unless he is assured that others are willing and able to give such care.

Duties of Doctors to Each Other

A doctor ought to behave to his colleagues as he would have them behave to him.

A doctor must not entice patients from his colleagues.

A doctor must observe the principles of “The Declaration of Geneva” approved by the World Medical Association.

APPENDIX 3 - Code of Ethics for Blood Donation and Transfusion*

International Society of Blood Transfusion

The preamble to the above Code, which was unanimously approved by the General Assembly of the International Society of Blood Transfusion, held during the Society's 16th Congress (Montreal, 16–22 August 1980), states that its purpose is to “define the principles and rules to be observed in the field of blood transfusion; these should form the basis of national legislation or regulations.” The text of the Code (which was approved by the League of Red Cross Societies) reads as follows:

The Donor

1. Blood donation shall in all circumstances, be voluntary, no pressure of any kind must be brought to bear upon the donor.
2. The donor should be advised of the risks connected with the procedure; the donor's health and safety must be a constant concern.
3. Financial profit must never be a motive either for the donor or for those responsible for collecting the donation. Voluntary non-remunerated donors should always be encouraged.
4. Anonymity between donor and recipient must be respected except in special cases.
5. Blood donation must not entail discrimination of any kind, either of race, nationality or religion.
6. Blood must be collected under the responsibility of a physician.
7. The frequency of donations and the total volume of the blood collected according to the sex and weight of the individual, as well as the upper and lower age limits for blood donation, should be defined by regulations.
8. Suitable testing of each donor and blood donation must be performed in an attempt to detect any abnormalities:
    a. that would make the donation dangerous for the donor.
    b. that would be likely to be harmful to the recipient.
9. Donation by plasmapheresis should be the subject of special regulations that would specify:
    a. the nature of additional tests to be carried out on the donor,
    b. the maximum volume of plasma to be taken during one session,
    c. the minimum time interval between two consecutive sessions,
    d. the maximum volume of plasma to be taken in one year.
10. Donations of leucocytes or platelets by cytapheresis should be the subject of special regulations that specify:
    a. the information to be given to the donor about any drugs injected and about the risks connected with the procedure,

b. the nature of any additional tests to be carried out on the donor,
c. the number of sessions within a given time frame.

11. Deliberate immunization of donors by any foreign antigen with the aim of obtaining products with a specific diagnostic or therapeutic activity should be the subject of special regulations that would specify:
    a. the information to be given to the donor about the substance injected and the risks involved,
    b. the nature of any additional tests which have to be carried out on the donor.

    N.B. The purpose of the special regulations in items 9, 10, and 11 above is to safeguard the donor. After being told about the nature of the operation and the risks involved, a statement of informed consent must be signed by the donor. For donors immunized against red cell antigens, a special card should indicate the antibodies and specific details as to the appropriate blood to be used in case the donors need to be transfused.

12. The donor must be protected by adequate insurance against the risks inherent in the donation of blood, plasma or cells, as well as the risks of immunization.

The Recipient

13. The object of transfusion is to ensure for the recipient the most efficient therapy compatible with maximum safety.
14. Before any transfusion of blood or blood products, a written request, signed by a physician or issued under his responsibility must be made, which specifies the identity of the recipient and the nature and quantity of the substances to be administered.
15. Except for the emergency use of type O blood or red blood cells, every red cell transfusion necessitates preliminary blood grouping tests on the recipient, and compatibility tests between the donor and the recipient.
16. Before administration, one must verify that blood and blood products are correctly identified and that the expiry date has not been passed. The recipient's identity must be verified.
17. The actual transfusion must be given under the responsibility of a physician.
18. In case of a reaction during or after the injection of blood or blood products, appropriate investigations may be required to ascertain the origin of the reaction and to prevent its recurrence. A reaction may require the interruption of the transfusion.
19. Blood and blood products must not be given unless there is a genuine therapeutic need. There must be no financial motivation on the part of either the prescriber or of the establishment where the patient is treated.
20. Whatever their financial resources, all patients must be able to benefit from the administration of human blood or blood products, subject only to their availability.
21. As far as possible the patient should receive only that particular component (cells, plasma, or plasma derivatives) that is needed. To transfuse whole blood into a patient who requires only part of it may deprive other patients of necessary components, and may carry some additional risks to the recipient.
22. Owing to the human origin of blood and to the limited quantities available, it is important to safeguard the interests of both recipient and donor by avoiding abuse or waste.
23. The optimal use of blood and blood products requires regular contact between the physicians who prescribe and those who work in blood transfusion centres.

The Controls

24. Appropriate controls should be required by the Health Authorities to verify that blood transfusion practices meet internationally accepted standards and that the guidelines or regulations issued in accordance with this code are effectively respected.
25. The following should be regularly checked:
    a. the proficiency of the staff,
    b. the adequacy of the equipment and premises,
    c. the quality of methods and reagents, source material and finished products.

Footnote

*The text of this Code (accompanied by an editorial describing its development) was originally published in the *ISBT Newslett* 1980;9 and in *Transfusion Int* 1981;26. It is reproduced here with the kind permission of the International Society of Blood Transfusion.

APPENDIX 4 - Statement on Ethics of Voluntary, Non-Remunerated Blood Donation: League of Red Cross and Red Crescent Societies

The following statement was adopted at the Third International Colloquium on Recruitment of Voluntary Blood Donors, held in Hanover on 22–24 August 1990; it was thereafter endorsed by the International Group of Red Cross Blood Transfusion Experts at its 19th Meeting, held in Los Angeles, CA, on 9 November 1990:

> The League of Red Cross and Red Crescent Societies, in keeping with its humanitarian principles, has always maintained uncompromising support for the concept of voluntary, non-remunerated blood donation.

The XXIVth International Conference of the Red Cross in Manila in 1981 reaffirmed the movement's commitment to voluntary nonremunerated blood donation, and in this context approved the Code of Ethics developed by the International Society of Blood Transfusion as consistent with the principles of the movement. The rapid changes of the last decade have thrown new light on the ethical aspects of blood donation.

Voluntary nonremunerated blood donation is considered among the safest kind of blood donation in terms of security to the recipient as the blood donor does not benefit from the transaction. The donor is expected to communicate without hesitation any contraindication that could have potential harmful effects on a recipient. In recent years, this responsibility has become more onerous, and the questions asked by the transfusion services have become of necessity more personal and more detailed than before.

For these reasons, there are increased ethical responsibilities, some old, some new, placed on transfusion services that collect

blood from voluntary nonremunerated donors. They include the following:

1. No coercion or pressure should be brought to bear on a potential donor to donate.
2. Every transfusion service should have current detailed criteria for donor selection and deferral, and these should be explained to the donor when an occasion for deferral arises.
3. Staff and volunteer who have donor contact should be carefully selected and trained to ensure that donors are handled sensitively and thoughtfully.
4. Donors should be made aware of the ethical responsibilities that they have toward the recipient(s) of their donation.
5. Donors should be assured by every available means that his or her donation is being utilized for patients in need without financial gain for any intermediate party.
6. Donors should be assured that blood and blood products made available through voluntary nonremunerated donation are being used optimally within hospitals for patients in need.
7. Donors should be assured that the transfusion service will treat in a confidential manner all personal donor details, including the results of all laboratory tests.

The Red Cross and Red Crescent Movement continues to support, as consistent with its principles, the ethical implications of voluntary, nonremunerated blood donation.

Section 2 Molecular Pathology

Molecular Pathology - Introduction

Serhan Alkan

Section Chief

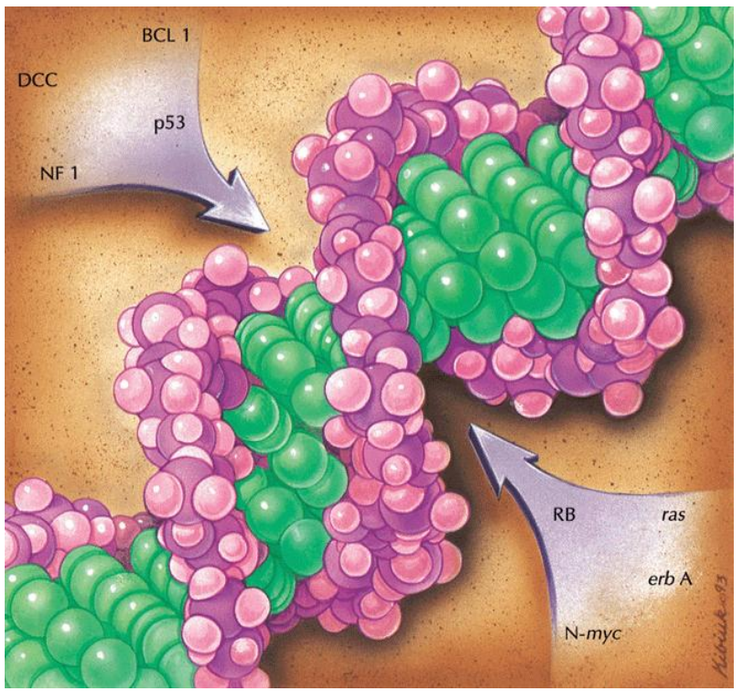


Figure.

7. Introduction to Molecular Diagnostics
8. Molecular Biology of Inherited Diseases
9. Molecular Biology of Solid Tumors
10. Molecular Biology of Infectious Disease
11. Clinical Applications of Molecular Biology Hematopoietic Disorders

The diagnosis and characterization of genetic diseases and of malignancy are increasingly dependent on analysis of specific genetic sequences that determine the pathologic processes in such disorders. This type of molecular genetic analysis is technology intensive; as with other, more established diagnostic methods, proper design and interpretation of these new diagnostic approaches requires understanding of the concepts behind the techniques and of the biological processes and molecular structures that are being examined. Chapter 7 addresses these issues by first providing fundamental information about the organization of genes. This chapter includes discussion of techniques used to detect specific gene sequences, their RNA products, and to determine abnormalities involving overall structure and sequences. This is followed by Chapter 8, a discussion of molecular biology of the commonly seen hereditary disorders.

Chapter 9 outlines the clinical and molecular pathogenesis of solid tumors, several illustrations for applications of molecular methods in the analysis of solid neoplasms in which mutational events are considered to be pathogenic and in some cases used for proper histologic classification.

Microbiology is one of the disciplines most dramatically affected by molecular techniques and Chapter 10 illustrates some of the examples of molecular diagnostic techniques used in clinical practice. Some of these topics particularly utilization of these assays are also covered in the microbiology section. Chapter 11 describes molecular techniques that are used to detect genetic abnormalities associated with hematopoietic malignancies and to characterize common abnormalities associated with coagulative disorders. Molecular characterization of lymphoid and myeloid malignancies represents a major focus of this chapter, as it is increasingly clear that this type of analysis is providing important therapeutic and prognostic information.

7

Introduction to Molecular Diagnostics

Andrea Ferreira-Gonzalez

David S. Wilkinson

Carleton T. Garrett

During the past decade, we have acquired a vast amount of knowledge about genes and their proteins. This new information directly impacts the way we view diseases and is beginning to have a significant impact on many areas of laboratory medicine. The well-defined boundaries that earlier separated medical specialties are disappearing rapidly. Nucleic acid probe technology provides information for diagnosis, determining prognosis, selecting therapeutic modalities, and monitoring disease progression. Examples abound for the application of this technology in microbiology, immunology, forensic science, genetics, and oncology (1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11). Following chapters will discuss current and future clinical applications of nucleic acid probe technology in laboratory medicine. This chapter will provide a framework for understanding the molecular techniques used to detect sequence changes in nucleic acids and the biological processes that such sequence changes can affect. This understanding is important for the meaningful evaluation and interpretation of results.

REVIEW OF NUCLEIC ACID BIOCHEMISTRY

Part of "7 - Introduction to Molecular Diagnostics"

Nucleic Acid Structure

Deoriboxynucleic acid (DNA) contains genetic information that is unique to the particular organism from which it is isolated. DNA consists of a chemically linked sequence of nucleotides. Each nucleotide consists of a nitrogenous base, a pentose sugar (2-deoxyribose) and phosphate. There are four separate bases: adenine (a) and guanine (G), which are purines; and thymine (T) and cytosine (C), which are pyrimidines (Fig. 7.1). The nucleotides are arranged along the strands of DNA sequentially, like beads on a string. In most DNA molecules, there are two DNA strands that wrap around each other in a highly ordered fashion such that a T on one strand always pairs with an A on the other strand and a C always pairs with a G. This pairing of A to T and C to G, referred to as base-pairing, is the basis for nearly all detection systems used to measure differences between fragments of DNA (Fig. 7.1). The resulting double-stranded, helical structure has the DNA bases in the interior connected to each other through hydrogen bonds, while the sugar-phosphate backbone groups lies on the outside (12). Nucleotides contain a phosphate group on the number 5′ carbon atom of the deoxypentose (5′ end), and a hydroxyl group on the number 3′ carbon atom (3′ end). Two or more nucleotides can be joined together through phosphodiester bonds, which link the 5′ carbon atom of the 2-deoxy-pentose from one nucleotide with the 3′ carbon atom of the 2-deoxy-pentose of the other nucleotide. The resulting chain of 2-deoxy-pentose sugars linked together by phosphodiester bonds forms the backbone of the DNA molecule orienting the bases towards the center of the DNA. Because the A:T and G:C pairs have an identical width, the pairing of A:T and C:G creates regularity to the width of the DNA double helix. The base pairs lie 3.4Å apart, and the helical turn repeats approximately every 34Å. The two strands of DNA are complementary to each other rather than identical. That is, for every A on one strand there is a corresponding T on the opposite strand, and a similar ‘complementary relationship’ exists for C and G. The two strands of nucleotides comprising the double helix also are oriented in an antiparallel fashion. That is, the nucleotide at one end of the double helix that contains a free hydroxyl group on the 5′ prime carbon atom (the 5′ end) pairs with a nucleotide that has a free hydroxyl group on the 3′ carbon atom of its 2-deoxy-pentose (the 3′ end) (12).

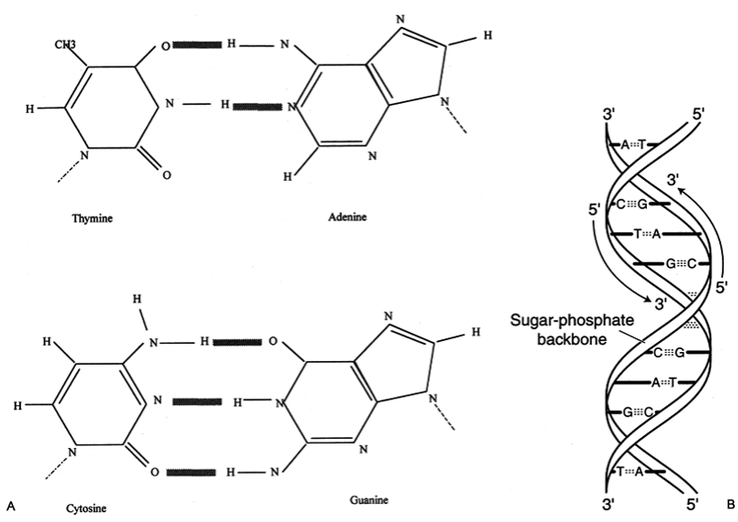


FIGURE 7.1. (**A**) Purine and pyrimidines. The shaded bars correspond to the hydrogen bonds formed between complementary bases. (**B**) DNA double helix. Solid bars correspond to bases and dotted bars correspond to hydrogen bonds between paired bases.

Genome Structure and Replication

In eukaryotic cells, almost all the genetic material is located in the nucleus. Each DNA molecule is organized into discrete units called chromosomes. The total genetic information stored in chromosomes is referred as the genome of the organism. The human genome contains approximately 3×10^9 nucleotide base pairs that are packaged into 23 chromosomes pairs. Twenty-two chromosome pairs are independent of whether the sex is male or female and are called autosomes. The remaining chromosome pair determines the sex and may be either X or Y. Human cells are diploid, which means that they contain 44 autosomes and 2 sex chromosomes, XX for females and XY for males (13). The structure of DNA facilitates its replication. First, the two nucleotide strands of the DNA double helix must separate (14). Separation is accomplished by breaking the hydrogen bonds that hold complementary bases together. The helix is unwound

through enzymatic action. Next, a second enzyme, called helicase, breaks the hydrogen bonds between base pairs, using the energy derived from its ATPase activity causing strand separation (15). The junction between the separated strands and the hydrogen-bounded strands is called the replication fork. Once both strands are separated, each strand can act as a template for replication. Ribonucleic acid (RNA) polymerase ads short nucleotide sequences of RNA (primer) to the separated strands. These primers bind to the DNA strands and provide sites for subsequent addition of complementary nucleotides by DNA polymerase (16). During DNA replication, both DNA strands are copied. The addition of bases occurs in a 5′ to 3′ direction. As noted above, the original strands of DNA are antiparallel. Thus, for one parent strand, the replicating complementary copy grows in the direction that the parent DNA molecule is unwinding and hence can be copied without interruption. This strand is called the leading strand. The remaining strand (lagging strand) must be copied in short pieces, which then are joined together. For the lagging strand, DNA polymerase excises the RNA primers that served as initiation sites for earlier DNA segments and fills in the nucleotide gaps left after their excision. DNA ligase then joins the separately created DNA segments. Because copying of the 3′ to 5′ oriented lagging strand takes place in interrupted fragments, DNA replication is said to be semidiscontinuous (12, 16, 17). Another step during the replication process, associated with the generation of the 3′ to 5′ strand, is proofreading of the newly created nucleotide sequence. If a base added to the chain is not complementary, a 3′, 5′ exonuclease excises the mismatched base and replaces it with the complementary base. Replication is terminated by the addition of characteristic sequences at the 3′ ends of the strands in the presence of the enzyme telomerase. The extraordinary fidelity of DNA replication is necessary to maintain the accuracy of the genetic information from one generation to another (16, 17).

Gene, Gene Structure, and Gene Expression

A gene is the smallest functional unit of genetic information. A gene consists of a sequence of bases that encodes a functional product, either RNA or protein (12, 18). The vast majority of eukaryotic genes contain nucleotide sequences called exons that code for proteins. In addition, there are sequences that are interspersed among the coding sequences called introns, which do not code for proteins (Fig. 7.2). A codon consists of a triplet of nucleotides that specifies one amino acid. There are 64 possible codons. Sixty-one triplets code for different 20 amino acids. One triplet specifies the beginning of a polypeptide, and three codons provide signals for the end of the coding sequence. The process of gene expression consists of transcribing the DNA by RNA polymerase into messenger RNA (mRNA) and translating this mRNA into protein. Codons are read in the 5′-to-3′ direction, however, genes are not transcribed indiscriminately by RNA

polymerase. Other proteins determine whether a particular gene is to be copied into mRNA (12). In addition to the exons and introns, there are regulatory sequences called promoters that govern transcription of a gene (12, 19). The promoters for the genes transcribed by the RNA polymerase I and II usually are upstream of the start point of the gene, while the promoter for RNA polymerase III lies downstream of the start point of the gene. An example of a prototypic eukaryotic protein-coding gene is shown in Figure 7.2. The most common promoter sequence is called the TATA box. It usually is located 25 base pairs (bp) upstream of the start point and it occurs in almost all eukaryotes. Its 8 bp consensus sequence consists of A:T base pairs ATATAA. The TATA box tends to be surrounded by CG-rich sequences, which may be important for function. The role of the TATA box could be to align the RNA polymerase via interactions with nuclear proteins called transcription factors, so that RNA polymerase initiates transcription at the proper site. Another promoter is called the CAAT box (19). The CAAT box (GGCCAATCT) often is located at approximately 80 bp upstream of the starting point. However, it can function at distances from the start point that vary considerably. It also may work in either orientation. Mutational analysis shows that this is one of the strongest promoter elements. The GC box promoter region contains the sequence GGGCGG and multiple copies may be present in the region of the promoter (12).

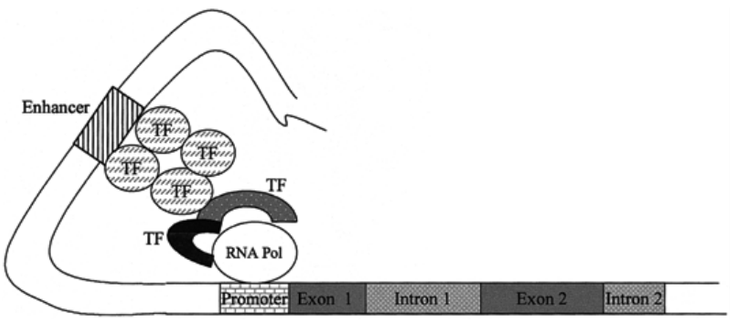


FIGURE 7.2. Gene structure and transcription initiation complex.

Enhancers are another category of regulatory sequences, distinct from promoters, which can increase the rate of transcription (12, 20). Their distance from the promoter can be variable, and their function is independent of orientation. The enhancers of immunoglobulin gene expression have been studied extensively. They are located within the transcription unit. These enhancers only are active in B lymphocytes; thus, enhancers may be part of the mechanism of tissue-specific gene regulation (12). All cells within a higher eukaryotic organism contain the same sequence of DNA, but there are profound phenotypic differences between cells (i.e., liver cell, kidney cell, etc). The phenotype of different cells is a result of differential gene expression within these cells. Complex mechanisms govern regulation of gene expression within each cell type. They cause constitutive expression of some genes, rapid fluctuations in expression of other genes in response to environmental or endogenous stimuli or cell injury, and permanent inactivation of yet other genes (18).

RNA polymerase cannot properly initiate transcription unless it is aided by a number of general transcription factors (20). These transcription factors are protein molecules that are required for transcription but are not part of RNA polymerase. They may bind directly to DNA, to other factors, or to RNA polymerase. Transcription factors provide recognition and specificity for minimal promoter and enhancer sequences. RNA polymerase II promoters have binding sites for transcription factors about 100 bp upstream of the start site. Distances between binding sites can be similar for different genes, but some are specific for a particular gene. Transcription factors bind to specific sequences and form a complex to which RNA polymerase then binds. Initiation of transcription follows formation of this complex (20).

There are two types of transcription factors. One type binds to RNA polymerase when the latter forms the transcription initiation complex (20). These factors are general transcription factors needed for mRNA initiation or termination. Other factors bind specific DNA sequences in the target promoter region. These factors can be general transcription factors or specific to a certain class of promoters. In addition, there are protein transcription factors that bind enhancer regions after recognizing a specific sequence present in the enhancer element. These enhancer regions may act as docking sites for regulatory proteins that move to the vicinity of the gene because of their affinity for the enhancer sequence (21). Once bound to the enhancer sequence, these proteins participate in the assembly of the transcription initiation complex at the promoter possibly by bending the DNA into a loop (Fig. 7.2). Typically, transcription factors contain a helical shaped region that lies adjacent to a cluster of positively charged amino-acid residues. This domain binds specifically to short sequences of double-stranded DNA present in promoters and in enhancer regions (22). Transcription factors do not act in isolation. They form regulatory networks through interactions with each other and by binding to different regions of the DNA to regulate gene transcription. A physically different domain within the transcription factor, called the trans-activator domain, mediates these cooperative associations. These transcription factors can be classified according to certain structural motifs of the DNA-binding domain. Most of the transcription factors contain one of four motifs. These are called the helix-turn-helix, zinc finger, leucine finger, and helix-loop-helix (22).

Gene expression involves the transcription of only one DNA strand (antisense strand). The first nucleotide of the newly created mRNA strand contains a free triphosphate group at its 5′ end, and subsequent nucleotides are linked together through the formation of the 3′ to 5′ phosphodiester bonds. The strand grows in a 5′ to 3′ direction through the addition of nucleotides that are complementary to the antisense strand until a termination signal is reached. The mRNA strand then is freed from the DNA template. The mRNA molecule has the same sequence as the sense strand of the DNA but uracil replaces thymine and ribose replaces 2-deoxyribose (12).

The newly synthesized mRNA molecule, also called immature or heterogeneous RNA (hnRNA), is first processed within the nucleus and then migrates to the cytoplasm where protein synthesis will occur (23). During processing, the hnRNA undergoes a series of post-transcriptional modification steps that result in a mature mRNA. The first post-transcriptional modification is the addition of a special nucleotide sequence at the 5′ end called a cap, which increases the efficiency of gene expression by helping the 40S ribosomal subunit to find and bind to the mRNA in the cytoplasm. At the 3′ end of the molecule, a nuclease recognizes a specific sequence in the hnRNA (AAUAAA) in the noncoding portion of the transcript and cleaves approximately 20 bases from the 3′ end of the molecule. A second enzyme, adenosine terminal transferase, adds a polyadenylate (polyA) tail of at least 200 adenylate nucleotides (23).

Although the entire gene is initially transcribed into hnRNA, it is not all used in assembling the final protein. The introns present in the hnRNA must be removed before translation of the mRNA in order for the appropriate polypeptide to be synthesized. The removal of the introns occurs through a process called splicing (24). Splicing is a highly regulated process and occurs with the help of protein-RNA structures called spliceosomes. Excision of the intron proceeds through a lariat mechanism. In this process, the intron forms a loop (lariat) that enlarges until the 5′ end of one adjacent exon is juxtaposed to the 3′ end of the other adjacent exon. At this point, the two exons form a link via a phosphodiester bond and discard the intron. Small nuclear RNAs participate in this process (25, 26). The splicing process requires absolute precision because deletion or addition of any bases during this process will throw the coding sequence out of frame (26).

Protein Synthesis

Spliced mRNA molecule produced through the preceding process carries the instructions from the gene for the synthesis of proteins (12). The process by which the mRNA is decoded into protein is called translation. Translation occurs in the cytoplasm and requires the involvement of two additional kinds of RNA molecule, transfer RNA (tRNA) and ribosomal RNA (rRNA) (27). The mRNA leaves the nucleus after transcription and enters the cytoplasm where it attaches to the 40S ribosomal subunit. The latter is composed of ribosomal proteins and 18S rRNA. This occurs through a complex series of interactions with proteins called initiation factors (IF), which form the translation initiation complex. As a result of these interactions, all key components are brought together and translation of the mRNA can start. In order for translation to begin, the 40S ribosomal subunit scans through the mRNA from the cap until it finds the correct in-frame start codon, usually AUG. Following this event, binding of the larger 60S ribosomal subunit occurs to form the 80S ribosomal initiation complex. After assembly of the complex, protein synthesis begins. The correct aminoacetyl-tRNA is brought to the ribosomal-mRNA complex by a recognition factor. Following peptide bond formation, translocation of the ribosome relative to the mRNA brings the next codon into position. When the termination codon is reached, the polypeptide chain is released and the ribosome detaches from the mRNA (28, 29).

Genomic Integrity

The double-helical structure of DNA and the complementary base pairing between the purine and pyrimidine bases maintain fidelity of the DNA sequence during its replication and its transcription into mRNA. Translation of mRNA via tRNA into a peptide sequence of amino acids depends upon the accurate reading of the template DNA. Despite safeguards present in the biological system, errors may occur (30). Misreading of the DNA template can occur if a base is not in its usual configuration, i.e., if tautomerization has occurred. Deamination of the bases also can result in erroneous base-pairing (31). If a base is inadvertently deleted, a frame shift mutation may result unless compensatory additions or deletions balance the loss (30). The change of a single base, rather than its loss, represents a point mutation. Mutations may result from chemical or radiation exposure (32). Certain types of DNA damage can be repaired directly by the intervention of endogenous enzymes. DNA photolyase, for example, splits pyrimidine dimers formed by ultraviolet radiation, and 6-O-methylguanine methyl transferase repairs methylated DNA. Less-specific avenues of repair also are available. These include DNA glycosylase, which excises a base from the phosphodiester backbone, 5′->3′ nuclease, which causes chain scission in a DNA strand next to the site of DNA glycosylase activity (incision), and 3′->5′ exonuclease, which excises an offending nucleotide from a strand of DNA (30). Other types of genetic damage, less amenable to repair, include translocation of entire DNA sequences from one chromosome to another, deletion of an entire chromosome (monosomy), or addition of a chromosome (trisomy) (12). Consideration of the mechanisms for these latter changes is beyond the scope of this chapter.

The uniqueness of the base sequence of an organism's genome makes it a highly specific marker for that organism. Moreover, it is becoming increasingly apparent that many disorders are a direct consequence of errors in DNA replication and genetic polymorphisms. These facts, coupled with the elegant simplicity of copying the nucleotide strands, have provided the impetus and current capabilities of molecular diagnostic techniques.

NUCLEIC ACID PROBE TECHNOLOGY

Part of "7 - Introduction to Molecular Diagnostics"

Nucleic acid probe technology represents a collection of techniques and reagents that have as their objective the detection of

a specific sequence of DNA and/or RNA. One of the major advantages of nucleic acid probe technology is that it makes use of nature's most specific marker for all living organisms, which is the sequence of nucleotides comprising the organism's genome. The order of the nucleotides comprising the DNA of one species differs from that in another species. The ability to distinguish the DNA of one person from another has obvious implications in the fields of transplantation and forensic sciences (33, 34). Moreover, because changes in DNA sequences are the cause of genetic and malignant diseases, the ability to measure nucleotide sequence changes in DNA reliably has significant implications for the diagnosis and management of human disease (35, 36). Until 1975, however, practical laboratory methods of detecting differences in nucleotide sequences had not been perfected. In the subsequent years, this situation has changed substantially.

Nucleic Acid Hybridization

Nucleic acid probe technology primarily is based on the simple concept of complementary base pairing between two strands of nucleic acid. Double-stranded DNA (dsDNA) can be denatured into two complementary single strands by simply applying heat, alkali, or other double-helix destabilizing agents, such as formamide. When incubated under the appropriate conditions, the complementary strands can reassociate to reform the duplex double helical structure. The term hybridization refers to this reformation of a sequence-specific, base-paired duplex. Hybridization also can occur between two complementary RNA molecules or complementary RNA and DNA molecules. The temperature at which a specific dsDNA sequence denatures is called the melting temperature or Tm. Thus, the Tm reflects the thermal stability of a double-stranded nucleic acid hybrid. The factors that affect the stability of a double-stranded nucleic acid helix have been studied carefully. The three major factors affecting the melting temperature are the ionic strength of the solution, the presence of specific denaturing agents, and the base composition of the particular nucleic acid sequence. The contributions of all these factors are combined to produce the following relationship:

$$Tm = 81.5 + 16.6 (\log M) + 0.41(\% G + C) - 0.61 (\%form) - 500/L$$

Where: Tm = melting temperature (C°); M = molarity of monovalent cations; %G + C is the percentage of guanine and cytosine present in the DNA; (%form) refers to the percentage of formamide in the solution; and L is the base pair length of the DNA.

Hybridization of two complementary nucleic acid sequences starts with the formation of a short duplex region between the correctly base-paired complementary sequences (nucleation) and this event is followed by a rapid base pairing of the remaining nucleotides on either side of the nucleation sequence (zippering up). The rate of hybridization is significantly dependent upon the rate at which nucleation occurs. The maximum rates of hybridization occur at temperatures approximately 20 to 25°C below the Tm of the nucleic acid duplex. At lower temperatures, the hybridization is slow and cross-hybridization between noncomplementary sequences tends to occur more frequently, because the mismatched duplexes are stable at the lower temperatures. On the other hand, at temperatures approaching the Tm of the duplex, hybridization also tends to be slow because perfectly matched duplexes tend to dissociate at a frequency that approaches the frequency at which nucleation is occurring. Other factors can profoundly affect the rate of hybridization. Low rate of hybridization occurs with hybridization solutions that contain low-salt concentration (low ionic strength) and high viscosity. DNA hybridizes at a higher rate in the presence of inert polymers like dextran sulfate and polyethylene glycol. Formamide often is added to the hybridization solutions to lower the Tm of the duplexes and enable hybridization to occur at lower temperatures. Other parameters that can affect hybridization are the base composition, the pH of the hybridization solution, and the length and concentration of the probe present in the hybridization solution (37).

Isolation of Nucleic Acid

The first task that must be accomplished in order to perform a molecular assay is the isolation of nucleic acid. A number of different procedures have been applied to the isolation of nucleic acid, whether DNA or RNA. Three factors that determine the isolation procedure to be used are the purity, integrity, and amount of nucleic acid that is required to successfully perform the assay. The greater the need to maximize any one or all of these factors, the more labor-intensive and complex must be the procedure utilized to extract the DNA or RNA.

The common starting point for all nucleic acid extraction procedures is the lysis of cells (38). A DNA extraction procedure is facilitated greatly by the use of Proteinase K, detergents, and chelating agents. The detergents dissolve cell membranes and denature proteins present in the cell. Proteinase K is a potent enzyme that degrades proteins. Chelating agents inhibit nucleases that degrade DNA by binding divalent cation ions that are necessary cofactors for the nucleases. After digestion with ribonuclease (RNAse), purification of the DNA from this mixture can be achieved by using organic extraction with phenol, chloroform, and isoamyl alcohol. Lipid contaminants and hydrophobic denatured proteins separate into the organic phase or interface. Nucleic acids partition into the aqueous phase and can be concentrated by precipitation with the addition of salt and alcohol. Organic extraction procedures produce a high yield of highly purified DNA with a molecular weight suitable for use in procedures that require high-molecular-weight DNA. As might be expected, however, the method is extremely labor intensive and uses hazardous chemicals (39).

There are a number of DNA extraction procedures that do not involve the use organic reagents. One method that is widely used is the salting out procedure for isolation of high molecular weight DNA. This procedure is based on the use of a saturated sodium chloride solution that dehydrates and precipitates the partially digested proteins from the Proteinase K treated whole cell lysate. This procedure has the advantage of using nontoxic reagents and is less time consuming compared to organic extraction. The method also allows for extraction of a larger number of cells at one time. One disadvantage is that in a small number of cases, the procedure fails to remove contaminants that subsequently can interfere with molecular testing (39).

More recently, several rapid extraction procedures have been described that can dramatically reduce the time and effort necessary for isolation of DNA (40). Some of these procedures are based on the use of chaotropic reagents to disrupt cellular components and at the same time inhibit exonucleases without causing major DNA shearing. Purification of the nucleic acid is achieved by using spin columns that contain a silica-based gel that can bind DNA in the presence of high salt. The DNA is eluted from the column with deionized water or a low-salt buffer after washing with buffers of decreasing salt concentration. The major advantage of these rapid procedures is minimal time and labor involved for the DNA extraction from a variety of specimens producing good-quality DNA. The trade-off is the small sample size that can be processed at one time, limiting the amount of nucleic acid that can be recovered (41).

Isolation of RNA is more challenging than isolation of DNA. A significant problem encountered during RNA extraction is the degradation of RNA by RNAses. RNAses are very difficult to inactivate and even resist boiling. In addition to their presence within the cells of the tissue being extracted, these enzymes can be introduced easily during laboratory manipulation from a variety of sources including sneezing, hair, dandruff, and fingerprints. Chaotropic agents such as guanidinium isothiocyanate generally are used in the purification of RNA (42). These agents are capable of dissolving cellular membranes, causing nucleoproteins to dissociate from nucleic acid and inactivating ribonucleases. Purification of RNA can be achieved by organic or nonorganic procedures (42, 43). The organic and nonorganic procedures for isolating RNA are similar to the ones already described for DNA. As a final step in purification, RNA should be stored in a solution containing 1 unit/μL of Rnasin.

Southern Hybridization and Restriction Fragment Length Polymorphism Analysis

One of the first nucleic acid probe assays used in the clinical molecular laboratory was the study of hematological and inherited diseases using restriction fragment length polymorphism (RFLP) analysis (44). An important concept for understanding RFLP analysis is that of polymorphism. Polymorphism refers to a variation in the nucleotide sequence at a specific locus in the DNA of two closely related organisms. The term frequently is restricted to nucleotide sequence variations that occur at a relatively fixed frequency and that are present in normal individuals. However, the term also has come to be used to indicate differences in DNA sequences that arise through mutation and through gene rearrangements (12). Restriction endonucleases are a class of enzymes that recognize specific sequences in DNA and cause chain scission of the DNA at or close to the recognition sequence (Fig. 7.3). The recognition sequences usually are 4 to 6 nucleotides in length (12). Recognition sequences for many different restriction endonucleases are present throughout the entire length of the DNA of any given organism. However, because the DNA from one cell is essentially identical to the DNA from every other cell of that same organism, the location of the recognition sites for a given restriction endonuclease essentially is the same in all the DNA molecules isolated from that organism. There are no predetermined rules governing how far apart restriction sites must be located. Thus, over some stretches of the DNA molecule, there will be only a few scattered sites, while over other portions the sites will be closer together. Therefore, when a DNA molecule as large as that present in humans is cut by a restriction endonuclease, fragments ranging in size from a few thousand base pairs (a few kilobases) to 30,000 or more base pairs in length are generated. These can be fractionated by size on an agarose gel. An important point is that the intervening sequence of nucleotides between any two restriction sites for a given fragment from any single cell should be the same as that for the same fragment from any other cell, because the recognition sites are located at identical positions in the DNA from each cell. If following digestion of DNA with *MstII*, the gene for ß-globin is found on a 1.2 Kb fragment of DNA from one cell then it will be on a 1.2 Kb fragment from every other cell. As shown in Figure 7.4, if one extracts DNA from a tissue, digests it with *MstII*, and separates the fragments on an agarose gel, the fragments of DNA with the ß-globin gene accumulate or concentrate at the location on the gel to which 1.2 Kb-long fragments of DNA migrate. This process of concentrating fragments of DNA that possess a specific nucleotide sequence at a specific location on an agarose gel raises their concentration to the level that they can be detected by DNA hybridization (45). Restriction endonucleases are used together with the techniques of gel electrophoresis, DNA transfer, and DNA hybridization to identify a restriction fragment length polymorphism (Fig. 7.4) (46). During gel electrophoresis, the negatively charged DNA or RNA migrates through an inert porous solid medium (i.e., agarose, acrylamide,

etc.) from the negative to the positive electrode. The degree of migration is inversely related to the length of the nucleic acid restriction fragment. Determination of the molecular weight of the nucleic acid fragments is achieved by analyzing controls consisting of fragments of nucleic acid of known molecular weight. These nucleic acid standards are called DNA or RNA ladders. By comparing the rate of migration of the unknown sample to the rate of migration of the ladder, one can determine the size of the nucleic acid fragment in question (45). When the DNA fragments are separated on an agarose gel, the pieces containing the sequence of interest come to lie at the same location on the gel. The DNA then is denatured and transferred out of the gel by gentle convection onto a filter where it becomes irreversibly bound. Denaturation is necessary both to assist the DNA in binding to the filter during transfer, and to enable the DNA on the filter to be recognized subsequently by a DNA probe. The process of transferring the DNA out of the gel and binding it to the filter is called Southern blotting (45). The process of transferring RNA out of the gel onto a filter is called Northern blotting, and transferring protein out of a gel onto a filter is called Western blotting (46). The location of the desired nucleic acid sequence on the filter is determined using a labeled probe by the process of hybridization. The probe usually is 1,000 to 6,000 nucleotides long (1 to 6 Kb) and is labeled with some reporter group, such as a radioactive isotope or by nonisotopic means. It is hybridized with the filter-bound nucleic acid and the target sequence detected by exposing the filter to x-ray film or by development of a colorimetric signal (Fig. 7.4).

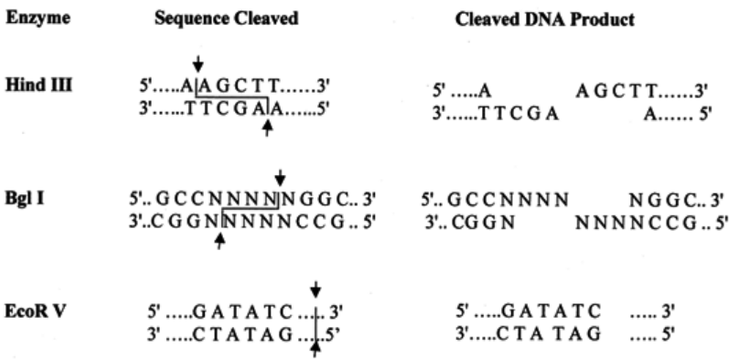


FIGURE 7.3. Restriction Endonucleases. Examples of DNA restriction endonucleases and sequence recognized by them (N represents any base).

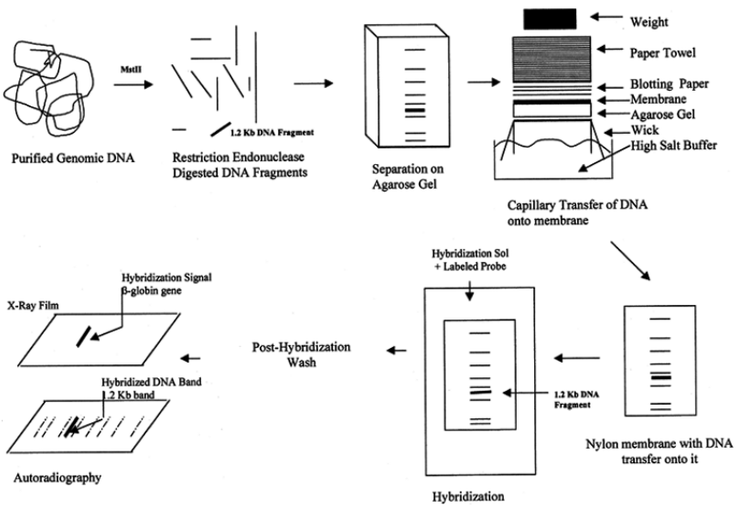


FIGURE 7.4. Representation of Southern Blot Hybridization Analysis. The essential steps for the procedure are represented. Purified genomic DNA is digested with restriction endonucleases and separated in an agarose gel. The DNA is fragmented into smaller pieces by depurination with HCl and the double-stranded DNA is denatured with alkali (NaOH). The denatured DNA in the gel is transferred out of the gel into a nylon membrane by capillary action in the presence of high salt. The nylon membrane is treated with blocking agents and hybridized with a solution containing a single-stranded DNA probe labeled with a radioisotope. After washing nonspecific hybridization products, the nylon filter is exposed to an x-ray film and exposed for autoradiography.

TABLE 7.1. NUCLEIC ACID AMPLIFICATION TECHNIQUES

	Technique	Target	Patent Holder
Target amplification	Polymerase chain reaction (PCR)	DNA	Hoffman-LaRoche
	Ligase chain reaction (LCR)	DNA	Abbott Laboratories
	Transcription mediated amplification (TMA)	RNA	Gen-Probe
	Nucleic acid sequence based amplification (NASBA)	RNA	Organon Tecknica
	Strand displacement (SD)	DNA	Becton-Dickinson
Probe amplification	Invader	DNA	Third Wave Technology
Signal amplification	Branched DNA (bDNA)	DNA	Chiron/Johnson & Johnson
	Hybrid capture	DNA	Digene

Based on the previous discussion, it might be concluded that a DNA probe to a specific target sequence should identify one and only one restriction fragment when hybridized to DNA isolated from the tissue of a single person. In fact, however, each human cell actually contains two complete copies of the human genome, as reflected by the presence of two sets of chromosomes. While one copy of the human genome is very similar to every other copy, the sequences at some loci vary because of the presence of a polymorphism at that specific location. If a variation in the DNA sequence occurs at a recognition site for a restriction endonuclease, the site usually will be eliminated. The restriction enzyme then will have to cut the DNA at a recognition site further from the target sequence in order to create a restriction fragment that contains the target. A second mechanism that can lead to restriction fragments of different lengths is the presence of a sequence consisting of a variable number of tandem repeats (VNTR) that occurs between two restriction endonuclease recognition sites (47). These VNTR sequences consist of a core sequence from eight to ten, to more than 50 nucleotides in length that is repeated from a few to many times. For a given

copy of the human genome, the size of the VNTR sequence is fixed. However, it can vary between different copies of the genome. Thus, because all human tissues except sperm and ova are diploid, examination of a single individual's DNA using a probe to a VNTR sequence may detect two fragments.

If a probe detects a target sequence that can exist in multiple-sized DNA fragments, the probe is said to detect a restriction fragment length polymorphism (RFLP) (48). An individual with two different-sized fragments is said to be heterozygous for that RFLP. Because each single copy of DNA (each set of chromosomes) was inherited from a different parent, this RFLP can be used to aid in paternity testing and for human identity testing. If a neoplastic or genetic disease alters the size of a restriction fragment, the target sequence of the nondiseased tissue may lie at a different location on the filter as compared to that of the diseased tissue. The disease then is said to induce a restriction fragment length polymorphism. Chromosomal rearrangements or deletions and point mutations involving a restriction endonuclease site can all produce RFLPs. RFLPs also can be used to help identify genes that carry mutations and are responsible for inherited diseases through a process called linkage analysis. The latter process relies on the fact that when an RFLP is found to be tightly associated with the development of an inherited disorder, the target sequence must lie relatively close to the gene causing the inherited disease (44, 47).

A number of practical considerations limit the usefulness of RFLP studies in clinical work. First, the methodology is cumbersome and labor-intensive, usually requiring 7 to 10 days to obtain a result. In addition, this methodology sometimes requires large quantities of highly purified DNA of high molecular weight. Second, for many purposes such as detecting point mutations or small deletions or insertions, this method is not sufficiently sensitive unless the mutation alters a known restriction endonuclease site. Even with these limitations, however, for certain disease and disorders, RFLP Southern hybridization analysis is still the method of choice.

In vitro Nucleic Acid Amplification Methods

The Polymerase Chain Reaction (PCR) was the first *in vitro* nucleic acid amplification method developed and it revolutionized life sciences (49). As the result of its development, advances have been made in defining the molecular pathogenesis of many diseases. In the last decade, we have seen development of additional methods of *in vitro* nucleic acid amplification and associated clinical applications. In this section, we will review the basic concepts underlying several of these *in vitro* amplification methodologies. A list of common *in vitro* nucleic acid amplification methods is given in Table 7.1.

Target Amplification Methods

Polymerase Chain Reaction

The PCR causes the *in vitro* amplification of a target DNA sequence by 100,000 to over one-million fold greatly enhancing the ability to detect low levels of target nucleic acid (50). The technique utilizes a set of oligonucleotides called primers (checkered black boxes Fig. 7.5). The primers generally are 19 to 21 nucleotides long and each primer is complementary to a sequence at the 5′ and 3′ ends of the target sequence, respectively. PCR utilizes repeated cycles of DNA polymerase activity to achieve amplification. For PCR to work, the DNA must be denatured. Denaturation is accomplished through use of high temperatures that break the hydrogen bonds that hold the two strands of DNA together. Next, the primers hybridize (anneal) to complementary sequences present at either end of the target sequence. After hybridization of the primers, DNA polymerase extends the primers by adding complementary deoxyribonucleotides beginning at the first base adjacent to the 3′ end of each primer. Through this process, the DNA polymerase creates two new strands of DNA, which are complementary to the target sequence (51). Following the first cycle of denaturation, primer annealing, and primer extension, four strands of DNA exist where initially there were only two. After two cycles, the four strands are converted to eight strands of target sequence. During the third cycle, the eight strands are converted into 16 strands. This geometric increase of the DNA continues for as many cycles as the reaction runs, resulting after ten cycles in about a thousand-fold increase, after 20 cycles in a million-fold increase, and after 30 cycles in a billion-fold increase (Fig. 7.5).

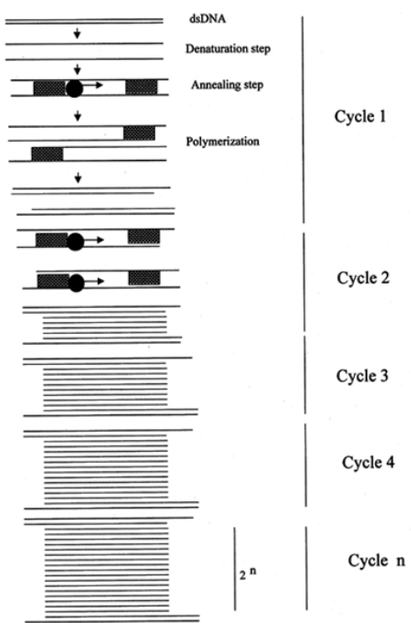


FIGURE 7.5. Polymerase Chain Reaction (PCR). Each cycle of the PCR is composed of denaturation of double-stranded DNA by heat, followed by annealing of primers (checker boxes) to the complementary target sequence. Annealing is favored by lowering the temperature of the reaction below the melting temperature (Tm) of the primer pair. The last step in the cycle is extension by a thermostable DNA polymerase. After one cycle, the number of target copies is doubled. After two cycles, the number of copies of the target sequence is quadruplicated. After 20 cycles, the number of target copies could reach one million.

Major advantages are derived from the enhanced sensitivity of *in vitro* amplification. By increasing the copy number of the target sequence, PCR allows the use of less sensitive detection systems that facilitate use of nonradioactive isotopic detection methods. Another advantage is that the amount of target nucleic acid required for each assay is reduced. The purity of the target sequence generally is less critical because as long as a few target sequences are able to bind the primers, the target can be amplified to a level where it ultimately can be detected. DNA extraction can thus be reduced to a simple, one-step incubation procedure. Detection of small mutations also is greatly simplified. One can amplify a portion of a gene by PCR and then directly

sequence the DNA product without ever having to purify the DNA fragment through cloning. If cloning of the gene is desired, then it is a simple matter to insert the amplified fragment into a vector (52). Sequencing is expected to soon become a standard of practice in the treatment of human immunodeficiency virus (HIV). Mutations in the *pol* gene of HIV-1 may predict resistance of the organism to treatment with zidovudine (AZT) and other therapeutic agents (53). As with any methodology, there are limitations for PCR. The sequence of the target gene must be known so that primers can be prepared to amplify a portion of it. A related problem in microbiology is that the genome of most microorganisms changes over time as a result of mutations. If the mutation occurs in the region of a primer, this mutation may prevent the primer from binding properly to the template and, thus, may inhibit the PCR.

Another limitation of PCR is the potential contamination of negative samples by amplified product (amplicon) (54). Amplified fragments of DNA, called amplicons, may be aerosolized easily into the work environment when sample tubes are opened following amplification; therefore, special precautions must be observed because contamination of a negative sample by even one of these small DNA fragments can cause a false-positive result. Pipetman and other similar pipets, gloves, fingers, clothing, equipment, reagents, “virtually everything with which you might work” can spread amplicon easily (55). Physical and chemical approaches have been introduced to avoid amplicon contamination. Physical separation of pre-PCR and post-PCR work areas is a basic requirement. Use of hydrophobic filtered tips or positive displacement pipettes is another standard. All reagent and sample preparation and handling of patient material should be confined to one area designed as the preamplification area. The sample then should be transported to a separate area, the postamplification area, in which the PCR technique and detection procedures are performed (56, 57). Contamination also can occur from a high background level of template DNA (58). This “high level” is, however, actually so low that its existence may go undetected until one begins to prepare samples for PCR in the contaminated area. This problem arises most often when samples are prepared in an environment where there has been biological amplification through culture or cloning of a target organism or sequence.

Chemical approaches to reduce amplicon contamination have been developed. The most common chemical method is the combined use of the enzyme Uracyl-N-Glycosylase (UNG) and the nucleotide dUTP (59). Inclusion of dUTP in the PCR reaction results in incorporation of dUMP in place of dTMP into the amplicon. The decontamination process consists of treating a sample with UNG prior to PCR. If amplicon containing dUMP contaminates the sample, this dUMP containing amplicon will be destroyed by the action of UNG. UNG will not destroy “true” DNA template in the sample because the latter does not contain any dUMP. The advantage of this system is that there is no need to manipulate the PCR product after amplification and analysis. Even though this approach seems ideal, it has been shown to have limitations with regard to eliminating the effects of amplicon contamination (60). The greater the size of the amplicon and the greater the number of copies of contaminating amplicon, the less likely it will be that the UNG-dUTP procedure will control for contamination.

Another chemical approach is the addition of psoralen compounds to the PCR reaction mixture before amplification. After amplification, the PCR product is treated with ultraviolet light to activate the psoralen compounds, which in turn bind and form adducts with the pyrimidines present in the PCR product. One major disadvantage of this system is that because the psoralins cross-link double stranded DNA irreversibly, the use of probes for PCR product detection no longer is feasible (61).

Reverse Transcriptase PCR

In order to use PCR technology with an RNA target, it is first necessary to use reverse transcriptase (RT) to create a complementary DNA (cDNA) copy of the RNA molecule (62). With this modification, RT-PCR provides an important tool to study gene expression, particularly low copy number mRNA. The first step in RT-PCR is to copy the mRNA into cDNA by RT. A primer is annealed to the mRNA providing a segment of double-stranded

nucleic acid with a free 3′ OH end that can be used for extension by the RT. The RT enzyme then extends the primer by adding complementary nucleotides in a 5′ to 3′ direction. The primer used for this process could be the downstream primer that subsequently will be used in the PCR reaction, or oligo (dT), or random hexamers. The action of the RT causes the formation of an RNA/DNA hybrid molecule. The cDNA strand of this hybrid then can serve as the DNA target molecule in a standard PCR assay. During the initial denaturation step, the RNA is removed and a second DNA strand is created by DNA polymerase from the single-stranded cDNA using the upstream primer. Subsequent PCR cycles then will generate more copies of the double-stranded cDNA target molecule (63, 64, 65). A potential problem that can arise with RT-PCR is that the reaction will amplify genomic DNA that may be contaminating the RNA sample. Even very small amounts of genomic DNA can produce a positive amplification signal in an RT-PCR reaction. One way to distinguish amplification signals of genomic DNA versus mRNA derived cDNA is to design primers that span an exon-exon junction in the cDNA. The presence of an intron in the genomic DNA template will yield a larger fragment than the one from the intronless cDNA template. Unfortunately, this approach cannot always be used, because the necessary information to design primers is not known for a particular gene, or because some functional genes lack introns, as do all processed pseudogenes present in genomic DNA. The use of a well-characterized internal control in the PCR amplification procedure provides a standard by which sample-to-sample comparisons can be made. By using this type of control, meaningful information regarding gene expression can be obtained by RT-PCR (66).

Nucleic Acid Sequence Base Amplification and Transcription Mediated Amplification

A different approach to nucleic acid amplification uses RNA as the target nucleic acid and isothermal amplification (67). One of these procedures is called Nucleic Acid Sequence Based Amplification (NASBA) (Fig. 7.6) (68). This method reverse-transcribes an RNA target into cDNA using a reverse transcriptase enzyme Avian Myeloblastosis Leukemia Virus Reverse Transcriptase (AMLV-RT). A second enzyme, DNA-dependent RNA polymerase subsequently uses this cDNA as a template for RNA synthesis resulting in approximately 100 copies of RNA from one molecule of the original target RNA. The addition of an RNAse H1 enzyme allows degradation of the RNA template thus liberating the cDNA for transcription of RNA. Organon Technika, the holder of the patent for this methodology, has developed a quantitative approach to this method by introducing a series of synthetic RNA molecules of known concentration during nucleic acid extraction (69). Quantitation of target RNA is performed on magnetic beads coated with a specific probe for each RNA control and the RNA target sequence. These probes are labeled with ruthenium, which following excitation by electrical current, causes the emission of light that is proportional to the amount of amplified product bound to the magnetic bead. Transcription mediated amplification (TMA) is a second nucleic acid amplification technique that uses RNA as a target and is isothermal (70, 71). It is very similar to NASBA in its method of amplification.

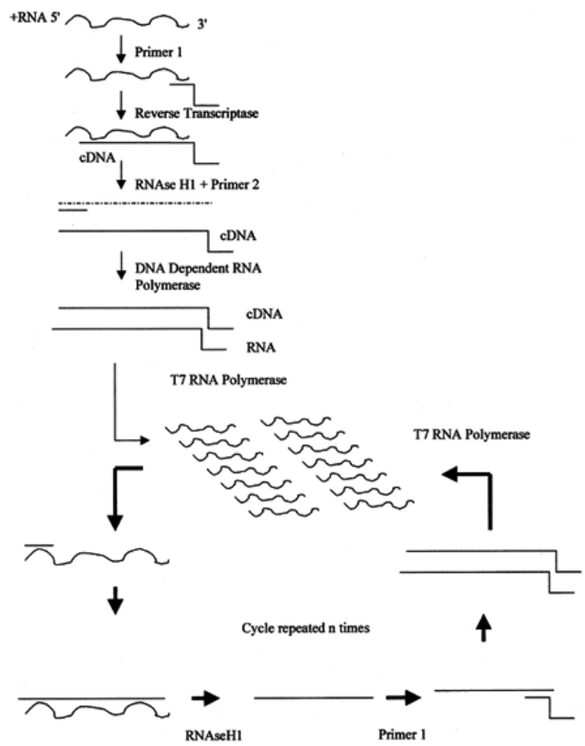


FIGURE 7.6. Nucleic Acid Sequence Based Amplification (NASBA). A ribonucleic acid (RNA) target is reverse transcribed to complementary DNA (cDNA) using Avian Myeloblastosis Leukemia Virus Reverse Transcriptase (AMLV-RT). Ribonuclease (RNAse) H1 degrades the RNA template so the cDNA can be transcribed to RNA. A second enzyme uses the cDNA for RNA synthesis.

Ligase Chain Reaction

The Ligase Chain Reaction (LCR) is an amplification method that first was developed for the detection of point mutations in genetic disease and cancer (Fig. 7.7) (72). Four primers are used in the procedure. Two primers are located adjacent to each other and are complementary to the opposite strand of the target sequence. Under these circumstances, ligase enzymatic activity will join the 3′ end of one primer to the 5′ end of the adjacent downstream primer in both sets of primers. The temperature of the reaction then is raised leading to denaturation and separation of the ligated primer pairs from the target molecule. When the reaction temperature is reduced, nonligated primers can anneal to both the target nucleic acid and the previously ligated primers. These primer pairs then are covalently linked through the action of the thermostable ligase. Because ligated primers subsequently can serve as a target for the complementary set of primers in the next reaction cycle, exponential amplification occurs. If there is a mismatch at the 3′ end of the upstream primers, ligation will not take place (73). More recently, a modification of the LCR was described called the gap LCR. This modification uses four oligonucleotide primers but the two primers in each set are separated by a gap of one or more bases when they are hybridized to the target nucleic acid sequence. After hybridization of the oligonucleotide primers, a thermostable DNA polymerase fills in

the gap followed by ligation of the extended primers by the thermostable ligase (74).

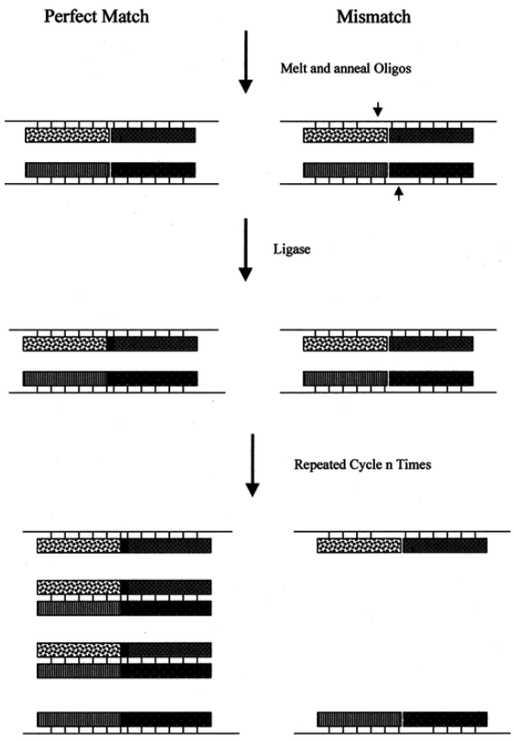


FIGURE 7.7. Ligase Chain Reaction (LCR). The shaded boxes represent primers. The ligase will ligate adjacent primers only if the primer sequence is exactly the same as the genomic DNA sequence at the junction of the primers.

NUCLEIC ACID PROBE SIGNAL AMPLIFICATION TECHNOLOGY

Part of "7 - Introduction to Molecular Diagnostics"

An alternative to target amplification for increasing the sensitivity of nucleic-acid-based testing is to increase the strength of the detection signal that occurs following binding of the probe to the target sequence. Branched DNA technology (b-DNA) uses this strategy called signal amplification (75). b-DNA technology detects and quantitifies the presence of a target nucleic acid sequence through hybridization of multiple probes to the target sequence. The first step in the b-DNA assay is to bind the target nucleic acid sequence to the well of a microtiter plate. This process uses capture probes that hybridize to the target and are themselves covalently attached to the base of the microtiter well. After this step, a series of oligonucleotide probes containing a region that is complementary to the target sequence are added and hybridized to the trapped target. These probes hybridize to sequences located in the target that do not contain any complementarity to the capture probes. In addition, these oligonucleotide probes contain sequences that are designed to anneal to a specially modified b-DNA molecule that produces the detection signal for the system. This branched DNA molecule contains up to 45 sites for oligonucleotides that have been linked covalently to alkaline phosphatase molecules. When this signal amplification system becomes fully saturated, the result is that as many as 1,700 enzyme molecules can become bound to the target sequence. After addition of alkaline phosphatase substrate, the presence of the target sequence is detected by light emission. Detection and quantitation of the target sequence is achieved by measuring the amount of light emitted from the sample compared to the amount of light generated by control samples (76).

Another signal amplification system is the hybrid capture technology. This method uses RNA probes for the detection of a DNA target sequence. In this procedure, one or more RNA probes are hybridized to the DNA target. The RNA/DNA hybrids then are captured in the wells of a microtiter plate using either monoclonal antibody to RNA/DNA hybrids that are attached to the walls of the well or oligonucleotides that are bound

to wells and are complementary to regions of the target. Next, monoclonal antibodies against RNA:DNA hybrids labeled with alkaline phosphatase molecules are added. A chemiluminescent substrate then is added and light emission is measured (77, 78).

Methods for Detecting Amplified Nucleic Acid Sequences

A variety of methods is utilized to determine whether the desired DNA sequence is present after *in vitro* amplification in the clinical sample. These methods are: (a) direct detection of amplicon by gel electrophoresis; (b) solid-phase hybridization; and (c) liquid-phase hybridization.

The simplest and most direct method to detect the presence of a target sequence in amplified product is by direct visualization after gel electrophoresis. Because the sequence of the gene chosen for amplification by PCR usually is known, so is the size of the DNA fragment that will be generated by the PCR reaction. To measure fragment size, the PCR product is placed on an acrylamide or agarose gel and after electrophoresis, the presence of the fragment is detected directly by immersing the gel in a solution of ethidium bromide, which binds preferentially to double-stranded DNA fragments. Direct visualization of the PCR product is achieved by exposing the stained gel to ultraviolet light. If radioactive nucleotides or fluorescently labeled primers have been incorporated into the PCR product, the PCR product can be detected by autoradiography or with a fluorescence scanner. Direct incorporation of radiolabeled nucleotides into the PCR product increases sensitivity and permits detection of DNA fragments on the order of 1/1000 of those detectable by ethidium bromide. Use of direct incorporation of radiolabeled nucleotides or fluorescently labeled primers has the disadvantage that the PCR sometimes can amplify fragments other than the desired target sequence. This method can be made more specific if a restriction endonuclease recognition site is present in the DNA fragment. In the latter case, the PCR product is subjected to restriction digestion. The digested PCR products then are put on a gel and examined to verify that the predicted restriction site is present in the fragment and that it is located appropriately based on the size of the restriction-digested products (46).

In the solid-phase and liquid-phase hybridization procedures, the detection of amplified product is achieved by hybridization of the amplified product with a specific probe, generally an oligonucleotide probe. This approach, when used in combination with gel electrophoresis, has the potential to provide information on both fragment size and homology. In performing hybridization studies with oligonucleotide probes, one must be aware that the melting temperature of the probe, and hence the optimum hybridization temperature for the reaction, is highly sensitive to changes in the base composition and the length of the probe.

In the solid phase hybridization, either the amplified DNA fragments or probes to the amplified target are bound to a solid phase support. A number of different solid supports have been used. Among these are nitrocellulose, nylon, chemically activated papers and, more recently, polystyrene surfaces like microtiter plates. In the past, the most widely used solid support was nitrocellulose, which binds high-molecular-weight DNA with high efficiency but retains nucleic acid fragments below 500 base pair with poor efficiency. In addition, nitrocellulose membranes are fragile and can become brittle. The introduction of nylon membranes as a solid support media for hybridization overcomes some of these problems. Nylon membranes can bind nucleic acid smaller than 500 base pairs with high efficiency. Recently, polystyrene surfaces have been used with good success as solid support media. However, the process of binding target or probe to polystyrene is more complicated than to nitrocellulose or nylon membrane. Polystyrene is the material from which microwell titer plates are constructed. Use of formats incorporating microwell titer plates has become very popular. This format can be automated easily, decreasing the usage of reagents (46).

Solid support hybridization methods include spot-blot, slot-blot, dot-blot and reverse dot-blot formats. In spot-blot, the PCR-amplified fragments are simply spotted onto a nitrocellulose or nylon filter. In the slot-blot and the dot-blot procedures, the PCR-amplified fragments are placed in a small well and then drawn under vacuum through the membrane filter. During this process, the DNA fragments bind to the filter. The filters then are treated either with heat under vacuum or by ultraviolet light to bind the DNA irreversibly to the membrane. Next, the filters are placed in a hybridization solution with a labeled probe. Another commonly used solid hybridization method is the reverse dot-blot hybridization procedure. In this procedure, the oligo probe is bound to the solid phase support and the denatured amplified product is hybridized to the bound oligo probe. In cases where agarose gel electrophoresis of the PCR product has been performed, one can transfer the PCR product onto a nylon membrane using a Southern blot procedure. The membrane then is hybridized with a labeled oligonucleotide probe. The combination of gel electrophoresis and hybridization with an oligonucleotide probe permits two independent means of verifying the identity of the product because one is able to measure the size of the product as well as judge complementarity to the probe.

One approach to detecting small genetic changes is through the use of short oligonucleotide probes called allele-specific oligomers (ASOs). These probes usually are about 16 to 21 nucleotides long. The main characteristic that determines the ability of an oligonucleotide probe to detect very small genetic changes is its size. When a DNA probe anneals to its complementary sequence during DNA hybridization, the energy that holds the two strands together comes from the number of bases that are in proper register with each other. With a very short (oligo) probe, however, all bases must match the target sequence, or the destabilization from one mismatched base is sufficient to prevent a stable hybrid from forming (46).

One method that has been developed to reduce signals resulting from nonspecific hybridization signals generated by the presence of nonspecific hybridization between probe and nontarget sequences is the sandwich hybridization technique (79, 80). In sandwich hybridization, a portion of the amplified product is hybridized to a probe that has been bound previously to a solid support. After the first hybridization, the amplified material that has been bound by the first probe is hybridized to a second probe that is labeled with radioactive nucleotide or other reporter molecule. This approach only can generate positive signal

if there is an appropriate sandwich between immobilized target probe and the labeled probe. The second probe could be labeled with a biotin molecule at the 5′ end. Detection of amplicon then is performed by adding streptavidin conjugated to an enzymelike horse radish peroxidase. After adding the substrate for the specific enzyme, color development occurs that can be measured in a colorimetric reader (80).

In the liquid hybridization system, a portion of the PCR-amplified fragments is mixed directly with a labeled probe. Following hybridization, the samples are placed on an agarose or acrylamide gel and the labeled nucleotides separated from the hybridized fragments via electrophoresis. The hybrids generally are detected by direct autoradiography of the gel (46). Recently, another way of performing liquid hybridization has been described. This method, called the 5′ nuclease assay, detects the accumulation of PCR product as it is being produced in the PCR reaction tube (detection real time) (Fig. 7.8). PCR product is detected through the hybridization of a fluorescently labeled probe and the target sequence. The probes actually are labeled with 2 fluorescent dyes, a 5′ fluorescent dye that acts as a reporter dye, and a 3′ fluorescent dye that acts as a quencher. The probes are 25 to 30 nucleotides in length. When the reported dye is excited, fluorescence emission is blocked by the quencher. During amplification, the probe containing the reporter and quencher dyes binds to the target sequence if the latter is present in the sample. During the extension step of the PCR cycle the polymerase encounters the probe as it extends the primer and cleaves the probe through use of its 5′ exonuclease activity. Following hydrolysis of the probe, the quencher is separated from the reporter dye leading to an increase in the fluorescence intensity from the reporter dye. Because the probe is removed during the extension step, PCR is not interrupted and polymerization can continue to the end of the target sequence. With each cycle more probe binds to the ever-increasing amount of amplified target sequence and subsequent cleavage of the probe results in increased fluorescence. The greater the initial amount of the target sequence in a sample, the more rapid will be the accumulation of fluorescent signal. Because fluorescence is being monitored during each cycle, the point at which a fluorescent signal is greater than the background is a measure of the initial quantity of the target sequence. Thus, the higher the starting copy number for a particular target sequence, the lower will be the number of cycles that the PCR reaction must execute before product is first detected. This is in contrast to traditional detection of PCR product, which is performed at the end of thermal cycling (end point analysis). Real-time analysis of PCR product has two major advantages over end point analysis. First, it eliminates the need for post-PCR manipulation reducing the risk of amplicon contamination; and, second, it increases sample throughput by reducing the time required for product detection (81).

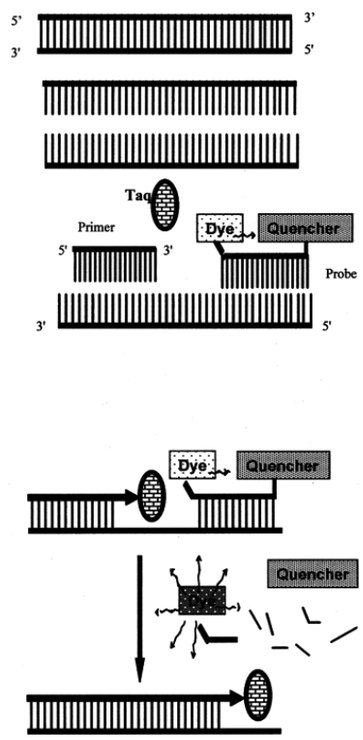


FIGURE 7.8. 5′ Exonuclease Assay (Real time PCR). PCR product is detected from the displacement of a fluorescent labeled probe. The probe is labeled with two fluorescent dyes, a 5′ dye acts as a reporter and a 3′ dye as a quencher resulting in no fluorescence emission from the reporter dye. During PCR amplification, the probe binds the target sequence and is hydrolyzed by the DNA polymerase. In the extension step, DNA polymerase extends the DNA until it finds the probe. The DNA polymerase contains a 5′ ->3′ exonuclease activity and hydrolyzes the probe. In the presence of target sequence, the quencher dye is separated form the reporter dye leading to an increase in fluorescence.

Detection of PCR product generally is facilitated by linking probe or product to some form of reporter molecule. Reporter molecules fall into two broad categories: radioactive and nonradioactive. Nucleotides that have been labeled with radioisotopes such as phosphorus 32 or 33, and sulfur 35 are the most common reporter molecules for research assays and clinical assays developed in house. The detection of this type of label following hybridization is performed commonly either by radioautography, by scintillation counter, or with phosphoimagers (42). The second category of reporter molecules are nucleotides that are covalently linked to one or more hapten molecules such as biotin, a fluorescent dye, or a reporter enzyme. These labels generally are nontoxic and safer than radioactive labels. Biotin-labeled probes generally are detected by using color development with streptavidin-conjugated peroxidases or alkaline phosphatase in the presence of the appropriate substrate. The fluorescent and chemiluminescent products are analyzed by fluorometry, auto-radiography, or luminometry (79).

Peptide Nucleic Acid Probes

Since the discovery in 1991, peptide nucleic acid (PNA) probes have attracted considerable interest as a tool for the analysis of nucleic acids. PNA oligomers are formed from repeating units of 2 amino-ethyl-glycine to which the purine or pyrimidine bases are attached via acetyl linkages. Despite radical differences in the chemical structure between nucleic acid and PNA, hybridization between PNAs and nucleic acids exhibit increased affinity and specificity compared to corresponding oligonucleotides. PNAs of approximately 20 bases are synthesized by manual or automated means. These PNA oligo-probes can be labeled with several different reporters like biotin, fluorescein, rhodamine, or acridine. PNAs can form stable hybrids with complementary DNA and RNA over a broad range of salt concentrations (82).

Direct DNA-Sequencing Procedures

There are several approaches to nucleic acid sequencing. The first method developed by Maxxan and Gilbert used chemical cleavage with four different chemical reactions of DNA strands that previously had been labeled with a radioactive nucleotide. Cleavage occurs at a particular base and reagents are adjusted such that each molecule is cleaved only once. This creates a set of nested DNA fragments that have a common starting point but each successive fragment is one base longer than the preceding fragment. The fragments are analyzed by polyacrylamide gel electrophoresis and autoradiography using gels capable of differentiating fragment sizes down to one base difference (83).

The most commonly used method of DNA sequencing is that of Sanger (84). This method uses four different enzymatic reactions that synthesize a complementary strand to the template DNA by using a primer and DNA polymerase. First, the single-stranded DNA is combined with primer, radioactive nucleotides and DNA polymerase. The polymerase extends the primer along the single-stranded template incorporating four to five labeled nucleotides. The sample of DNA and primers then are divided into four aliquots and each aliquot is combined with a second set of nucleotides that contains three deoxynucleotides and one dideoxynucleotide. The dideoxynucleotide differs from the normal deoxynucleotide in that it does not possess a 3′ hydroxyl group on the pentose sugar. For this reason, when it is incorporated into the extending DNA strand, the polymerase is not able to continue to add additional nucleotides to the growing strand. Termination of the DNA strand therefore occurs at the site of incorporation of the adenine, cytosine, guanine, or thymidine dideoxynucleotide. This procedure results in a series of single-stranded fragments of varying lengths. In the case of strands generated in the presence of dideoxythymidine, all of the strands will terminate with a thymidine. Because the incorporation of the dideoxynucleotide is random, some strands will have terminated after extending nearly to the end of the template while others will terminate very early during the extension process. When each of the four dideoxynucleotide reaction mixtures is subjected to electrophoretic separation on a denaturing acrylamide gel, the autoradiogram of these four mixtures results in what is called a sequencing ladder. By examining the autoradiogram either directly or with the aid of a image analyzing system, the sequence of the target DNA can be directly determined.

The Sanger sequencing method formerly required that the target sequence first be purified from all other possible contaminating sequences through a process termed cloning. This latter procedure could require weeks to months to accomplish and was highly labor-intensive. This made using DNA-sequencing for clinical purposes essentially unrealistic. However, through the use of PCR, the lengthy cloning procedures required to perform DNA-sequencing now can be circumvented through the process of direct amplification of the desired sequence from the DNA in the patient's specimen (85, 86, 87, 88). Once the amplified fragment has been obtained using PCR, it is subjected to further rounds of denaturation, primer annealing, and extension in which only one primer (sequencing primer) is used in conjunction with appropriate dideoxynucleotides. This procedure is called cycle sequencing and results in the generation of many single-stranded copies of the target sequence that can be examined on sequencing gels.

An additional improvement in the technology of this field has been the development of automated DNA-sequencing devices (84). Some improvements in the methodology have been introduced but the most important has been the introduction of robotic workstations to set up all reactions and automate loading of DNA-sequencing devices. A number of these devices use fluorescent-labeled primers or labeled dideoxynucleotides to tag the newly synthesized DNA fragments. Two formats have been used. One format utilizes a sequencing primer that is used in all four reactions and that is labeled with the same fluorescent dye. After the reactions are completed, each sequencing reaction is analyzed in a separate lane on the sequencing gel. The other format uses four fluorescent-labeled primers where the primer in each sequencing reaction has a unique dye. After the reactions are completed, the four different reactions are combined and resolved in a single lane in a sequencing system. In both approaches, DNA is detected using laser-activated fluorescence during the separation process. The automated laser fluorescent DNA sequencer from Pharmacia (Piscataway, NJ) uses a single dye labeled primer and the four-lane separation approach. In this instrument, a fixed red helium neon laser beam is located at the bottom of the sequencing gel and emits light through the entire width of the gel. As the sequencing reactions are electrophoresed in a sequencing gel, the tagged DNA strands pass through the laser beam and emit fluorescent light. The fluorescent light is then captured by 40 photodiode detectors. These detectors, placed at a right angle to the beam at the face of the gel, capture the signal and send it to a computer where the sequences are determined. Since four sequencing reactions are required per sample and the sequencing gel contains 40 lanes, a maximum of ten samples can be analyzed at one time.

Another sequencing device is the Model 377 ABI Prism Automatic DNA Sequencer from Applied Biosystems Division of Perkin-Elmer (Foster City, CA) (Fig. 7.9). This sequencing device uses the four-dye, single-lane approach previously described. An argon laser scans back and forth along the bottom of 36 lanes of the sequencing gel. Recent updates to the system include the capability of gels with 96 lanes. When the tagged DNA reaches the area where the laser is located, it emits light that passes through a prism. The prism separates the light emitted from each dye, which then is sent to a photomultiplier tube that subsequently

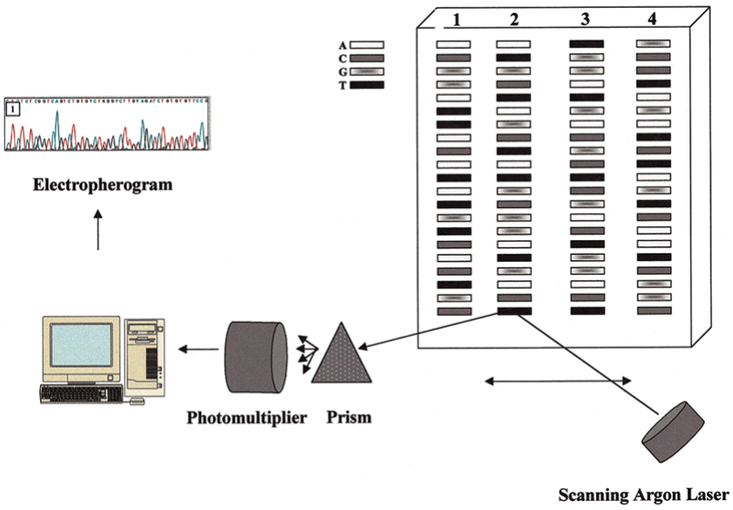


FIGURE 7.9. Automatic DNA Sequencing. Single-lane automatic DNA sequence analysis. The 377 ABI Prism automatic DNA sequencer uses four fluorescent labeled primers or four fluorescent labeled dideoxynucleotides to tag the newly synthesized DNA fragments. The four reactions are combined into a single tube and are separated by electrophoresis in a single lane. When the tagged DNA reaches the area where the laser is located, it emits a light that passes through a prism. The prism divides the light emitted from the dyes and sends it to a photo-multiplier and subsequently to a computer where the sequence is determined. The order of the colors as they pass through the laser reflects the DNA sequence.

sends the appropriate signal to a computer where the sequence is determined. The 310 ABI Prism Automatic DNA Sequencer uses the same detection method as the 377 model, but separates the sequencing reaction by capillary electrophoresis (89, 90). After the sequencing reactions are finished, they are placed in a sample tray and almost all subsequent aspects of analysis are automated. This device uses preformulated polymers that are injected automatically into capillaries. This eliminates the time consuming need to pour slab gels. An electrokinetic injection system is used to load each sample into the capillary for electrophoresis. The automatic loading system ensures consistent and reproducible handling of samples. The 310 Genetic Analyzer simultaneously detects four different fluorescent dyes in a single capillary with the ability to operate unattended for 48 hours.

Even though the devices described above are major improvements over manual sequencing, these systems still have a limited throughput. Alternative approaches are being developed to produce sequence results faster and more economically than currently possible. One novel approach is sequencing by hybridization (SBH) on chips, also called “DNA chips.” These DNA chips are composed of a large number of oligonucleotides of known sequence arranged on a solid support. The location of each specific sequence on the chip is mapped and known. A labeled DNA fragment is applied to the solid support and hybridizes to some of the oligonucleotide sequences. Based on the location of the hybridized DNA fragments, a computer program deduces its sequence (91, 92, 93). In order for SBH to be practical, miniaturization of the solid support is necessary. Several different approaches to SBH have been developed using micro-arrays. One of these approaches called the GeneChip (Affymatrix, Santa Clara, CA) contains a high-density array of oligonucleotide probes placed on a miniature silicon solid support. By combining photolithography technology used in semiconductors and solid phase chemical synthesis, light-directed oligonucleotide synthesis can be carried out on the surface of the silicon chip. The efficiency of synthesis is approximately 95%, which limits the length of the oligonucleotide that can be synthesized currently to 30 nucleotides in length.

At present, the Affymatrix chip contains 400,000 different probe regions. Because DNA is comprised of four bases, the different oligonucleotides are arranged in clusters of four for each base position. The labeled target hybridizes to the oligonucleotides that contain a sequence complementary to portions of the target molecule. By using laser confocal microscopy and computer-assisted analysis, the location and thus the sequence of the hybridized fragments can be deduced (90, 92).

The challenges facing SBH are common to a number of nucleic acid probe technologies. These include probe selection, target preparation, and probe immobilization as well as expensive equipment used to make and analyze this type of micro-array. Another approach developed by Nanogen, Inc. (San Diego, CA) captures prefabricated oligonucleotides onto electro-active spots on silicon wafers that are controlled independently. Each of these independent locations has a unique oligonucleotide attached to the surface where the electric potential can be controlled independently. By applying an electrical potential to each location of the micro-array, the target DNA is pulled toward the spot and hybridization can occur in minutes instead of hours. DNA chip technology appears extraordinarily promising, but it is far from becoming a routine tool for use in the molecular diagnostics laboratory at this juncture (91, 93).

AUTOMATION AND TECHNOLOGICAL ADVANCES IN MOLECULAR DIAGNOSTICS

Part of "7 - Introduction to Molecular Diagnostics"

Utilization of molecular testing has steadily increased in all areas of the laboratory and this phenomenon shows no signs abating. A number of molecular tests have become standard of practice (e.g., HIV-1 viral load, DNA tests for cystic fibrosis, and other inherited disorders). This has led to an increase in test volume and consequently to the need for more automation, better detection systems, and more standardization of current manual procedures. In the past, radioactivity has been the most popular method for labeling and detecting nucleic acid probes. Radioactivity is inexpensive and sensitive, but disposal of radioactive material is expensive and hazardous. Recently, new reporter molecules have been introduced that are sensitive and nonhazardous. Some of these have come into routine use for a number of molecular assays. Fluorescence now is the basis of almost all DNA chip-based technology and luminescence also has become more widely used because of the wide range of its linear response. Use of these reporter methods has facilitated automation of the detection of *in vitro* amplified product.

Automation of other aspects of molecular testing is likewise increasing (94). Thermocyclers are an example of an automated, single-function instrument that is in common use. Automatic nucleic acid extractors are another example of single-function instruments. A number of automatic nucleic acid extractors are commercially available using different extraction principles.

More recently, multifunction devices that are capable of performing more than one step of the testing process have become available. An example of one such device is the Roche Cobas Amplicor instrument. This device contains thermocyclers for performing multiplex PCR and reflex testing in addition to a three-axis cartesian robotic pipeting station for automated processing, detecting, and quantifying of amplicon present in each sample. Full-function molecular diagnostics automated devices that perform extraction, amplification, and detection are limited at present. One such device from Roche Molecular Systems is composed of a sample preparation unit (Ampliprep) coupled to a completely enclosed amplification, detection, and quantification system.

The development of micro-array devices is expected to have a major impact on automation in molecular diagnostic testing. The applications of DNA chip technology focus on the study of genomic structure and the study of gene expression. High throughput studies of genomic structure are expected to become important as single nucleotide polymorphisms (SNPs) are identified that relate to the metabolism of therapeutic agents and that predict our susceptibility to develop a host of common diseases (95, 96). The other application for micro-arrays is to determine gene expression profiles for cells and tissues. Gene expression arrays hold great promise for the characterization of the cancer cells for diagnostic, prognostic, and even therapeutic purposes (97, 98, 99). Cancer cells usually contain a number of genetic alterations that may be more easily detected at the level of gene expression than through direct examination of the cell's genome.

Microfluidics and chip technology likely will contribute substantially

to the automation of molecular testing. The concept of lab-on-a-chip has been proposed in which the chip would perform a large number of the processes currently performed by laboratory personnel (100). These chips would begin with whole blood and perform nucleic acid extraction, restriction enzyme digestion, gel electrophoresis, hybridization, and detection without the need for human intervention at any step. An early application of this concept has been incorporated by Clinical Microsensors (Pasadena, CA) into a hand-held, point-of-care instrument. At the core of their technology is the bioelectronic detection of nucleic acid, either DNA or RNA in a variety of different samples. They use microchips that contain a number of electronically active “pads” coated with specific DNA capture probes. The specific probes on the microchip surface directly capture complementary target DNA sequences. A signal probe labeled with ferrocene is added and hybridized to the target sequence. The ferrocene molecule acts as an electron donor. When a slight voltage is applied to the sample following hybridization, the electrons of the ferrocene molecules rapidly tunnel through the double-stranded DNA yielding an electronic signal that can be detected through the microelectrode.

Finally, another technology that holds promise to facilitate automation is matrix-assisted laser desorption and ionization mass spectrometry (MALDIMS) (101). In this technique, the target is immobilized as a spot, and simultaneously probed with oligonucleotides that are labeled with a unique mass marker. A laser beam is used to vaporize and ionize a small amount of material, which in turn is analyzed by a mass spectrophotometer (102). The methodology combines the potential for high sensitivity, specificity, and fast turn-around time, but the complexity and the cost of the instrumentation make it impractical for clinical laboratory use at this juncture.

REFERENCES

1. Hruban RH, Offerhaus GJ. Molecular diagnosis of cancer and micrometastases. *Adv Anat Pathol* 1998;5:175–178.

2. Sawyers CL. Chronic myeloid leukemia. *N Engl J Med* 1999;340:1330–1340.

3. Cleary ML, Chao J, Warnke R, et al. Immunoglobulin gene rearrangement as a diagnostic criterion of B-cell lymphoma. *Proc Natl Acad Sci USA* 1984;81:593–597.

4. Reeder DJ. Impact of DNA typing on standards and practice in the forensic community. *Arch Pathol Lab Med* 1999;123:1063–1065.

5. Brettell TA, Inman K, Rudin N, et al. Forensic science. *Anal Chem* 1999;71:235R–255R.

6. Yin JA, Tobal K. Detection of minimal residual disease in acute myeloid leukaemia: methodologies, clinical and biological significance. *Br J Haematol* 1999;106:578–590.

7. Madrigal JA, Arguello R, Scott I, et al. Molecular histocompatibility typing in unrelated donor bone marrow transplantation. *Blood Rev* 1997;11:105–117.

8. Powell SM, Petersen GM, Krush AJ, et al. Molecular diagnosis of familial. *N Engl J Med* 1993; 329:1982–1987.

9. Fredricks DN, Relman DA. Application of polymerase chain reaction to the diagnosis of infectious diseases. *Clin Infect Dis* 1999;29:475–486.

10. Handel-Fernandez ME, Vincek V. Early cancer detection by microsatellite marker analysis. *Med Hypotheses* 1999;53:114–117.

11. von Knebel Doeberitz M, Lacroix J. Nucleic acid based techniques for the detection of rare cancer cells in clinical samples. *Cancer Metastasis Rev* 1999;18:43–64

12. Lewin B. *Genes IV*. Cambridge UK, Oxford University Press, 1990.

13. Sommers MS, Beery T. Foundations of genetics: genetic structure, function, and therapeutics. *AACN Clin Issues* 1998;9:467–482.

14. Lohman TM, Bjornson KP. Mechanisms of helicase-catalyzed DNA unwinding. *Annu Rev Biochem* 1996;65:169–214.

15. Marians KJ. Helicase structures: a new twist on DNA unwinding. Structure. 1997; 5:1129–1134.

16. Brautigam CA, Steitz TA. Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr Opin Struct Biol* 1998;8:54–63.

17. Jeffery J. Enzymes: chemistry and biochemistry. *EXS* 1995;73:79–104.

18. Bissell MJ, Weaver VM, Lelievre SA, et al. Tissue structure, nuclear organization, and gene expression in normal and malignant breast. *Cancer Res* 1999;59:1757–1763S.

19. Burley SK. The TATA box binding protein. *Curr Opin Struct Biol* 1996;6:69–75.

20. Latchman DS. Transcription factors: an overview. *Int J Biochem Cell Biol* 1997;29:1305–1312.

21. Littlewood TD, Evan GI. Transcription factors 2: helix-loop-helix. *Protein Profile* 1995;2:621–702

22. South TL, Kim B, Hare DR, et al. Zinc fingers and molecular recognition. Structure and nucleic acid binding studies of an HIV zinc finger-like domain. *Biochem Pharmacol* 1990;40:123–129.

23. Newman AJ. Pre-mRNA splicing. *Curr Opin Genet Dev* 1994;4:298–304.

24. Newman A. RNA splicing. *Curr Biol* 1998;8:R903–905.

25. Varani G. RNA-protein recognition during RNA maturation and development. *Nucleic Acids Symp Ser* 1997;36:36–38.

26. Adams MD, Rudner DZ, Rio DC. Biochemistry and regulation of pre-mRNA splicing. *Curr Opin Cell Biol* 1996;8:331–339.

27. Noller HF, Green R, Heilek G, et al. Structure and function of ribosomal RNA. *Biochem Cell Biol* 1995;73:997–1009.

28. Bommer UA, Lutsch G, Stahl J, et al. Eukaryotic initiation factors eIF-2 and eIF-3: interactions, structure and localization in ribosomal initiation complexes. *Biochimie* 1991;73:1007–1019.

29. Kimball SR, Mellor H, Flowers KM, et al. Role of translation initiation factor eIF-2B in the regulation of protein synthesis in mammalian cells. *Prog Nucleic Acid Res Mol Biol* 1996;54:165–196.

30. Modrich P. Mistmatch repair, genetic stability, and cancer. *Science* 1994;91:1959–1960.

31. Featherstone C, Jackson SP. DNA-dependent protein kinase gets a break: its role in repairing DNA and maintaining genomic integrity. *Br J Cancer* 1999;80:14–19.

32. Morgan WF, Day JP, Kaplan MI, et al. Genomic instability induced by ionizing radiation. *Radiat Res* 1996;146:247–258.

33. Little AM, Marsh SG, Madrigal JA. Current methodologies of human leukocyte antigen typing utilized for bone marrow donor selection. *Curr Opin Hematol* 1998;5:419–428.

34. Tsongalis GJ, Wu AH, Silver H, et al. Applications of forensic identity testing in the clinical laboratory. *Am J Clin Pathol* 1999;112:S93–S103.

35. Macintyre EA, Delabesse E. Molecular approaches to the diagnosis and evaluation of lymphoid malignancies. *Semin Hematol* 1999;36:373–389.

36. Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N Engl J Med* 1997;337:529–534.

37. Dangler CA. *Nucleic Acid Analysis: Principles and Bioapplications*. New York: Willey-Liss, 1996.

38. Jeanpierre, M. A rapid method for the purification of DNA from blood. *Nucleic Acids Research* 1987;15:9611–9614.

39. Coleman WB, Tsongalis GJ. *Molecular Diagnostics: For the Clinical Laboratorian*. Totowa, N.J.: Humana Press, 1997.

40. Dixon SC, Horti J, Guo Y, et al. Methods for extracting and amplifying genomic DNA isolated from frozen serum. *Nat Biotnol* 1998;16:91–96.

41. Attal J, Puissant C, Houdebine LM. An improvement of a rapid method using Qiagen columns to purify plasmids. *Biotechniques* 1990;8:269–271.

42. Chemczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques*. 1993;:532–534, 536–537.

43. Krafft AE, Duncan BW, Bijwaard KE, et al. Paraffin-embedded tissue: the Armed Forces Institute of Pathology experience and literature review. *Mol Diagn* 1997;2:217–230.

44. Chehab FF, Doherty M, Cai S, et al. Detection of sickle cell anemia and thalassaemias. *Nature* 1987;329:293–294.

45. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *L Mol Biol* 1975;98:503–517.

46. Sambrook J, Fritsch, EF, Mariatis T. *Molecular Cloning. A Laboratory Manual*, 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989:9.1–9.32, 11.45–11.48.

47. Nakamura Y, Koyama K, Matsushima M. VNTR (variable number of tandem repeat) sequences as transcriptional, translational, or functional regulators. *J Hum Genet* 1998;43:149–152.

48. Blomeke B, Shields PG. Laboratory methods for the determination of genetic polymorphisms in humans. *IARC Sci Publ* 1999;148:133–147.

49. Saiki RK, Scharf S, Falcona F, et al. Enzymatic amplificaiton of betaglobin genomic sequence and restrictoin site analysis for diagnosis of sickle cell anemia. *Science* 1985;230:1350–1354.

50. Mullis KB and Falcona FA. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 1987;55:335–350.

51. Oste, C. Polymerase chain reaction. *Biotechiques* 1988;6:162–167.

52. Oste C. Controlling false positives in PCR. *Diagnostics and Clinical Testing* 1989;27:50.

53. Larder BA, Kemp SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* 1989;246:1155–1158.

54. Kwok S, Higuchi R. Avoiding false positives with PCR. *Nature* 1989;339:237–238.

55. Porter-Jordan K, Garrett CT. Source of contamination in polymerase chain reaction assay. *Lancet* 1990;1:1220.

56. Ferreira-Gonzalez A, Garrett CT. Establishing a molecular diagnostics laboratory. In: Henry JB, ed. *Clinical Diagnosis and Management by Laboratory Methods*, 19th Ed. Philadelphia: W. B. Saunders Co.

57. Ferreira-Gonzalez A, Garrett CT. Pitfalls in establishing a molecular diagnostic laboratory. *Hum Pathol* 1996;27:437–440.

58. Porter-Jordan K, Rosenberg EL, Keiser J, et al. Nested polymerase chain reaction assay for the detection of cytomegalovirus overcomes false positives due to contamination with fragmented DNA. *J Med Virol* 1990;30:35–59.

59. Udaykumar, Epstein JS, Hewlett IK. A novel method employing UNG to avoid carry-over contamination in RNA-PCR. *Nucleic Acids Res* 1993;21:3917–4853.

60. Espy MJ, Smith TF, Persing DH. Dependence of polymerase chain reaction product inactivation protocols on amplicon length and sequence composition. *J Clin Microbiol* 1993;31:2361–2365.

61. Jinno Y, Yoshiura K, Niikawa N. Use of psoralen as extinguisher of contaminated DNA in PCR. *Nucleic Acids Res* 1990;18:6739–6759.

62. Salomon RN. Introduction to reverse trasncriptase polymerase chain reaction. *Diag Mol Pathol* 1995;4:2–3.

63. Baumforth KR, Nelson PN, Digby JE, et al. Demystified… the polymerase chain reaction. *Mol Pathol* 1999;52:1–10.

64. Hagen-Mann K, Mann W. RT-PCR and alternative methods to PCR for in vitro amplification of nucleic acids. *Exp Clin Endocrinol Diabetes* 1995;103:150–155.

65. Cale JM, Shaw CE, Bird IM. Optimization of a reverse transcription-polymerase chain reaction (RT-PCR) mass assay for low-abundance mRNA. *Methods Mol Biol* 1998;105:351–371.

66. Lion T. Appropriate controls for RT-PCR. *Leukemia* 1996;10:1843.

67. Van Belkum A, Niesters HG. Nucleic acid amplification and related techniques in microbiological diagnostics and epidemiology. *Cell Mol Biol* 1995;41:615–623.

68. Romano JW, van Gemen B, Kievits T. NASBA: a novel, isothermal detection technology for qualitative and quantitative HIV-1 RNA measurements. *Clin Lab Med* 1996;16:89–103.

69. Hodinka RL. The clinical utility of viral quantitation using molecular methods. *Clin Diagn Virol*. 1998;10:25–47.

70. Nelson NC. Rapid detection of genetic mutations using the chemiluminescent hybridization protection assay (HPA): overview and comparison with other methods. *Crit Rev Clin Lab Sci*. 1998;35:369–414.

71. Busch U, Nitschko H. Methods for the differentiation of microorganisms. *J Chromatogr B Biomed Sci Appl* 1999;722:263–278.

72. Wiedmann M, Wilson WJ, Czajka J, et al. Ligase chain reaction (LCR)—overview and applications. *PCR Methods Appl* 1994;3:S51–S64.

73. Ben-Ezra JM. Amplification methods in the molecular diagnosis of genetic diseases. *Clin Lab Med*. 1995;15:795–815.

74. Lee HH. Ligase chain reaction. *Biologicals*. 1996;24:197–199.

75. Nolte FS. Branched DNA signal amplification for direct quantitation of nucleic acid sequences in clinical specimens. *Adv Clin Chem*. 1998;33:201–235.

76. Wilber JC. Branched DNA for quantification of viral load. *Immunol Invest*. 1997;26:9–13.

77. Lorincz AT. Hybrid Capture method for detection of human papillomavirus DNA in clinical specimens: a tool for clinical management of equivocal Pap smears and for population screening. *J Obstet Gynaecol Res* 1996;22:629–636.

78. Comparison of the second-generation digene hybrid capture assay with the branched-DNA assay for measurement of hepatitis B virus DNA in serum. *J Clin Microbiol* 1999;37:2461–2465.

79. Mansfield ES, Worley JM, McKenzie SE, et al. Nucleic Acid detection using non-radioactive labelling methods. *Mol Cell Probes* 1995:9:145–156.

80. Guesdon JL. Immunoenzymatic techniques applied to the specific detection of nucleic acids. A review. *J Immunol Methods*. 1992;150:33–49.

81. Oldach D. "Real-time" polymerase chain reaction. *Gastroenterology*. 1999;116:763–764.

82. Hyrup B, Nielsen PE. Peptide nucleic acids (PNA): synthesis, properties and potential applications. *Bioorg Med Chem* 1996;4:5-2

83. Alphey L. *DNA Sequencing: From Experimental Methods to Bioinformatics*. Oxford, UK: BIOS Scientific Publishers; New York: Springer, 1997.

84. Kant JA. Direct DNA sequencing in the clinical laboratory. *Clin Chem*. 1995;41:1407–1409.

85. Kretz K, Callen W, Hedden V. Cycle Sequencing. *PCR Methods Appl* 1994;3:S107–S112.

86. Nelson K, Selander RK. Analysis of genetic variation by polymerase chain reaction-based nucleotide sequencing. *Methods Enzymol* 1994;235:174–183.

87. Phear GA, Harwood J. Direct sequencing of PCR products. *Methods Mol Biol* 1994;31:247–256.

88. Slatko BE. Thermal cycle dideoxy DNA sequencing. *Methods Mol Biol* 1994;31:35–45.

89. Righetti PG, Gelfi C. Capillary electrophoresis of DNA for molecular diagnostics. *Electrophoresis* 1997;18:1709–1714.

90. Kheterpal I, Mathies RA. Capillary array electrophoresis DNA sequencing. *Anal Chem* 1999;71:31A–37A.

91. Kurian KM, Watson CJ, Wyllie AH. DNA chip technology. *J Pathol* 1999;187:267–271.

92. Hacia JG, Collins FS. Mutational analysis using oligonucleotide microarrays. *J Med Genet* 1999;36:730–736.

93. Kurian KM, Watson CJ, Wyllie AH. DNA chip. *J Pathol* 1999;187:267–271.

94. Clewley JP. Automation of the polymerase chain reaction. Part I. Prospects and problems. *Commun Dis Public Health*. 1999;2:71–72.

95. Khan J, Bittner ML, Chen Y, et al. DNA microarray technology: the anticipated impact on the study of human disease. *Biochim Biophys Acta* 1999;1423:M17–M28.

96. Lipshutz RJ, Fodor SP, Gingeras TR, et al. High density synthetic oligonucleotide arrays. *Nat Genet*. 1999;(1 Suppl):20–24.

97. Afshari CA, Nuwaysir EF, Barrett JC. Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res* 1999;59:4759–4760.

98. Granjeaud S, Bertucci F, Jordan BR. Expression profiling: DNA arrays in many guises. *Bioessays* 1999;21:781–790.

99. Kozian DH, Kirschbaum BJ. Comparative gene-expression analysis. *Trends Biotechnol* 1999;17:73–78.

100. Effenhauser CS, Bruin GJ, Paulus A. Integrated chip-based capillary electrophoresis. *Electrophoresis* 1997;12–13:2203–2213.

101. Crain PF, McCloskey JA. Applications of mass spectrometry to the characterization of oligonucleotides and nucleic acids. *Curr Opin Biotechnol* 1998;9:25–34.

102. Murray KK. DNA sequencing by mass spectrometry. *J Mass Spectrom* 1996;11:1203–1215.

8
Molecular Biology of Inherited Diseases

Jeffrey A. Kant

Genetic diseases are passed in simple and complex patterns from parents to children — except for perhaps ‘insanity.’ We now understand the simple patterns well, including underlying molecular genetic defects that cause many disorders. Inherited conditions that exhibit more complex patterns, to include multifactorial disorders, behavioral traits, and perhaps even variant expressions of well-understood single gene disorders still await resolution at the molecular level. This chapter will concentrate on disorders with simple patterns transmitted through the germline of parents. *Somatic disorders* associated with mutations that arise in cells after conception and are not transmitted reproductively will be discussed under oncologic disorders. The full elucidation of the human genome sequence through the Human Genome Project and other efforts, combined with information from sequencing projects directed at other genomes such as the mouse, promise a veritable explosion of additional clinical applications over the next decade for simple, and ultimately complex, inherited disorders.

TERMINOLOGY

Part of "8 - Molecular Biology of Inherited Diseases"

There are a number of important terms to understand in a discussion of inherited diseases. Important terms or concepts are highlighted in *italics* throughout this chapter.

The medical manifestation(s) of a disease are referred to as the disease *phenotype*. A patient with cystic fibrosis (CF) may demonstrate symptoms of pulmonary infection resulting from inspissated secretions and failure of normal respiratory tract clearance mechanisms or maldigestion and failure to thrive because of loss of pancreatic exocrine function. There also may be characteristic laboratory or cellular abnormalities associated with the basic defect in the disease, such as improper chloride transport in CF that results in an elevated sweat chloride level.

The specific molecular alterations that lead to absence of or defects in a functional gene product in a patient are referred to as the disease *genotype*. CF patients who have inherited the three base pair deletion mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) known as ΔF508 from each parent generally display pulmonary and digestive disease in early childhood. Conversely, male patients whose genotype consists of a single ΔF508 mutation paired with a missense arginine to histidine change at CFTR residue 117 (R117H) frequently do not exhibit severe manifestations of CF but may have isolated abnormalities of the genitourinary tract such as bilateral or unilateral absence of the vas deferens that are undetected until adulthood — or not at all. For reasons that are only partially understood, two patients with the same genotype may display different phenotypes or timing in demonstrating a particular phenotype.

Physicians often are confused by the terms *penetrance* and *variable expressivity*. Penetrance refers to whether manifestations that permit a diagnosis of the inherited condition are present during the lifetime of a patient who has inherited a genotype that generally leads to disease. Variable expressivity refers to phenotypic variability in a disorder in a family where patients generally possess the same genotype.

The term *mutation* is semantically tricky. This term connotes a genotypic change that produces an adverse phenotypic effect, either by itself in dominantly transmitted disorders or in combination with another mutation in recessively transmitted disorders, which will be discussed below. The human genome constantly undergoes processes that produce and enhance genetic diversity. These include gross events such as crossovers between sister chromatids in meiosis as well as changes involving the deletion, insertion, or substitution of one or a few nucleotides on all chromosomes. If molecular changes do not lead to an adverse phenotypic effect, such changes are best termed genetic *polymorphisms* (or variants) and not mutations. Another commonly used term is *allele*, which refers to a gene or chromosomal site with known genetic variants, whether polymorphic or mutant.

Homozygous is a term used to describe the genotype of individuals who carry identical sequences on both autosomes (or both X chromosomes) at the site of a known disorder; of course, most people are homozygous for a nonmutant sequence, commonly referred to as ‘normal’ or ‘wild type.’ *Heterozygous* refers to individuals who carry one mutant and one normal autosome or X chromosome. Depending on whether the disorder is dominant or recessive, heterozygous individuals may or may not demonstrate a disease phenotype. A *compound heterozygote* is an individual bearing different mutations in the same gene on the respective autosomes or X chromosomes. *Hemizygous* refers to genotypes of the X and Y chromosomes in males, where there generally is not an opportunity for a diploid gene dosage.

PATTERNS OF INHERITED DISEASES

Part of "8 - Molecular Biology of Inherited Diseases"

Recognized simple patterns of inherited diseases include *autosomal recessive*, *autosomal dominant*, *X-linked*, and *mitochondrial* forms (Table 8.1). Almost all cases of such disorders are a result of molecular defects affecting single genes. By way of background, autosomes constitute chromosome pairs 1–22; X-linked disorders involve genes on the X chromosome that generally are present in two copies in females and a single copy in males. Mitochondria have been known for some time to carry their own discrete and separately replicating circular deoriboxynucleic acid (DNA), which contains genes whose mutation can lead to mild or severe disorders. Figure 8.1 provides representative family pedigrees for each of these patterns.

TABLE 8.1. SOME COMMON INHERITED DISORDERS BY PATTERN OF INHERITANCE

Inheritance Pattern	Disorder	Gene(s)
Autosomal recessive	α thalassemia	α-globin
	ß thalassemia	ß-globin
	Cystic fibrosis	CFTR
	Gaucher disease	Glucocerebrosidase
	Hereditary hemochromatosis	HFE
	Spinal muscular atrophy (autosomal recessive)	SMN-telomeric (? Others)
	Tay-Sachs disease	Hexosaminidase A
Autosomal dominant	Charcot-Marie Tooth disease, type 1A	Peripheral myelin protein 22
	Hereditary nonpolyposis colon cancer (HNPCC)	MSH-2, MLH-1, PMS-1, PMS-2
	Hereditary pancreatitis	Cationic trypsinogen
	Hereditary breast cancer	BRCA1/BRCA2
	Huntington disease	Huntingtin
	Marfan syndrome	Fibrillin
	Myotonic dystrophy	Myotonin kinase
	Neurofibromatosis, type I	Neurofibromin
	Neurofibromatosis, type II	MERLIN
	Osteogenesis imperfecta (OI, Types I and II)	Type I Collagen, α1 or α2
X-linked recessive	Duchenne/Becker muscular dystrophy	Dystrophin
	Fragile X syndrome	FMR-I
	Glucose 6-phosphate dehydrogenase deficiency	G6PD
	Hemophilia A	Factor VIII
	Hemophilia B	Factor IX
X-linked dominant	Hypophosphatemic rickets	PHEX
Mitochondrial	Leber's hereditary optic neuropathy (LHON) MERFF, MELAS, NARP, KSS	Many to few mutations in tRNA or other genes, mitochondrial rearrangements, etc.

CFTR, cystic fibrosis transmembrane conductance regulator
MERFF, Mitochondrial enchephalopathy with ragged red muscle fibers
MELAS, Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes
NARP, Neurogenic muscle weakness, ataxia, and retinitis pigmentosa
KSS, Kearn's-Sayre Syndrome

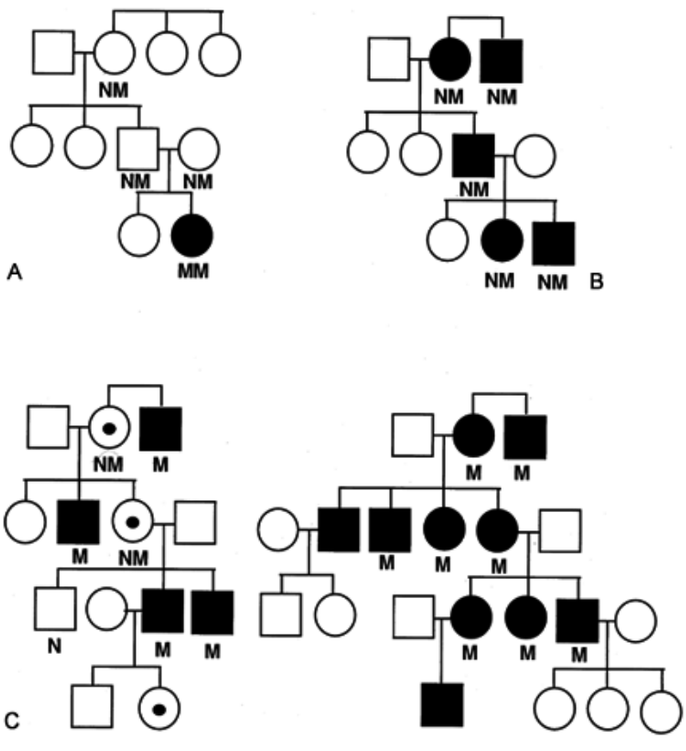


FIGURE 8.1. Major patterns of inheritance for molecular genetic disorders. (**A**), Autosomal recessive; (**B**), Autosomal dominant; (**C**), X-linked recessive; (**D**), Mitochondrial.

Autosomal recessive disorders frequently appear without warning in children of normal couples; neither parent exhibits signs of disease, and in many cases neither may be able to identify close or even distant family members with the disease that afflicts their child. This is because disease only is seen when an individual has inherited a mutant copy of the same gene from each parent (Fig. 8.1A). The presence of a single normal gene copy, as is the case for each parent, provides a sufficient level of the essential gene product to prevent the manifestations of disease. Individuals with one normal and one mutant gene copy are termed *carriers* for the particular autosomal recessive disease. The likelihood two carriers will each transmit an autosome carrying a mutant gene to conceive an affected offspring is 25% (1 in 2 from father × 1 in 2 from mother). In general, autosomal recessive disorders are

close to 100% penetrant and may show mild to modest phenotypic variability for identical genotypes.

In contrast to autosomal recessive disorders, *autosomal dominant disorders* (at least those that do not result in the death of affected individuals before they are able to reproduce) generally are seen in each generation of a family and affect both sexes (Fig. 8.1B). Both sexes also transmit the disease to offspring, an important difference from mitochondrially transmitted disorders. Because the disorder is dominant, the likelihood offspring will be affected is 50% with each pregnancy. The dominant nature of these disorders reflects a molecular mechanism in which the presence of a single mutant gene product is either sufficient to produce disease in the presence of a normal gene copy (as in a number of connective tissue disorders) or the loss of the normal gene product is sufficiently likely (as in hereditary breast cancer because of BRCA1/2 gene mutations) that the preexisting loss of a normal autosomal gene dramatically increases the patient's risk of developing cancer.

X-linked recessive disorders, a result of mutations affecting genes on the unequally distributed X chromosome, affect males but rarely affect females (Fig. 8.1C) because a normal gene on the second X chromosome is protective. Females therefore are similar to carriers of an autosomal recessive disorder, except that 50% of their male offspring demonstrate disease. Cytogenetic disorders (e.g., Turner's syndrome) that remove entire X chromosomes or portions may give rise to X-linked disorders in females, as can a marked skew in the mandatory X chromosome inactivation that occurs in each cell with more than one X chromosome. If an affected male reproduces, a mutant X chromosome may be transmitted to a daughter, but never to sons. An interesting feature of lethal X-linked recessive disorders in populations at equilibrium for the disorder is a high germline mutation rate of 33% (*Haldane's hypothesis*). A woman whose first male child demonstrates such an X-linked disorder has only a two in three risk of being a somatic carrier. Most X-linked disorders are recessive, but rare *X-linked dominant disorders* (e.g., hypophosphatemic rickets) exist with expected patterns of inheritance. Male-to-male transmission does not occur. 50% of children born to affected females are affected by disease. Affected hemizygous males show a more consistent and severe disease course, and heterozygous affected females demonstrate greater phenotypic variability.

Mitochondrially inherited disorders are from mutations associated with the 37 genes found on the 16.5 kb circular mitochondrial chromosome that include transfer and ribosomal ribonucleic acid (RNA) genes as well as several proteins involved as subunits in enzymes associated with oxidative phosphorylation. Not surprisingly, mitochondrial DNA mutations are associated with encephalopathies and myopathies; diseases of organs which show a greater dependence on oxidative phosphorylation. Because sperm contribute few mitochondria to a fertilized ovum, mitochondrial disorders are transmitted only by females but affect all offspring of both sexes (Fig. 8.1D). Mitochondria replicate their DNA independent of nuclear replication and segregate independently to daughter cells in mitosis. As a result, the percentage of mitochondria carrying a specific mitochondrial DNA mutation may vary between somatic cells and organs, a phenomenon known as *heteroplasmy*. Heteroplasmy can lead to interesting threshold effects and variabilities of disease expression in mitochondrially inherited disorders.

Occasionally, inheritance defies expected Mendelian patterns, as when two or more affected children are born in a family without prior history of an autosomal-dominant or X-linked recessive disorder. Such phenomena have been shown in many cases to arise from *germline mosaicism* where a spontaneous mutation presumably arose early enough in gametogenesis that it is present in a modest percentage of an individual's germ cells. The ability to predict risk to subsequent pregnancies in such situations is difficult to estimate, and generally is based on empiric experience, which may not be extensive.

Estimation of Simple Genetic Risks

The prospect of having an initial (or subsequent) affected child naturally is of concern in families with a history of a genetic disorder. Physicians caring for affected patients and genetic counselors frequently are asked to provide information on the risk of an individual being a disease carrier or a couple conceiving an affected child. Needless to say, risk assessment is most desirable before pregnancy occurs to aid in reproductive planning and, as we will discuss below, the opportunity for molecular genetic testing to provide improved risk information can be extremely helpful to this process.

The risk of conceiving an affected child in a nuclear family with one of the four major types of Mendelian inheritance has been discussed above, and extension of these principles to close relatives is relatively straightforward. For example, the likelihood of inheriting a mutation-bearing autosome from either parent or a mutation-bearing X chromosome from a known carrier mother is one in two. Thus, the risk of a fully penetrant autosomal dominant disorder is 50% to all children, similar to that for an X-linked recessive disorder to male children. For autosomal recessive disorders, two carrier parents have a one in four risk of conceiving an affected child. Several additional examples of simple risk assessment with explanation are provided in Figure 8.2.

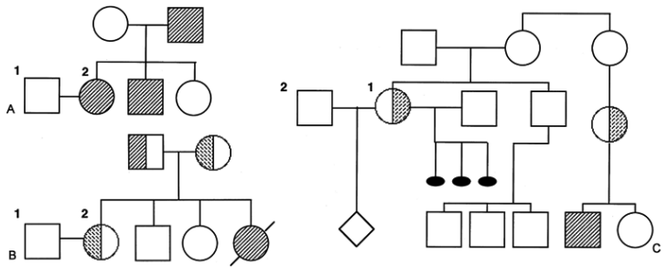


FIGURE 8.2. Estimation of simplex genetic risks. (**A**), What is the risk the couple (individuals 1 and 2) will have a child affected by the fully penetrant autosomal-dominant disorder affecting this family? The risk is 50%. The likelihood the affected mother will transmit a mutant autosome is one in two. The father contributes little or no risk because he is unaffected clinically, and the likelihood of a new mutation in his germline is much less than the genetic risk of transmission from the mother. (**B**), What is the risk the couple (individuals 1 and 2) will have a CF-affected child? Individual 2 is a known CF carrier by DNA testing. Her husband, who is of Northern European origin, has not been tested. The mixed white population carrier frequency for CF is ˜1 in 25. The wife's carrier risk is known to be 100% from prior testing. 1/25 × 1 = 1/25 is the likelihood both are CF carriers. 1/2 × 1/2 = 1/4 is the likelihood each will contribute a mutant CF chromosome if both are CF carriers. 1/25 × 1/4 = 1/100. This risk could be lowered further by molecular testing of the husband. A negative result from a molecular screening assay that has 90% sensitivity of detecting CF mutations in this population would reduce his *a priori* carrier risk from 1 in 25 to 1 in ˜240. (**C**), What is the risk that the fetus of individuals 1 and 2 will be affected by CF? The *a priori* carrier risk of individual 2 is 1 in 25. The *a priori* carrier risk of individual 1 is 1 in 8. The risk her aunt is a carrier is 1 in 2. The risk her mother is a carrier is 1 in 4. Therefore her risk is 1 in 8; because the father's carrier risk is sufficiently lower (1 in 25) than the mother's 1 in 4 it is 'ignored.' The risk individuals 1 and 2 are both CF carriers is 1 in 200 (1/8 × 1/25). If both were carriers, the risk of both transmitting a mutant CF chromosome is 1 in 4, and the risk the fetus will be CF-affected is 1 in 800 (1/200 × 1/4). Obviously, direct molecular testing would be useful in this family. In fact, individual 1 was demonstrated to be a ΔF508 mutation carrier, but despite this because individual 2 was negative, the final risk to the fetus was 1 in ˜960 (see 2B above).

Because individuals being considered for genetic risk analyses in reality do or do not carry a mutation of concern, prior reproductive experience or laboratory tests may be helpful as *conditional parameters* in modifying the basic genetic risk of having an affected child or of being a carrier for a disorder. Risk calculations that incorporate such conditional probabilities are termed *Bayesian calculations*, from Bayes theorem. Table 8.2A and Table 8.2B demonstrate risk modification calculations in X-linked recessive inheritance to assess the carrier likelihood for a woman who has two clinically unaffected sons; her mother and sister are obligate carriers for Duchenne muscular dystrophy (DMD) (Fig. 8.3). This woman's *a priori* carrier risk of 50% is lowered significantly to less than 10% by her 'favorable' birth experience with male children to date (Table 8.2A) and normal serum levels of creatine kinase, an enzyme elevated in two thirds of obligate DMD carriers but uncommonly elevated in noncarriers (Table 8.2B).

MOLECULAR DIAGNOSTICS OF INHERITED DISORDERS—GENERAL APPROACHES

TABLE 8.2. BAYESIAN RISK ANALYSIS (based on pedigree shown in Figure 8.3)

A. LIKELIHOOD WOMAN IS OR IS NOT A CARRIER FOR DMD BASED ON A *PRIORI* GENETIC RISK AND BIRTH HISTORY

Probability Term	Carrier for DMD	Not a Carrier for DMD
Prior probability (genetic) 50% risk of inheriting mutant X chromosome from mother-obligate carrier	1/2	1/2
Conditional probability	1/2 × 1/2 = 1/4	1 × 1 = 1 (100%)
Likelihood of 2 unaffected males		
Joint probability (prior × conditional)	1/2 × 1/4 = 1/8	1/2 × 1 = 1/2
Posterior probability	(1/8)/(1/8 + 1/2) = 1/5	(1/2)/(1/8 + 1/2) = 4/5
Individual probability/total		
Risk (likelihood) = posterior probability	20%	80%

B. LIKELIHOOD WOMAN IS OR IS NOT A CARRIER FOR DMD BASED ON GENETIC RISK, BIRTH HISTORY, AND SERUM CREATINE KINASE LEVELS (SEE PRIOR RISK CALCULATIONS)

Probability Term	Carrier for DMD	Not a Carrier for DMD
Prior probability (genetic risk + birth history), from Table 8.2A	1/5	4/5
Conditional probability	1/3	19/20 (95%)
Normal serum creatine kinase × 2		
Joint probability (prior × conditional)	1/5 × 1/3 = 1/15	4/5 × 19/20 = 76/100
Posterior probability	(1/15)/(1/15 + 76/100)	(76/100)/(1/15 + 76/100)
Individual probability/total		
Risk (likelihood) = posterior probability	8%	92%

Abbreviations: DMD, Duchenne Muscular Dystrophy

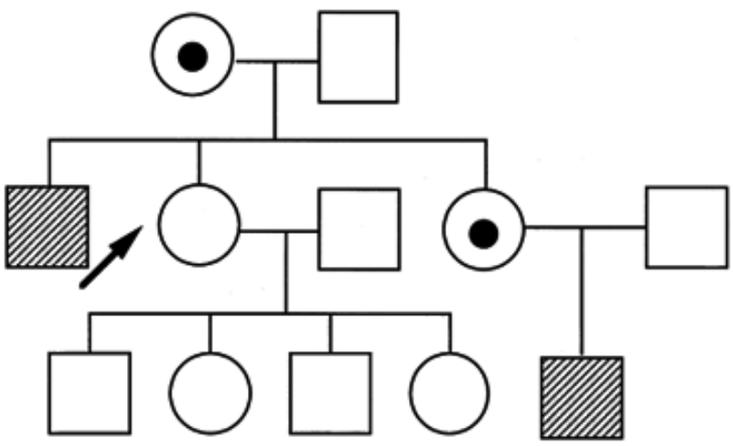


FIGURE 8.3. Family with history of Duchenne muscular dystrophy (DMD). See also Table 8.2A and Table 8.2B.

Molecular Changes in Genes: Naming Conventions

Common conventions for naming mutations and polymorphisms are discussed below and additional examples are provided in Table 8.3, using examples from the CFTR protein. Most of these focus on single nucleotide changes and small insertions or deletions rather than large-scale deletion or duplication of exons or whole genes. Proteins are composed of amino acids numbered 1 to N from the amino terminus. A glossary of single-letter amino acid codes is used in conjunction with the numerical position of an altered amino acid to describe point mutations resulting from single nucleotide substitutions; for example, R117H indicates substitution of histidine (H) for the normally occurring arginine (R) at amino acid 117 of CFTR. Mutations where nucleotide substitution generates a translation stop codon are coded as 'X,' for example, G542X–STOP for glycine at amino acid 542. Mutations in introns, which generally occur at exon-intron junctions and compromise normal messenger RNA (mRNA) splicing, are designated by a minus (-) or plus (+) sign to indicate the number of nucleotides the mutation follows or precedes the exon-intron junction, which is indicated by the number of the nucleotide in mature mRNA that starts or ends the exon, for example, 621+1G>T describes a substitution of the invariant first guanine of intron 4 of CFTR by thymidine. The term Δ (delta) indicates loss (deletion) of an amino acid, hence ΔF508 means the loss of the amino acid phenylalanine at residue 508 of CFTR from a three nucleotide deletion in the gene and subsequent mRNA.

TABLE 8.3. A RANGE OF MUTATION TYPES IN THE CFTR GENE

Mutation	Type	Description
R117H, G551D	Point (missense)	Amino acid codon to new amino acid codon in exons 4 and 11; R117H is phenotypically mild, G551D severe
G542X, W1282X	Point (nonsense)	Amino acid codon to stop codon in exons 11 and 20
ΔF508	Deletion (non-frameshift)	Loss of amino acid phenylalanine at CFTR residue 508
3659del C*	Deletion (frameshift)	Amino acid codon to frameshift in exon 19
3100insA*	Insertion (frameshift)	Amino acid codon to frameshift in exon 16
621+ 1 G > T*	Splice junction	G > T at first nucleotide of intron 4
3849+ 10kbG > T*	Splice junction	Mutation at considerable distance (10 kb) from splice junction activates cryptic splicing in intron 19
CF50kbdel#1	Large deletion	Complex deletion involving exons 4–7 and 11–18.

* For these mutations, the numbers used are nucleotides numbered from the 5′ end of mature mRNA, not amino acid numbers; 'silent' nucleotide polymorphisms that do not lead to a change of amino acid are also numbered using nucleotide numbers.
CFTR, cystic fibrosis transmembrane conductance regulator.

Techniques for the Study and Clinical Diagnosis of Inherited Diseases

Refer to the introductory chapter in this section for additional information on specific molecular techniques. The major molecular biologic technique employed today in clinical testing for inherited diseases is polymerase chain reaction (PCR) amplification, commonly followed by analysis on electrophoretic gels with or without prior restriction endonuclease digestion. However, different approaches to amplification of sample DNA or methods that do not involve amplification of the target sequence, as in PCR, are being used increasingly in the clinical laboratory. Southern blot hybridization also is useful for the analysis of certain genetic disorders, generally where larger regions of DNA are being interrogated. Although time-consuming and technically demanding, direct nucleic acid sequencing following PCR is being used more often in the clinical laboratory for inherited diseases.

Inherited disorders are studied for the most part using two approaches: an indirect approach known as *linkage analysis*, which has become less common as the molecular pathology of various disorders is elucidated, and *direct identification* of specific molecular changes associated with disease. Direct identification has a number of advantages over linkage analysis, particularly the ability to conduct definitive testing (within the sensitivity limitations of the test) on individual adult and prenatal samples usually without the need for samples from other family members. The ability to directly identify common mutations in common disorders also has opened up the possibility of widespread population-based genetic screening.

Linkage Analysis and Polymorphisms

The term polymorphism used earlier in this chapter is an essential element of linkage analysis. A polymorphism is a chromosomal site that shows variability among individuals, usually an alteration of one to a few nucleotides. With assays that detect that variability, it is possible to track individual chromosomes through a family pedigree. A requirement for useful tracking is that polymorphisms be reasonably 'near' (or linked with) a disease locus on the chromosome so they are unlikely to be affected by crossover between sister chromatids.

Restriction fragment length polymorphisms (*RFLPs*) were the first molecular polymorphisms identified. Nucleotide changes in the genome were found to either abolish or create recognition sequences for restriction endonucleases, enzymes that cut double-stranded DNA at precise nucleotide sequences, usually four to six nucleotides in length. In fact, as we shall see, RFLP analysis also is a useful assay for disease-causing mutations. A problem with RFLPs is they are relatively uninformative, rarely allowing distinction of chromosomes in more than 50% of individuals in a population, and often far fewer.

Di-, tri- and tetranucleotide repeat polymorphisms are particularly useful and have largely supplanted RFLPs for linkage analysis. These typically are termed *microsatellite* polymorphisms, and are scattered about the genome, often in introns, with size variations resulting from the number of repeating motifs (e.g., $CA_{(n)}$ in the case of dinucleotide repeats). These markers are far more variable, and thus informative, than RFLPs, frequently demonstrating a dozen or more distinct alleles. A third type of polymorphism, based on *variable numbers of tandem nucleotide repeats* (*VNTR*) of larger size, also is highly informative but less common than microsatellite polymorphisms. Highly informative polymorphisms with large numbers of alleles have wide applications today in identifying relationships among individuals or in linking individuals to criminal evidence in forensics.

Figure 8.4 demonstrates the use of microsatellite polymorphic markers to track the normal and mutation-bearing chromosomes in an autosomal recessive disorder spinal muscular atrophy (SMA). Here there are polymorphic markers on both the centromeric and telomeric sides in this disease locus, an important advantage if informative because recombination can be identified except in rare instances of double recombination. The likelihood a patient carries a disease-causing chromosome, or an affected individual carries two, can be calculated based on whether there are informative markers, their distance from the disease locus, and whether informative markers are present on both sides of the disease locus. A low or high likelihood usually can be intuited from the inheritance pattern; precise risk calculations are best performed by computer programs or qualified genetics professionals.

Polymorphic markers, within and outside the gene, also can be useful secondary analysis tools in diseases where the causative gene and specific mutations have been identified, when direct identification of 100% of mutations is not possible. This includes many patients with Duchenne muscular dystrophy, hemophilia A and B, neurofibromatosis-1, and other disorders.

Direct Mutation Identification

Most inherited disorders now can be examined, at least in a primary mode, for specific molecular alterations that are characteristic of a disease. In some cases, specific mutations are associated with specific effects (genotype-phenotype correlation) although mutations more often convey a general (e.g., severe versus mild) impact on phenotype.

Many different types of mutations have been delineated. Single '*point*' mutations (usually in exons, but sometimes in introns or regulatory/other control regions) are most common. Other types include small *insertions* or *deletions* of one to several nucleotides that shift the 'reading frame' for protein translation, and larger-scale deletions (or insertions) involving portions of, or even whole, genes. An interesting class of mutations responsible for a variety of neurodegenerative disorders is expansions of microsatellite *trinucleotide repeats* within or near certain genes. Another interesting group are disorders of *imprinted genes* in which mutation of a locus normally active only on the chromosome inherited from mother or father can have significant phenotypic effects if inherited from the 'wrong' parent. Table 8.4 presents an overview of the common types of mutations and some disorders associated with them, several of which are discussed in greater detail below. The reader will notice that some disorders (e.g., DMD, hemophilia A) have multiple types of mutations, and more than one type of laboratory assay may need to be employed to screen for the full spectrum of mutations.

TABLE 8.4. COMMON DISORDERS ASSOCIATED WITH DIFFERENT TYPES OF GENETIC MUTATIONS

Type of Mutation	Disease	Genes Affected
'Point' mutations (to include small insertions or deletions of one or more nucleotides)	α thalassemia	α-globin
	ß thalassemia	ß-globin
	Cystic fibrosis	CFTR
	Duchenne/Becker muscular dystrophy	Dystrophin
	Gaucher disease	Glucocerebrosidase
	Hemophillia A	Factor VIII
	Hereditary hemochromatosis	HFE
	Hypercoagulability – factor V Leiden	Factor V
	Hypercoagulability – factor II	Prothrombin (Factor II)
	Neurofibromatosis, type I	Neurofibromin
	Sickle cell anemia	ß-globin
Splice junction mutations	ß thalassemia	ß-globin
	Cystic fibrosis	CFTR
Intragenic inversion	Hemophilia A	Factor VIII
Partial or full gene deletion	α thalassemia	α-globin
	ß thalassemia	ß-globin
	AR spinal muscular atrophy (SMA)	Survival motor neuron (SMN)-telomeric
	Duchenne/Becker muscular dystrophy	Dystrophin
	Hemophilia A	Factor VIII
	Neurofibromatosis, type I	Neurofibromin
Partial or full gene duplication	Charcot-Marie Tooth disease, type 1A	Peripheral myelin protein 22 (PMP22)
	Duchenne/Becker muscular dystrophy	Dystrophin
Trinucleotide repeat expansion	Dentatorubro-pallidoluysian atrophy (DRPLA)	Atrophin (CAG)**
	Fragile X syndrome	FMR-I (CGG)
	Friedreich's ataxia	Frataxin (GAA)
	Huntington disease	Huntingtin (CAG)
	Myotonic dystrophy	Myotonin kinase (CTG)
	Spinocerebellar ataxia, type I	Ataxin (CAG)
	X-linked SMA (Kennedy's disease)	Androgen receptor (CAG)
Hybrid genes due to chromosomal crossover	ß thalassemia	ß and γ globin
Imprinted genes*	Prader-Willi syndrome (PWS), Maternal imprint	? several
	Angelman's syndrome (AS), Paternal imprint	? several
	Beckwith-Wiedmann syndrome	unknown

* Imprinting is believed to be associated with methylation and lack of expression of a critical gene(s); loss of a functional paternal (PWS) or maternal gene (AS) by deletion, uniparental disomy, abnormal methylation, or possibly point mutation, leads to disease.
** Repeated trinucleotide given in parentheses
AR, autosomal recessive
CFTR, cystic fibrosis transmembrane conductance regulator

Because molecular abnormalities can generally be assayed precisely, assays for individual mutations usually show 100% sensitivity and specificity. However, for disorders with many mutations, clinical laboratory assays must select a subset of mutations; thus, the detection sensitivity of the overall assay will be less than 100%. Apart from procedural errors in assays such as a specimen switch, incorrect assay conditions, or reagents, there are few reasons for false negative or false positive results in molecular assays. Unknown polymorphisms that alter nucleotides at the 3′ end of PCR amplification primers may lead to poor or absent amplification of an allele with loss of information about that allele. In theory, contamination by amplicons from another positive sample

could lead to a false-positive result. Prenatal samples, particularly chorionic villus samples, must be dissected free of maternal tissue, lest a chimeric genotype be obtained, which can be catastrophic for proper interpretation.

MOLECULAR DIAGNOSTICS OF INHERITED DISORDERS—SPECIFIC DISORDERS

Part of "8 - Molecular Biology of Inherited Diseases"

The current application of molecular approaches to specific disorders is presented; most applications involve direct mutation detection.

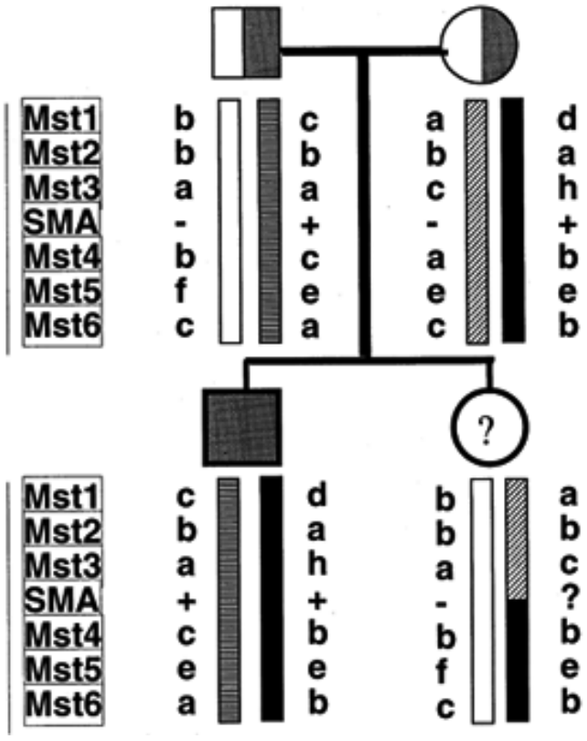


FIGURE 8.4. Linkage analysis using microsatellite polymorphism markers. Linkage analysis was performed in this family with autosomal recessive spinal muscular atrophy (SMA) to determine if the fetus (a daughter) of 16 weeks gestation has the same chromosomal pattern as the affected son. A list of six microsatellite polymorphisms in the region of the SMA locus is indicated; SMA indicates the locus on chromosome 5q responsible for this disorder; Mst1 is furthest away from SMA on the centromeric side, Mst6 furthest on the telomeric side. Alleles a through h are indicated based on relative size for each Mst locus and arranged in haplotypes corresponding with the mutant (+) and nonmutant (-) SMA chromosomes. Linkage analysis was performed for this family before more specific direct analysis for SMA became available. The fetus has inherited a recombinant chromosome from her mother. Whether this chromosome carries a mutant locus cannot be determined because recombination occurred between Mst3 and Mst4, which encompasses the disease-causing region. Fortunately, the fetus has inherited the nonmutant chromosome from her father, and has low likelihood of demonstrating SMA.

Cystic Fibrosis

Cystic fibrosis (CF) is a prototypical autosomal-recessive disease with a carrier frequency of ˜1 in 25 to 30 for mixed whites, but rarer in Hispanics, African-Americans, and Asians. Roughly one in 2,500 births are children affected with CF. A three nucleotide deletion that removes sequences encoding a phenylalanine residue at position 508 of the CFTR protein, known as the ΔF508 mutation, comprises ˜70% of mutant CF chromosomes in the United States. Unfortunately, there are over 700 other CF mutations, none more than 5% prevalent in the general U.S. population, although certain mutations are more common in certain ethnic groups, such as the W1282X, which comprises ˜40% of CF mutations in individuals of Ashkenazic Jewish background. The number of mutations screened currently in clinical assays ranges as high as 90, but many laboratories test for a core of 15 to 30 mutations that yield assay sensitivity in the low 90% range for individuals of mixed Caucasian origin. A proposed uniform panel of ˜35 mutations soon will be recommended for population-based screening in the United States. Panels of 90% sensitivity should identify mutations in both individuals from slightly more than 80% of couples who present for screening. For families in whom only one of two causative

mutations can be identified, a variety of flanking and intragenic polymorphisms are useful.

Duchenne/Becker Muscular Dystrophy (DMD/BMD)

Approximately two thirds of all Duchenne/Becker Muscular Dystrophy (DMD/BMD) patients show deletion (usually) or duplication of a moderately large region (thousands to millions of nucleotides) of the dystrophin gene on the short arm of the X chromosome. The other one third appear to have mutations involving one or a few nucleotides that are extremely difficult to find in the dystrophin gene, the largest of all human genes with 79 exons spanning two megabases of DNA. Mutations produce two clinical phenotypes: a severe illness in which patients are wheelchair-bound early in life (DMD) and a milder form where there is much less muscle weakness (BMD). Partial gene deletions or duplications that cause DMD juxtapose exons in a way that disrupts the normal translation reading frame of dystrophin mRNA while deletions or duplications causing BMD, even though they may be quite extensive, retain the reading frame (Figure 8.5).

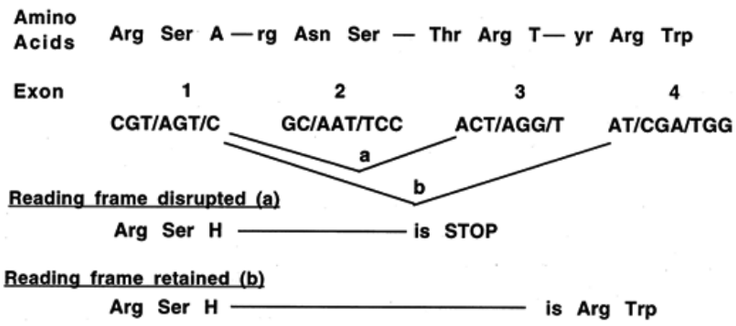


FIGURE 8.5. Effects of partial gene deletion on gene product produced. Four exons (1–4) are shown for a short hypothetical gene. Two examples of partial gene deletion are shown, one (a) that deletes exon 2, and the other (b) that deletes exons 2 and 3. Amino acids coded by the full-length mRNA from exons 1–4 are given above the nucleotide sequence. Amino acids encoded by the mRNAs resulting from partial gene deletions (a) and (b) are given below. Note that deletion (a) disrupts the normal translation reading frame by juxtaposing the final nucleotide of exon 1 (the first in a codon) next to the first nucleotide of exon 3 (also the first in a codon); this quickly leads to a translation STOP codon. Deletion (b), by contrast, which juxtaposes the final nucleotide of exon 1 next to the first of exon 4 (the 2nd in a codon) maintains the translation reading frame, resulting in a protein missing several amino acids and with one new amino acid (His) but otherwise identical in sequence to the normal protein.

Diagnostic approaches to DMD/BMD generally first look for partial gene deletion by screening genomic DNA in males for the presence or absence of specific regions of the dystrophin gene. A demanding but commonly used assay combines two groups of nine pairs of PCR primers in multiplex assays for 18 regions of the dystrophin gene. The loss of regional signals in males confirms partial dystrophin gene deletion and a diagnosis of DMD/BMD. Linkage analysis usually is used next for families in which gene deletion cannot be demonstrated or affected males are deceased. Finally, Southern blot analysis using complimentary DNA (cDNA) probes from the dystrophin gene may be employed to look for deficiencies or excesses of exon dosage to identify individuals with duplications as well as to identify carrier females in families with known (and sometimes unknown) deletions.

Disorders Associated with Trinucleotide Repeat Expansions

Expansions of trinucleotide microsatellite sequences, usually in the nucleotide format CXG, where X is any nucleotide, are unique from single nucleotide changes or gene deletions that predominate in many genetic disorders. Although these mutations are transmitted through the germline, they may undergo further expansion (and rarely contraction) in somatic or germ cells postfertilization. Trinucleotide repeat disorders for the most part reflect large-scale or small-scale expansions.

A good example of a large-scale trinucleotide expansion is *myotonic dystrophy* (DM), which affects young to middle-aged adults. The disease is autosomal dominant with variable expressivity, and within individual families exhibits a phenomenon termed *anticipation*, where the age of onset decreases and disease severity increases in successive generations. The molecular cause of DM and the basis for anticipation involves variable expansion of a CTG repeat in the 3′ untranslated region of the myotonin kinase gene on chromosome 19. Normal individuals show 5 to 30 CTG repeats, while adult DM patients exhibit several hundred to over one thousand repeats. Children of affected adults who exhibit the severe congenital form of DM have CTG repeats of 2,000 or more. Transmitting parents of adult patients generally show a small expansion of this region, usually 40 to 150 repeats, called a premutation. These individuals demonstrate few or no symptoms and require electromyographic studies to disclose myotonia.

A similar defect involving variable amplification of a CGG triplet region just 5′ of the FMR gene on the long arm of the X chromosome characterizes the *fragile X syndrome*, a common cause of inherited mental retardation in males. Normal individuals show up to ˜50 repeats. Affected patients demonstrate several hundred to several thousand repeats, with disease severity correlating roughly to repeat size. The chromosomal region surrounding the FMR gene becomes highly methylated in affected males, presumably silencing expression of the FMR gene. Males with mild symptoms but expansions expected to cause disease may demonstrate lower degrees of methylation. As in DM, unaffected females and clinically unaffected males may transmit the disease through premutations with small to modest CGG expansions of 50 to 200 repeats. The likelihood that a premutation allele will expand increases significantly in relation to size in the range from 60 to 100 CGG repeats. A few point mutations of the FMR gene also have been identified to cause the fragile X syndrome.

Several other disorders exhibit trinucleotide repeat expansions of a more limited size (Table 8.5). These disorders usually

are inherited in a dominant fashion for autosomal loci, and recessively for X-linked loci. As an example, in Huntington disease, there is expansion of a CAG triplet in the 1st exon of the Huntingtin gene on the short arm of chromosome 4. The CAG region encodes a polyglutamine stretch in the resultant Huntingtin protein, which, when expanded, is thought to have a toxic effect on neurons. Affected individuals demonstrate 36 to as many as 100 or more repeats. The number of repeats correlates with age of onset as well as disease severity. For many of these disorders, there still is uncertainty about the phenotypic effects or likelihood of expansion in transmission to offspring of rare alleles that fall between documented normal or expanded ranges.

TABLE 8.5. REPRESENTATIVE TRINUCLEOTIDE REPEAT DISORDERS

Disorder	Repeat	Normal^	Premutation^#	Affected^
DRPLA	CAG	<34	N/A	48–83
Huntington Disease	CAG	<35	N/A	36–125
Spinocerebellar ataxia, type 1	CAG	<36	37–44#	>45
Spinocerebellar ataxia, type 2	CAG	<37	32–34#	>35
Spinocerebellar ataxia, type 3	CAG	<36	37–60#	>61
Spinocerebellar ataxia, type 6	CAG	<16	17–20#	>21
Spinocerebellar ataxia, type 7	CAG	<17	18–37#	>38
X-linked SMA**	CAG	<30	N/A	36–62
Fragile X syndrome	CGG	<50	50–200	>200
Myotonic dystrophy	CTG	<37	40–150	>150
Spinocerebellar ataxia, type 8	CTG	<91	N/A	100–160
Freidrich's ataxia	GAA	<33	N/A	>66

** Dentato-rubro-pallidoluysian atrophy
* Also known as Kennedy's disease
# , Inconclusive range
N/A, Not applicable or uncertain
^ , Ranges may shift slightly as more information is collected; ranges given are intended to be representative

Hereditary Cancer Syndromes

It is clear that some portion of a number of malignant neoplasms are caused by hereditary alterations in specific genes. These include hereditary breast cancer (BRCA1, BRCA2, and ATM genes) malignant melanoma (p16 gene), hereditary colon cancer (APC as well as several DNA mismatch repair genes), retinoblastoma (rb gene), and more complex combinations such as the Li-Fraumeni syndrome (p53 tumor suppressor gene) with tumors of multiple organ systems.

About 5% of breast cancer in the United States is attributed to autosomal dominant inheritance of two genes: BRCA1 on chromosome 17 and BRCA2 on chromosome 13. Both are complex with >20 exons, large mRNAs, and many point mutations scattered the entire length of each gene. For a laboratory to offer ‘full service’ BRCA1/2 testing DNA sequencing is required, although some partial gene deletions will not be identified. Because of patent limitations (see below), full sequencing of BRCA1/2 is an option only through a single provider. However, laboratories may screen (with a license) for selected groups of BRCA1/2 mutations. Most commonly, a group (del185AG and 5382insC in BRCA1; 6174delT in BRCA2) is screened in individuals of Ashkenazic Jewish background, where the cumulative carrier frequency for these is approximately 1 woman in 40.

The interpretation of DNA sequencing results of BRCA1/2 offers an excellent example of the need to assemble databases associating nucleotide changes with disease for hereditary cancers. In addition to many disease-associated mutations, many clear-cut or likely polymorphisms also have been identified in BRCA1/2. An instructive example is a STOP change found at the 3′ end of the BRCA2 gene, which nevertheless turns out to be a polymorphism; presumably it occurs sufficiently carboxy in the BRCA2 polypeptide that it does not adversely affect protein function. An extremely important aspect of counseling patients in breast cancer families is to make sure that women who appear to have *not* inherited a mutant BRCA1/2 locus understand they still have the general population risk (˜10%) of developing breast cancer from other causes.

Studies of the APC gene in hereditary polyposis patients as well as several DNA mismatch repair genes (usually MSH2 and MLH1) associated with hereditary nonpolyposis colonic cancer (HNPPC) also have revealed multiple disease-causing mutations. By contrast, the c-ret oncogene associated with familial multiple endocrine neoplasia, type 2 syndromes with cancers of the thyroid and other organs shows a more restricted range of mutations confined to several exons.

Hereditary Hemochromatosis

Hereditary hemochromatosis (HH) is the most common autosomal recessive disorder in whites with a carrier frequency of 1 in 8 to 1 in 10. Affected individuals demonstrate significant iron overload with deposition in a number of organs leading to dysfunction and failure, particularly in the cardiac and endocrine systems. Two genetic variations of the HFE gene, which is closely linked to the HLA locus on chromosome 6, are most commonly associated with the molecular basis of this disorder.

The first is a single nucleotide change altering residue 282 from cysteine to tyrosine (C282Y), which interferes with the binding of HFE to transferrin receptors usually leading to increased iron uptake. This mutation is found in the homozygous state in 65% to 100% of affected individuals depending on the study. HH appears to be partially penetrant; a modest number of 282Y homozygotes do not have clinical or biochemical evidence

of disease, so there are almost certainly other (likely genetic) factors that affect iron metabolism.

The second variation is a point mutation that changes histidine to aspartic acid at HFE residue 63 (H63D). This change is in linkage disequilibrium with the 282Y mutation — they are not associated. An interesting feature of compound heterozygotes (one copy each) for the 282Y and 63D variants is an increase in serum iron concentration and an excess of HH disease over that seen with other combinations of alleles including heterozygous carriers of 282Y.

The heterozygous state of HH may not be entirely benign, as studies have suggested an increased risk of cardiac and cerebrovascular disease in individuals heterozygous for the C282Y mutation including acute myocardial infarcts and cerebrovascular disease. There also are familial forms clearly unlinked to HFE, so the genetics of iron overload appears to be subject to a variety of influences.

Mutations Associated with Thrombosis

Defects in the protein C anticoagulant pathway account for most currently known genetic risk factors for venous thrombosis including deficiencies in protein C or S. A poor anticoagulant response to activated protein C (APC), termed APC resistance, is detected in 20% to 60% of venous thrombosis patients. The genetic basis of APC resistance usually is a point mutation in the gene for coagulation factor V, which results in the substitution of glutamine for arginine (R506Q), also known as Factor V Leiden. The mutation reduces the ability of APC to inactivate Factor V, resulting in a more thrombophilic state.

The Factor V R506Q mutation can be assayed using a variety of techniques to detect point mutations and may be present in heterozygous or homozygous states. The heterozygous state is relatively common (3% to 5% of the U.S. population) and is associated with a 5- to 10-fold increased risk of venous thrombosis. Approximately 20% of individuals presenting with a first deep venous thrombosis, and 60% of patients with recurrent venous thromboses or thromboembolic events, are Factor V Leiden heterozygotes. Heterozygous women using oral contraceptives have a 20- to 30-fold risk of thrombosis compared to women with normal genotypes who are not taking contraceptives. Moreover, women who experience second trimester spontaneous abortions, young women with myocardial infarcts, and women with complications of pregnancy such as preeclampsia, all appear to carry Factor V Leiden at a rate above the background population frequency. One half of offspring of heterozygous carriers would be expected to share an increased thrombosis risk. Thus, testing raises genetic counseling issues for tested individuals and their close relatives. The homozygous state is uncommon but associated with a much increased relative risk of venous thrombosis.

Direct mutation assays to determine Factor V Leiden genotype offer several advantages over the functional APC resistance assay: they are not compromised by anticoagulant therapy, are unaffected by lupus anticoagulants, can differentiate between heterozygous and homozygous individuals, and are unambiguous without significance thresholds as may occur with APC resistance assays. Of individuals classified as activated protein C resistant, approximately 94% are Factor V Leiden heterozygotes.

A G>A nucleotide change in the 3′-untranslated region of the prothrombin gene may result in elevated prothrombin levels and a several-fold increase in risk of venous thrombosis. 1% to 2% of the general population appears to be heterozygous for this mutation. Like Factor V Leiden, the prothrombin mutation appears to carry additional risk of venous thrombosis in women in states mimicking pregnancy, and thrombosis risk is additive for individuals heterozygous for both Factor V Leiden and prothombin 20210A.

Finally, a missense C>T change at nucleotide 677 of the methylene tetrahydrofolate reductase gene (MTHFR) can lead to decreased levels of MTHFR and increased serum homocysteine in individuals homozygous for the thermolabile 677T variant. Increased serum homocysteine has been shown by several studies to be a risk factor for coronary artery disease, and because the MTHFR 677T variant is found in 5% of the population, this could be an underlying genetic contributor to cardiovascular disease, one that may be manageable by diet because folate therapy often is helpful in suppressing serum homocysteine.

Genetic Alleles Associated with Risk for Specific Diseases

This is an expanding group of disorders, likely to grow significantly as the genetics of more complex disorders are better understood. An increased risk for developing a variety of diseases has been associated with histocompatibility genes on chromosome 6; these include ankylosing spondylitis (B27) and juvenile diabetes mellitus (DR3, DR4). Because many of these disorders involve immune-mediated mechanisms, such associations are understandable.

Individuals heterozygous or homozygous for the e4 allele of the apolipoprotein E gene show ˜3- and 8-fold increased risks of developing adult-onset Alzheimer's disease compared to individuals without this allele, while the presence of an e2 allele may be protective. While this is not useful to clinically unaffected individuals, the sensitivity and specificity of testing for the ApoEe4 allele in patients with early memory loss and features of Alzheimer's disease appears to be a useful diagnostic adjunct.

Applications of Genes to Confirm Individual Identity

As previously discussed, highly polymorphic loci, particularly when used in combination, serve as excellent individual identifiers to suggest a high probability of relatedness or definitive nonrelatedness when two sources of DNA are compared. Common uses include exclusion or inclusion for paternity or criminal investigation, familial association for immigration, and resolution of potential sample mix-ups in pathology laboratories. Identifying alleles also are essential in selecting optimal donors for bone marrow or other transplants where superiority of molecular over serologic typing methods clearly is established, as well as in the measurement of engraftment and relapse in allogeneic bone marrow transplants. A variety of approaches including sequence-specific amplification, dot blots, reverse dot blots, and direct DNA sequencing are employed.

BUSINESS, REGULATORY, AND ETHICAL ISSUES ASSOCIATED WITH TESTING FOR INHERITED DISORDERS

Patents

The granting of *patents* by the United States Patent and Trademark Office for molecular alterations that cause inherited diseases as well as methods to detect them, and the licensing and use of these patents has affected the way diagnostic services are offered for a number of these disorders. The broadest example is the patent for PCR, which is licensed broadly for clinical use; most laboratories pay a minimum 9% royalty on revenue collected for tests using this technique. Ironically, many patents for specific genetic disorders are owned by university medical centers that may license the patent exclusively to a single (generally commercial) laboratory thereby preventing other academic laboratories from developing testing. Examples of tests for which patents have been enforced include: Charcot-Marie-Tooth disease, type 1A; hereditary hemochromatosis, several of the spinocerebellar ataxias, Friedrich's ataxia, and hereditary breast cancer genes BRCA1 and BRCA2.

Regulation of Genetic Testing

The longer-term status of *regulatory oversight of genetic testing* is unclear at the moment. Virtually all molecular genetics tests currently are regulated by the Food and Drug Administration (FDA) under their lowest level of scrutiny (class I), which is applied to ‘home-brew’ tests developed and validated in individual laboratories. Some have called for stratification of genetics tests into several categories as well as prospective evaluation and validation of individual tests for inherited disorders by the FDA (or designates) to establish clinical utility before tests can be performed for reimbursement by Medicare and other third-party payors.

Informed Consent, Genetic Counseling, and Genetic Discrimination

It is common for laboratory physicians to provide information before and after testing about assays for inherited disorders including limitations. This is important information for physicians or other health-care personnel who obtain informed consent from the patient for testing by explaining the laboratory testing, the genetics of the disorder, its natural history including possible variable expressivity, treatments available for affected individuals, physical and psychological effects others have experienced in caring for a patient with a disorder, and other risks such as discrimination based on genetic information. Health-care professionals also communicate the results to patients following testing. The entire consent and disclosure process is complicated with pitfalls for the uninitiated. Issues such as nondirective counseling of the patient and other family members, addressing possible revelation of nonpaternity, and privacy issues can be mismanaged by those with little experience, and the use of a genetics counselor or medical geneticist sometimes may need to be strongly recommended by the laboratory physician. The definition of a ‘genetic test’ has been somewhat controversial, and there is wide variability currently in whether formal informed consent is obtained and for which disorders. Patients undergoing presymptomatic testing for Huntington's disease are invariably consented, whereas it is uncommon to see informed consent obtained for mutations associated with hypercoagulable states (Factor V Leiden, Prothrombin 20210A, and thermolabile MTHFR). It is likely in the future that laboratories performing genetic testing will be required to demonstrate that informed consent has been obtained, in at least an agreed-upon subset of genetic disorders.

Requests for certain genetic tests also may place laboratory directors in conflict with clinicians or parents. It is generally agreed that *testing of children for genetic disorders* is not indicated unless a therapeutic intervention is anticipated. If information is not essential to the child's health, it is argued the individual at risk should choose to know or not know this information (with informed consent) after becoming of legal age. There also is concern that children may become stigmatized by parents or others if a result is positive. A host of other difficult ethical issues may be encountered, and the involvement of a genetics professional or bioethicist may be warranted.

At this time, there is not universal statutory protection from discrimination in provision of health insurance or employment based on genetic information, although bills have been discussed in Congress for several years, and there are limited protections that have been enacted for individuals in certain types of health plans. Several states have passed more comprehensive bills. Although there are a few horrific anecdotes, the few studies that have looked recently at the issue of *discrimination based on genetic information* have not indicated this is happening to any significant degree. Genetic testing enjoys the same privacy and confidentiality afforded other types of medical information in the United States, including the expansions required under recently enacted Health Insurance Portability Assurance Act (HIPAA).

SUGGESTED READING

Books and Internet Sites

Brown DL, Connor JM, Pyeritz RE. Emery and Rimoin's Principles and Practice of Medical Genetics, 3rd ed. Churchill Livingstone, 1997.

Coleman WB, Tsongalis GJ. Molecular Diagnostics for the Clinical Laboratorian, Humana Press, 1997.

Connor M, Ferguson-Smith M. Essential Medical Genetics, 5th ed. Blackwell Science Publications, 1997.

Gelehrter TD, Collins FS, Ginsburg D. Principles of Medical Genetics 2nd ed. Williams & Wilkins, 1998.

Harper PS. Practical Genetic Counseling. 5th ed. Oxford. Butterworth Heinemann, 1998.

Jameson JL, ed. Principles of Molecular Medicine, Humana Press, 1998.

Strachan T, Read AP. Human Molecular Genetics 2nd ed. Wiley-Liss, 1999.

http://www.faseb.org/genetics/ashg/policy/pol-00.htm. Policy statements on contemporary genetics issues from the American Society of Human Genetics.

http://www.faseb.org/genetics/acmg/pol-menu.htm. Policy statements on contemporary genetics issues from the American College of Medical Genetics

http://www.ncbi.nlm.nih.gov/Omim/ McKusick V. Mendelian Inheritance in Man, web version (OMIM), the definitive source for information on genetic syndromes and related information.

http://www.nhgri.nih.gov/ National Human Genome Research Institute (National Institutes of Health).

http://www4.od.nih.gov/oba/sacgt.htm. Secretary's Advisory Committee on Genetic Testing (SACGT).

Medical Literature

Brinkman RR, Mezei MM, Theilmann J, et al. The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. *Am J Hum Genet*, 1997;60:1202–1210.

Collins FS. Shattuck lecture — Medical and societal consequences of the Human Genome Project. *New Eng J Med* 1999;341:28–37.

Cummings CJ, Zoghbi HY. Fourteen and counting: unraveling trinucleotide repeat diseases. *Human Molecular Genetics* 2000;9:909–916.

DeStefano V, Martinelli I, Mannucci PM, et al. The risk of recurrent deep venous thrombosis among heterozygous carriers of both factor V Leiden and the G20210A Prothrombin mutation. *New Eng J Med* 1999;341:801–806.

Fodor FH, Weston A, Bleiweiss IJ, et al. Frequency and carrier risk associated with common BRCA1 and BRCA2 mutations in Ashkenazi Jewish breast cancer patients. *Am J Hum Genet* 1998;63:45–51.

Gudnason V, Stansbie D, Scott J, et al. C677T (thermolabile alanine/valine) polymorphism in methylenetetrahydrofolate reductase (MTHFR): its frequency and impact on plasma homocysteine concentration in different European populations. *Atherosclerosis* 1998;136:347–354.

Hoffman EP. Muscular dystrophy: identification and use of genes for diagnostics and therapeutics. *Arch Path Lab Med*. 1999;123:1050–1052.

International RET Mutation Consortium Analysis, The relationship between specific RET protooncogene mutations and disease phenotype in multiple endocrine neoplasia type 2, *JAMA*, 1996;276:1575–1579.

Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. *Human Molecular Genetics* 2000;9:901–908.

Kant JA. Patent$, the human genome and the clinical laboratory. *ASCP News* 2000; November:7.

Koenig M, Beggs AH, Moyer M. et al. The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am J Hum Genet* 1989;45:498–506.

Krutovskikh V, Yamasaki H. Connexin gene mutations in human genetic diseases. *Mutation Research* 2000;462:197–207.

Kupferminc MJ, Eldor A, Steinman N, et al. Increased frequency of genetic thrombophilia in women with complications of pregnancy. *N Engl J Med* 1999;340:9–13.

Lakich D, Kazazian HH Jr, Antonarakis SE, Gitschier J. Inversions disrupting the factor VIII gene are a common cause of severe hemophilia A. *Nature Genet*. 1993;5:236–241.

Lieberman AP, Fischbeck KH. Triplet repeat expansion in neuromuscular disease. *Muscle & Nerve* 2000;23:843–850.

Mahowald MB, Verp MS, Anderson RR. Genetic Counseling: clinical and ethical challenges. *Ann Rev Genet* 1998;32:547–559.

Malone KE, Daling JR, Thompson JD, et al. BRCA1 mutations and breast cancer in the general population. *JAMA* 1998;279:922–929.

Mann MR, Bartolomei MS. Towards a molecular understanding of Prader-Willi and Angelman syndromes. *Human Molecular Genetics* 1999;8:1867–1873.

Martin JB. Molecular basis of the neurodegenerative disorders. *New Eng J Med* 1999;340:1970–1980.

Mathew CG. DNA diagnostics: goals and challenges. *British Medical Bulletin* 1999;55:325–339.

Merz JF, Cho MK, Robertson JH, et al. Disease gene patenting is a bad innovation. *Molecular Diagnosis* 1997;2:299–306.

Mickle JE, Cutting GR. Genotype-phenotype relationships in cystic fibrosis. *Med Clin North America* 2000;84:597–607.

Newman B, Mu H, Butler LM, et al. Frequency of breast cancer attributable to BRCA1 in a population-based series of American woman. *JAMA* 1998;279:915–921.

Nolin SL, Lewis III FA, Ye LL, et al. Familial transmission of the FMR1 CGG repeat. *Am J Hum Genet* 1996;59:1252–1261.

Olynyk JK, Digby JC, Aquilia, BA, et al. A population-based study of the clinical expression of the hemochromatosis gene. *New Eng J Med* 1999;341:718–724.

Rahman N, Stratton MR. The genetics of breast cancer susceptibility. *Annu Rev Genet* 1998;32:95–121.

Thornton C. The myotonic dystrophies. *Seminars in Neurology* 1999;19:25–33.

Wallace DC. Mitochondrial diseases in man and mouse. *Science* 1999;283:1482–1488.

Wagener C. Molecular diagnostics. *J Molecular Medicine* 1997;75:728–744.

Zlotogora J, Bach G, Munnich A. Molecular basis of mendelian disorders among Jews. *Mol Genet Metab* 2000;69:169–180.

9

Molecular Biology of Solid Tumors

Andrea Ferreira-Gonzalez

Arlene Buller

Mary E. Barcus

David S. Wilkinson

Carleton T. Garrett

Our understanding of the molecular changes responsible for malignant transformation has increased significantly over the past decade. As yet, however, the direct impact of this information on the diagnosis and management of patients with solid tumor neoplasia has been rather limited. Over the next decade, however, three factors are poised to alter this situation dramatically. First is the exponential increase in our knowledge of the human. Second is the dramatic increase in high throughput technology for detection and analysis of molecular changes in human tumors. Technology such as real-time polymerase chain reaction (PCR) and deoriboxynucleic acid (DNA) chips (1) are increasing the possibility that molecular testing can be performed with the turn-around times and the degree of reliability needed for performance in a clinical laboratory. Third is the enormous investment of private and public funds to identify potential molecular targets for new cancer therapies. This chapter will review some of the molecular changes that have been reported to contribute neoplastic transformation and tumor aggressiveness in solid neoplasms. These in turn will serve as potential targets and opportunities for future diagnostic, prognostic, and therapeutic interventions in solid tumors as well as provide some rationale for why a particular molecular change potentially may be important in each or all of these three areas (2).

MOLECULAR PATHOGENESIS OF NEOPLASTIC TRANSFORMATION

Part of "9 - Molecular Biology of Solid Tumors"

The central hypothesis underlying our current view of the mechanism of neoplastic transformation is that this process reflects a progressive adaptive change in cell regulation that is a result of progressive accumulation of changes in the DNA comprising the genome of the cell (3). These genetic changes result in activation of genes called *oncogenes* whose normal function is to induce growth or inhibit cell death (apoptosis); and, inactivation of genes called *tumor suppressor genes* whose normal functions are to suppress cell replication, induce cell death, and safeguard the integrity of the cellular genome (4). Beginning in the 1980s it became possible to transform nonneoplastic rodent and avian cells into tumor cells by introducing mutant and/or overexpressed versions of oncogenes into these cells in culture by noninfectious means (5, 6). Recently, nonneoplastic human cells were transformed to neoplastic cells through transfection of the cells with telomerase (hTERT), SV-40 large T antigen and mutant H-ras.[7] These studies implicate activation of telomerase and ras as well as inactivation of p53 and the retinoblastoma (RB) protein, as a minimum set of genetic changes for neoplastic transformation of some types of human cells (7). However, other molecular changes may be required for other types of human cells (8).

Oncogene Activation

Oncogenes originally were identified as genes that were present in acutely transforming retroviruses and were required for the virus to cause transformation of the infected cell in tissue culture (9). It subsequently was determined that these genes were derived from the host cell's DNA and that the genes had become incorporated into the retroviral genome through some recombinational event (10). Moreover, these genes were found frequently to play a role in normal cell replication and development in the absence of any retroviral involvement (11). Later, following development of methods to transfer noninfectious DNA obtained from primary tumors into susceptible eukaryotic cells, additional genes capable of inducing neoplastic transformation were identified including HER2/neu (12). Other nonretroviral oncogenes such as N-myc have been identified by virtue of their homology to retroviral-associated oncogenes as well as being amplified in certain human tumors (13).

Oncogenes have been found to cause transformation via two general mechanisms. The first is by inducing or facilitating cell replication and the second is by inhibiting apoptosis (Table 9.1). Facilitating cell replication can come about through stimulation of ‘upstream’ events beginning with binding of growth factors to growth factor receptors and extending through activation of transcription factors that turn on genes required for cell replication (14). Alternatively, some oncoproteins induce cell replication by directly interacting with the ‘downstream’ proteins that directly control the cell cycle (15). For a cell to divide, coordinated

replication of all of its various components together with the segregation of chromosomes into two daughter cells must occur. This complex process is controlled by the progressive phosphorylation and degradation of a group of molecules called cyclins through oligomerization between the cyclins and cyclin-dependent kinases (16). Over expression of these cyclins can cause the cell to enter into cellular division and promote the development of cancer (15, 16).

TABLE 9.1. CLASSIFICATION OF ONCOGENES

Class	Type	Examples
1	Growth factors	Sis, α-TGF,
2	Growth factor receptors	Erb (EGFR), HER2/neu,
3	Cytoplasmic signal Transduction modulators	H-, K-, N-ras, src, raf,
4	Nuclear transcription factors	N-, c-, L-myc, c-jun, fos
5	Cell cycle enhancers	PRAD1 (cyclin D)
6	Inhibitors of apoptosis	bcl-2, bcl-X

Apoptosis and programmed cell death refer to a process in which a cell causes its own death. It is an essential process for normal organ development and for the development of the immune system (17). Stimuli that induce apoptosis include the binding of specific ligands to certain types of receptors (18), the absence of certain growth factors in cells that have been stimulated to undergo cell division (19), and the inability of a cell to repair damaged DNA (20). Apoptosis is regulated by a series of competing proteins, some of which promote cell death and others of which inhibit it (21). Loss of apoptosis contributes directly to tumor development as well as tumor resistance to treatment with radiation and chemotherapy (22).

Inappropriate expression or over expression of oncogenes may occur through several mechanisms. At the genetic level, the oncogene may undergo mutations that enhance or constituitively activate the protein as is the case for the members of the ras gene family (23). Alternatively, the gene may become activated through gene rearrangement where it or an active portion of it is translocated into a region of the cell's genome that normally is up-regulated in that tissue (24). In the case of Ewing's sarcoma, a portion of the EWS gene, which is responsible for its oncogenic activity, joins to the FLI-1 gene as a result of a translocation between chromosomes 11 and 22 (25). A third mechanism of oncogene activation is demonstrated by c-erbB-2 and N-myc, which becomes over expressed in some breast cancers and neuroblastomas respectively through amplification of the oncogene (26, 27). Not all cases of over expression or inappropriate expression of an oncogene are associated with detectable mutations of the oncogene. However, in those cases where this is the case, the mutation represents a potential tumor marker of high specificity because it will be present in the tumor cells but not normal cells.

Genes Involved in Tissue Invasion, Metastases, and Angiogenesis

Malignant solid tumors derive most of their potential to destroy their host, not through direct expansion of the lesion, but through their ability to extend either by direct invasion or metastasis into other regions of the host's body. A second important feature of malignant tumors is their ability to induce the development of a new blood supply to the tumor cells (28). The ability of cancer cells to invade adjacent normal tissue and stimulate angiogenesis is in part because of their expression and secretion of extracellular proteinases (29). One group of extracellular matrix degrading enzymes that have been studied for their role in metastasis of tumor cells are the matrix metalloproteinases (MMPs). These include MMP-1, MMP-2 (72 kd type IV collagenase), MMP-3 (stromelysin), MMP-7 (pump-1, PUMP, matrilysin), MMP-9 (92 kd type IV collagenase) and MMP-10 (stromelysin-2). Induction of metalloproteinases has been demonstrated in several normal cells under normal physiological states such as invading trophoblasts (30), developing human mononuclear cells (31) and in normally cycling endometrium (32). Expression has been observed in some normal cells following stimulation by growth factors (33, 34). MMPs have been detected in both cancer cells and cancer cell stroma (35, 36) as have their naturally occurring inhibitors [tissue inhibitors of metallo-proteinases (TIMPs)] and tumor invasion has been regarded by some to result from an imbalance between MMPs and TIMPs (37, 38). A second family of proteinases, the serine proteases and in particular the plasminogen activator/plasmin system, also has been implicated in tumor cell invasion and metastases (39). Expression of the urokinase-type plasminogen activator (uPA) often is increased in malignant tumors (40). Extracellular proteinases appear to play an important role in both the stimulation and inhibition of angiogenesis (41).

Tumor Suppressor Genes

Based on studies that examined the phenotype of cells created from the fusion of malignant and benign cells, it has been recognized that certain genes can suppress the malignant or neoplastic phenotype and that these genes are, in fact, dominant in their biological behavior (42, 43). The first tumor suppressor gene to actually be isolated and its mechanism of action defined was the retinoblastoma gene (44) and its isolation served as a model for isolation of many subsequent tumor suppressor genes. The key to isolating the retinoblastoma gene was the observation that it appeared in familial cases to be transmitted in association with markers on chromosome 13 (44). Subsequent studies narrowed the region to a single gene on 13q14. This gene was shown to possess mutations in patients who developed retinoblastoma while family members who inherited the wild type, that is, normal retinoblastoma gene, did not develop tumors. It subsequently was demonstrated that the tumors arising in these individuals tended to retain the mutant gene but to lose the normal gene presumably through abnormal chromosomal segregation during mitosis (45). This process involving the physical loss of one of the alleles of a tumor suppressor gene during the course of neoplastic transformation is known as loss of heterozygosity (LOH). The process fits with an earlier prediction by Knudson regarding the mechanism for transmission of genetic cancers (46). It now is used as the major experimental approach to identifying new tumor suppressor genes (47).

Tumor suppressor genes appear to exert their antitumorigenesis activity by three basic mechanisms: (i) suppression of cellular

and DNA replication; (ii) promotion of apoptosis; and, (iii) maintenance of genomic integrity through DNA repair and chromosome stability. The terms ‘Gatekeepers’ and ‘Caretakers’ have been proposed to classify the effect of a tumor suppressor gene's activity (4, 48) though not specifying the mechanism through which this phenotypic effect is achieved. Thus, ‘Gatekeepers’ are genes that prevent cell replication when the cell's genomic integrity has been compromised while ‘Caretakers’ act to ensure genomic integrity through maintenance of chromosome stability and repair of DNA damage (4). In general, genes that appear to suppress tumorgenicity through molecular mechanisms one and two fall into the category of ‘Gatekeepers’ while those that utilize primarily mechanism three belong to the ‘Caretaker’ genes.

The overall result of the effects of all tumor suppressor genes appears to be to contribute to the maintenance of the integrity of the cellular genome. Genomic instability appears to be expressed in two forms. The first form might be characterized by ‘subtle’ instability in which changes in DNA such as point mutations and small deletions and insertions occur but where the number of chromosomes and general integrity of the genome remains unperturbed. This type of instability arises from either inactivation of genes in the nucleotide excision repair (NER) pathway, which result in the clinical syndrome of *xeroderma pigmentosa*; or inactivation of one of the DNA mismatch repair (MMR) genes, which have been implicated in the development of cases of colon cancer arising in hereditary nonpolyposis coli (4). This form of genomic instability is referred to as either nuclear excision repair instability (NIN) or mismatch repair instability (MIN) and cell lines with these genetic disorders show an expected increased mutation rate when compared to normal cells (4, 49).

TABLE 9.2. CLASSIFICATION OF TUMOR SUPPRESSOR GENES

Class	Type	Examples	Tumors	Ref
1	Suppressors of cell cycle	p16, p21	p16 (many solid tumor malignancies)	54
2	Enhancers of apoptosis	bax, FHIT	NSCLC (FHIT)	55
3	Protectors of genomic stability*			
	Mediators of DNA integrity			
	Nucleotide excision repair	Xeroderma pigmentosa DNA repair enzymes	UV induced skin cancer	56
	Mismatch repair	hMSH2, hMLH1, BRCA1, BRCA2	Colon cancer	57,58
	Mediators of chromosome stability			
	Number of chromosomes			
	Spindle check point proteins	hMAD2	Breast cancer, T-cell leukemia	59 60
		hBUB1, hBUBR1	Colorectal cancer	61,62
	DNA-damage check point proteins	ATM, ATR,		63,64
		BRCA1, BRCA2,		65
		p53		66
	Proteins modulating centrisome number	Aurora2/STK15	Small cell lung cancer	67
	Protect against translocations			
	DNA-damage check point proteins	(see above)		
	Normal recombination proteins	RAG1, RAG2		68
	Protect against amplification			
	DNA-damage check point proteins	p53		69

* Based on reference 4
Abbreviations:
ATM, ataxia telangiectasia mutated
ATR, ataxia and rad-related
BRCA1, breast cancer 1
BRCA2, breast cancer 2
FHIT, fragile histidine triad
hMSH2, human Mut 5 homolog 2
NSCLC, non-small cell lung cancer
RAG1, recombination activating gene 1
RAG2, recombination activating gene 2
UV, ultraviolet

The second form of genomic instability is identified by the presence of chromosomal instability (CIN). The CIN phenotype is characterized by a gain or loss of chromosomes (aneuploidy) and losses in heterozygosity even in cases where two alleles for a particular gene exist (4). This type of instability occurs in the majority of solid tumors and is more common than NIN or MIN (4, 47, 50, 51 and 52). Two characteristics of tumors that express the CIN phenotype are of interest. First is the fact that this phenotype appears to be dominant. Fusion of cells demonstrating CIN with those demonstrating MIN result in cells in which the mismatch repair deficiency is corrected but chromosomal instability still is present (4, 53). This suggests that only a single hit may be needed to induce the CIN phenotype. The second important characteristic of this phenotype is that a surprisingly large number of genes appear to be potentially involved in the creation of this phenotype [Table 9.2 (54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68 and 69)]. The CIN phenotype is postulated to arise as a result of a loss of function of ‘Gatekeeper’ genes that act as check points for each of the many processes that are involved in replication of the genome during cell division (Table 9.2). Disruption of these checkpoints presumably

permits development of the broad range of mutations (Table 9.2) that inevitably permit tumor cells to develop and express the malignant phenotype.

CLINICAL SIGNIFICANCE OF MOLECULAR CHANGES IN SOLID TUMORS

Part of "9 - Molecular Biology of Solid Tumors"

Use of Microsatellite and Other Molecular Changes in Tumor Diagnosis

The current standard approach for establishing a diagnosis of cancer is to obtain a tissue biopsy or cytologic sample of the suspected lesion and determine on the basis of morphological criteria whether it is benign or malignant. However, samples characterized by scanty atypical cells or subtle morphological differences between benign and malignant lesions can prove to be diagnostic dilemmas. For these reasons, the spectrum of somatic mutations associated with neoplastic transformation is being evaluated as markers of the neoplastic cells. In general, for solid tumors, detection of changes in microsatellites, point mutations, and viral nucleic acids appear to be promising as new targets for detecting and identifying tumors [Table 9.3 (25, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, and 84)].

TABLE 9.3. MOLECULAR CHANGES OF POTENTIAL USE IN TUMOR DIAGNOSIS

Molecular Marker	Examples	Reference
Microsatellite analysis		
LOH and MSI	Bladder cancer cytology	70,71 and 72
Oncogenes or TSGs mutations		
Ha-ras mutation	Bladder cancer cytology	73
Ki-ras mutation	Bronchopulmonary cytology	74,75
Oncogene translocations		
EWS-FLI-1 [t(11;22)]	Ewings sarcoma	25,76
EWS-ERG [t(21;22)]	Ewings sarcoma, pPNETS	77
PAX3-FKHR [t(2;13)]	Alveolar rhabdomyosarcoma	78
EWS-WT1 [t(11;22)]	Desmoplastic small round cell tumor	79
SYT-SSX [t(X;18)]	Synovial sarcoma	80,81
Oncogene expression		
Telomerase	Bladder cancer cytology	82
Detection of viral targets		
HPV	Endocervical neoplasia cytology	83,84

LOH, loss of heterozygosity
MSI, microsatellite instability
TSG, tumor suppressor gene
pPNETS, peripheral primitive neuroectodermal tumors of childhood

Microsatellites potentially are one of the most useful molecular markers for use in diagnosing cancer. They consist of regions of DNA that are composed of a ‘core’ sequence of from two to five nucleotides that are repeated sequentially from two to greater than 50 times. Thousands of microsatellites have been identified within the human genome and their locations mapped throughout the 22 autosomes and the two sex chromosomes. Microsatellites are the usual marker for studying loss of heterozygosity. They also may become mutated when there is inactivation within a tumor of one or both of the DNA mismatch repair genes. These mutations are expressed as new microsatellite alleles that differ from the normal germline alleles through an increased or decreased number of repeats. Such a finding is called microsatellite instability.

Alterations of microsatellites have been compared with conventional cytology for the diagnosis of both primary and recurrent transitional cell carcinoma (TCC) of the bladder (70, 71 and 72). Molecular examination showed increased sensitivity compared to cytologic evaluation, identifying microsatellite changes in 95% vs 50% of those individuals with bladder cancer. Two similar studies by Mourah et al. and Steiner et al. detected recurrent TCC in 83% and 91% of patients evaluated by microsatellite analysis (71, 72). This technique appears to represent a sensitive and noninvasive diagnostic modality that may supplement routine screening for this neoplasm. Alterations in microsatellites also have been proposed as aids in the evaluation of cytologic specimens from the head, neck, and lungs.

Detection of mutant oncogenes or mutant tumor suppressor genes also may prove useful in identifying tumor cells. In a study by Fitzgerald et al., H-ras mutations were observed in nearly half of patients evaluated with bladder tumors (73). In addition, molecular testing for H-ras mutations appeared to confer increased sensitivity as a screening method, detecting a significantly greater proportion of low-grade tumors than cytologic screening alone.

K-ras mutations have been identified in a number of tumors including lung adenocarcinomas. Evaluation of cytologic specimens such as sputum and bronchoalveolar lavage fluid by PCR-based molecular assays has shown promising results, identifying 50% to 80% of mutations shown to be present in resected material (74, 75). Such studies may prove to be a valuable complement to conventional cytologic examination in screening for neoplasia.

In addition to the detection of mutations as delineated above, characteristic translocations have been reported in some solid tumors, most notably sarcomas such as Ewings sarcoma, rhabdomyosarcoma, and synovial sarcoma (25, 76, 77, 78, 79, 80 and 81). Demonstration of these translocations by RT-PCR or cytogenetic studies may elucidate diagnoses of these challenging lesions and facilitate the detection of residual or recurrent disease. As with other ancillary techniques, a negative finding must be interpreted in conjunction with other clinical and morphologic data and does not invariably exclude a given lesion.

Finally, molecular methods may be employed in the detection of viral oncogenes known to be associated with malignancies. For example, with the use of techniques such as *in situ* hybridization, human papillomavirus (HPV) has been identified as a potential etiologic agent associated with cervical cancer. While the implementation of screening for cervical dysplasia by the Papanicolaou (PAP) smear has dramatically reduced the incidence of cervical cancer, the PAP smear carries a known risk of false negatives. Furthermore, because of the inability to predict, based upon morphology alone, which lesions will behave most aggressively and progress to carcinoma, inconclusive and abnormal results necessitate close follow-up. Some studies suggest that HPV testing may detect a subset of patients with high-grade cervical dysplasia not detected by conventional screening by PAP smear (83). In addition, HPV testing may serve to distinguish a subset of patients at greatest risk for progression of cervical dysplasia to carcinoma (84).

USE OF MOLECULAR CHANGES IN DETERMINING TUMOR PROGNOSIS

Part of "9 - Molecular Biology of Solid Tumors"

The prognosis of a patient following the diagnosis and treatment of malignancy traditionally has been based on the clinical and pathologic stage of disease and histological grade of the tumor. This analysis enables some estimation of disease outcome and stratification of patients into treatment regimens but it remains imperfect when attempting to apply the criteria to an individual patient. Molecular pathology has the potential to improve the assessment of tumor outcome through more precise staging (molecular staging) of the extent of disease and through identification of molecular changes that are associated with adverse clinical outcome.

Microsatellite and mutational analyses have demonstrated utility as prognostic indicators as illustrated in Table 9.4 (85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109 and 110). LOH at particular loci in carcinomas of the cervix, gastrointestinal tract, head, neck, and lung appears to denote a subset of high-grade tumors or a poor prognosis. Lee et al. have reported a three-fold increased risk of death in patients with head and neck squamous cell carcinomas (HNSCC) that demonstrate LOH of 14q (99). In a separate study of HNSCC, LOH at 17p appeared to correlate with poor response to chemotherapy (101).

TABLE 9.4. ASSOCIATIONS BETWEEN MOLECULAR CHANGES AND PROGNOSIS

Tumor	Alteration	Association	References
Bladder cancer	LOH RB	High grade/muscle invasion	86
	Genomic alterations 2q-, 5p+, 5q-, 6q-, 8p-, 10q-, 18q-, 20q+	Higher grade	87
Breast cancer	Plasma DNA similar to tumor DNA	Poor prognosis	88
	Allelic loss at 1p22–p31	Lymph node metastasis, tumor size 2 cm	89
Cervical carcinoma	LOH on chromosome 1	Advanced stage	90
Colorectal cancer	LOH at 18q21	Recurrence/poor survival (Dukes B & C)	91,92
	p53 expression	Recurrence/poor survival (Dukes A)	92
	MI and K-*ras* mutations in normal-appearing colonic mucosa micrometastases	Predictive of colorectal cancer	93
		Decreased survival (50 vs. 91%)	94
	P16-hypermethylation	Shorter survival in T3N0M0 tumors	95
Gastric cancer	LOH p53	Invasive potential	96
	LOH of 7q (D7S95)	Poor prognosis	97
Gliomas	Chromosome 22q loss	Astrocytoma progression	98
HNSCC	LOH of 14q	Poor outcome	99
	LOH on 2q	Poor prognosis	100
	LOH at 17p	Chemoresistance	101
Melanoma	LOH in plasma	Advanced stage/tumor progression	102
Neuroblastoma	N-*myc* amplification	Poor prognosis	103
	TRK-A expression	Good prognosis	103
Neuroblastomas, 4s	N-*myc* amplification, 1p deletion,17q gains, elevated telomerase activity	Poor outcome (*not independent*)	104
NSCLC	Allelic imbalances on 9p	Poor prognosis	105
	Reduced Fhit protein expression	Poor prognosis (stage I)	106
	LOH 11p13	Poor prognosis	107
PNET	LOH of 17p	Metastatic disease	108
	c-*myc* amplification	Poor prognosis	108
Prostate cancer	LOH on 13q	Advanced stage	109
Retinoblastoma	LOH at RB1 locus	Tumoral differentiation, absence of choroidal invasion	110

(a) Modified from Table 1 in reference 85
Abbreviations: HNSCC, head and neck squamous cell carcinoma
LOH, loss of heterozygosity
NSCLC, non-small cell lung cancer
PNET, primitive neuroectodermal tumor
RB, retinoblastoma gene

Amplification of several oncogenes appears to have prognostic relevance in certain tumors. In neuroblastoma, N-*myc* amplification has been shown to be predictive of poor prognosis (103). Brinkschmidt et al. have also reported poor outcomes in stage 4s patients with deletions of 1p, gains on 17q, and elevated telomerase activity (104). In contrast, TRK-A expression appears to

confer a favorable prognosis (103). Another well-studied example of where oncogene amplification may have clinical relevance is the amplification of the c-*erb*B-2 or HER-2/*neu* proto-oncogene. This oncogene encodes a growth factor receptor with tyrosine kinase activity (27). Amplification of this oncogene has been identified in neoplasms of the breast, stomach, endometrium, ovary, and lung (27). c-erbB-2 amplification appears to correlate with poor prognosis and chemoresistance in some neoplasms (111).

Molecular staging is another approach to evaluating the prognosis in patients with solid tumors (94, 112, 113 and 114). This technique frequently is referred to as examination for minimal residual disease when applied to the study of hematological malignancies. The standard approach uses RT-PCR to detect the presence of tumor cells in lymph nodes or bone marrow that cannot be identified by standard histologic or immunohistochemical evaluation. Shivers et al. (112) and Bostick et al. (113) have reported the detection of micrometastases by RT-PCR in melanoma patients and have demonstrated a statistically significant increased rate of recurrence in those individuals positive by the molecular assay.

In a comparable study by Liefers et al. (94), RT-PCR for carcinoembryonic antigen (CEA) identified occult lymph node metastases in approximately 50% of stage II colorectal carcinoma patients (node-negative by histology). The presence of micrometastases correlated with poor prognosis, identifying a subset of patients with 50% 5-year survival compared to 91% in those individuals negative for metastases by both histology and RT-PCR. In addition, RT-PCR for CEA has been employed for the detection of metastatic esophageal cancer in histologically negative lymph nodes (114).

Increased telomerase activity appears to represent a promising molecular marker of neoplasia and is useful as a marker in molecular staging. For example, Soria et al. reportedly have measured telomerase activity as a means to identify circulating tumor cells in the peripheral blood of breast cancer patients with metastatic disease (115). Following the extraction of epithelial cells from blood using immunomagnetic beads coated with antibodies specific to epithelial cells (BerEP4), telomerase activity was assayed using PCR and ELISA methodology. Eighty-four percent of stage IV breast cancer patients had detectable telomerase activity in the isolated epithelial cells compared to none of the tested control cases. Telomerase activity as determined by the telomeric repeat amplification protocol (TRAP) also has been shown to be predictive of poor prognosis in neuroblastoma (104).

Therapeutic Implications of Molecular Changes

The identification of the multitude of genomic alterations that characterize neoplasms invites the development of new molecular therapies for solid tumors. Current approaches to gene therapy include the use of “suicide gene” to activate drugs at the tumor, transfer of tumor suppressor genes, and the inhibition of activated or amplified oncogenes with the use of antisense oligonucleotides and monoclonal antibodies (116). In addition, the development of microarrays, which enable a rapid assessment of presumed therapy response based upon genetic make-up, has the potential to allow more individualized treatment of oncology patients, conferring optimal response yet minimizing drug toxicity (117). While still in their infancy, the fields of gene therapy and pharmacogenomics undoubtedly will make significant contributions to the management of neoplastic disease over the next decade.

Perhaps the most recognized example of a new therapy based on the presence of a molecular change is the use of Herceptin (Genentech, Inc., San Francisco, California) in the treatment of breast cancer (118, 119, 120, 121 and 122). This recombinant humanized monoclonal antibody to the Her2 receptor protein has shown efficacy in clinical trials for use in women with metastatic breast cancers that overexpress the protein. Studies have demonstrated a survival benefit when administered either independently or in combination with conventional chemotherapeutic agents (121, 122). Antisense oligodeoxynucleotides (ODNs) also are being examined as potential antitumor agents that, by virtue of sequence, could be targeted specifically to induce decreased oncogene expression through binding and degradation of RNA. Initial investigations of HER2/neu antisense ODNs by Roh et al. reveal increased tumor response to chemotherapy in the presence of these agents (123).

Therapy directed at the restoration of tumor suppressor gene function through transfer of wild-type genes is another potential avenue for therapies based on identification of specific molecular lesions. Various studies have focused on the delivery of p53, p16, and PTEN to tumor cells, although the optimal vector for gene import remains undetermined (124, 125, 126 and 127). For example, adenovirus-mediated transfer of p16INK4A protein expression has demonstrated a significant therapeutic effect against HNSCC in cell lines and mouse studies (125). Studies evaluating p53 gene transfer in patients with advanced NSCLC have shown stabilization of disease in 64% of subjects and partial response in 8% of subjects receiving multiple intratumoral doses (126). Preliminary results of clinical trials evaluating adenovirus-mediated transfer of p53 in advanced head and neck tumors also reveal promising results (127).

SPECIAL CONSIDERATIONS IN HEREDITARY CANCER SYNDROMES

Part of "9 - Molecular Biology of Solid Tumors"

Although the majority of cancer is sporadic, clustering of cases often is observed within families. In recent years, an increasing number of cancer-predisposing genes segregating in families have been isolated. Accounting for approximately 5% of all cancer, familial cancer syndromes most commonly result from mutations in tumor suppressor and DNA repair genes. A person carrying a germline mutation in a cancer-susceptibility gene generally is predisposed, but not mandated, to develop cancer. Although all cells in the body contain the mutation, only a fraction of cells in specific tissues may develop cancer. Moreover, some individuals with the mutation will remain disease free. Multiple factors, some not yet understood, control disease development. Complicating factors include variability in the penetrance, or the likelihood that a mutation in a gene will be expressed, and the expressivity of the disease, which includes the severity and age of onset. Penetrance and expressivity are dependent

on the type of mutation, cell type, effects of modifier genes, and environmental effects such as diet and carcinogen exposure. Variable expressivity and penetrance are hallmarks of all of the known familial cancer syndromes defined thus far, and have a direct effect on how the physician will manage the patient and the patient's family (128).

Genetic testing is available for inherited predisposition to a number of common as well as unusual types of cancer. A list of inherited cancers and their cancer predisposition genes is shown in Table 9.5 (128, 129, 130, 131, 132, 133, 134, 135, 136 and 137). Inherited breast and ovarian cancer syndrome provides a detailed illustration of the important features of genetic predisposition to malignancy. Five to 10% of breast and ovarian cancer have been linked to mutations in the BRCA1/2 tumor suppressor genes on chromosomes 17 and 13, respectively. Most families with multiple cases of breast and ovarian cancers and approximately half of the early-onset breast cancer families have mutations in BRCA1 (128, 138, 139). Colon and prostate cancers also are observed at a higher frequency than in the general population (128, 139, 140). Mutations in BRCA2 are responsible for approximately one third of inherited breast cancer cases and are associated with breast cancer in males (141). Testing for mutations in these genes is offered only to individuals with a family history suggestive of inherited breast/ovarian cancer syndrome. When a genetic alteration is identified in an individual, several factors are taken into account in assessing the risk of actually being affected with cancer. First of all, is the genetic change indeed a deleterious mutation? It is always desirable to test an affected member of the family first to determine whether a particular genetic change is responsible for the disease. Hundreds of genetic changes have been identified in BRCA1/2, many of which are harmless polymorphisms that do not affect the function of the genes. If the mutation is present only in the affected person's DNA and not in healthy members of the family, then it can be presumed that the mutation is responsible for the disease. If it is not possible to test an affected member of the family, then any genetic changes that are detected must be compared to a genetic database to attempt to determine whether the change is the causative mutation. Once a mutation is confirmed, then a person's risk is calculated based on the number of affected relatives, the occurrence of bilateral versus unilateral breast cancer, the presence of ovarian cancer, and the age of onset. Penetrance and expressivity are variable.

TABLE 9.5. GENES ASSOCIATED WITH INHERITED CANCER SYNDROMES

Condition	Gene(s)	Gene Classification	Location/Type of Tumors	References
Breast and/or ovarian cancer	BRCA1 and BRCA2	Tumor suppressor	Breast*, ovary*, colon, prostate,	128
Familial adenomatous polyposis (FAP)	APC	Tumor suppressor	Colon*, small bowel, stomach, bone, CNS, soft tissue	128,129 and 130
Hereditary nonpolyposis colorectal cancer (HNPCC)	hMSH2, hMLH1, PMS1, PMS2, hMSH3, hMSH6	Mismatch repair	Right side of colon*, endometrium*, stomach, ovary, small bowel, transitional cell, sebaceous, CNS	128,129 and 130
Renal cancer	VHL	Tumor suppressor	Kidney*, hemangioblastoma, phaeochromocytoma	128
Neurofibromatosis types I/II	NF1, NF2	Tumor suppressor	CNS*, myeloid leukemia	128
Nevoid basal cell carcinoma	PTC	Development and cellular proliferation	Skin*, CNS	128
Inherited prostate cancer	BRCA1, HPC1, others	Tumor suppressor	Prostate*	128,131
Familial melanoma	MLM1, CDKN2A, CDK4	Cell cycle regulator, oncogene (CDK4)	Skin*, pancreas	128,132
Multiple endocrine neoplasia (MEN2A/B)	RET	Oncogene	Thyroid* and adrenal glands	128,133,134
Li-Fraumeni syndrome	P53	Tumor suppressor	Bone, soft tissue, breast, brain, adrenal gland, leukemia	128,135,136
Tuberous sclerosis	TSC1, TSC2	Tumor suppressor	Hamartomas of skin, CNS, heart, kidney	128
Retinoblastoma	RB1	Tumor suppressor	Eye*, bone, soft tissue, brain, skin	128
Rhabdoid predisposition	HSNF5/INI1	Tumor suppressor	CNS, kidney	137
Xeroderma pigmentosum	XP	DNA repair	Skin	128
Bloom's syndrome	BLM	Helicase (DNA ligation)	Leukemia, GI carcinoma	128
Ataxia telangiectasia	ATM	Cell cycle regulator	Leukemia, lymphoma, GI carcinoma	128

* Most common location of tumors

Determining that a person carries a cancer-predisposing mutation can greatly affect his/her medical management. There is significant controversy in the literature regarding this issue for inherited breast/ovarian syndrome (142). Due to the variable

penetrance of mutations in BRCA1/2, there exists the possibility some individuals with a positive test result will never develop the disease. Therefore, carriers of BRCA1/2 mutations should be familiarized thoroughly with the advantages and disadvantages of each treatment option.

Genetic testing has been well-documented to offer treatment alternatives in the medical care of at-risk individuals. Prophylactic colectomy is the treatment of choice for individuals at risk for familial adenomatous polyposis (FAP) because mutations in the adenomatous polyposis coli (APC) tumor suppressor gene have a high penetrance (129). Therefore, a negative genetic test result will prevent such drastic measures from being necessary. In hereditary nonpolyposis colon cancer (HNPCC), carriers of mutations in one of several DNA repair genes (MLH1, MSH2, MSH6, PMS1 and PMS2) are at an 80% lifetime risk of colorectal cancer, and are advised to undergo surveillance by colonoscopy every 1 to 3 years, or prophylactic colectomy (130). Studies have indicated that increased surveillance in HNPCC mutation carriers increases life expectancy by seven years compared to no surveillance (129).

In multiple endocrine neoplasia type 2 families (MEN2), individuals with a 50% risk of inheriting mutations in the RET proto-oncogene must undergo prophylactic thyroidectomy and/or annual biochemical screening at an early age because medullary thyroid carcinoma (MTC) can develop as early as 6 years of age (134). Genetic testing can determine at an early age those individuals who require increased surveillance and/or thyroidectomy, while sparing those individuals who have not inherited a causative mutation.

The era of genetic testing has unveiled a plethora of ethical dilemmas, not the least of which is the use of archival material for DNA studies. Requiring informed consent for the use of archival materials in genetic research serves to protect both the researcher and the source of the samples from legal, social, and psychological sequelae. A balance must be developed in which the potential benefits of the use of stored tissues for genetic research are weighed against the rights of the individual not to have his/her genetic material scrutinized unknowingly. Although it is mandated by federal regulations that informed consent must be obtained when an identifiable sample is used for genetic studies, unidentifiable samples are exempt from this requirement (143). A question arises regarding stored tissues as to whether consent obtained at the time of sample collection is valid for further studies at a later date. There is consensus that all identifiable stored samples require an informed consent before use in a research protocol (143, 144). For samples that will be anonymized, the recommendations for obtaining consent vary with the protocol to be used (143). Guidelines regarding informed consent are in the process of development and certainly will be an important issue in the future.

CONCLUSIONS

Part of "9 - Molecular Biology of Solid Tumors"

While recent findings in molecular biology have greatly increased our understanding of the molecular mechanisms responsible for expression of the malignant phenotype, no single molecular change group of molecular changes has been identified to date that permits a clinician to diagnose, predict the behavior, or treat all or even most cancers. This hardly is surprising given the complexity of the molecular changes observed to occur with neoplastic transformation. Perhaps because of this complexity, molecular information will not be of practical value to the clinical management of solid tumors until technology evolves to a point where very small samples of tumor cells can be simultaneously and rapidly evaluated for 50 to 100 or more genetic changes. However, this may be an exaggeration of the degree of genetic information needed for clinical utility. For example, identification of a single clonal mutation that is present in an individual's tumor cells offers itself as potential marker for distinguishing normal from neoplastic cells in that individual. Moreover, some mutations such as loss of heterozygosity on 17p and 9p are present in a significant fraction of the more common solid tumors and thus offer themselves as potential tumor markers. Whether one or a limited small number of genetic mutations will prove to be independently useful in predicting prognosis of individual tumors is less clear although some encouraging findings have been observed with regard to N-myc amplification in neuroblastomas and c-erbB-2 (HER2/neu) over expression in breast cancers. Information is likely only to increase as sequencing of the human genome is completed and the genes associated with malignancy further defined. Advances in technology in the form of microfluidics and high density DNA and complementary DNA (cDNA) arrays will provide platforms that potentially will allow the complex changes present in individual tumors to be evaluated in real time. The time frame for all of this to become integrated into routine clinical laboratory practice is unclear but the rate of discovery of new information as well as development of new technology continues to accelerate.

REFERENCES

1. Khan J, Bittner ML, Chen Y, et al. DNA microarray technology: the anticipated impact on the study of human disease. *Biochim Biophys Acta* 1999;1423:M17–M28.

2. Cairns P, Sidransky D. Molecular methods for the diagnosis of cancer. *Biochim Biophys Acta* 1999;1423:C11–C28.

3. Fearon ER, Cho KR, Vogelstein B, et al. A genetic model for colorectal tumorigenesis. *Science* 1990;247:49–56.

4. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998;396:643–649.

5. Reddy, EP, Reynolds RK, Santos E, et al. A point mutation is responsible for the acquisition of transforming properties of the T24 human bladder carcinoma oncogene. *Nature* (London) 1982;300:149–152.

6. Land H, Parada LF, Weinberg RA. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature* (London) 1983;304:596–602.

7. Hahn WC, Counter CM, Lundberg AS, et al. Creation of human tumor cells with defined genetic elements. *Nature* 1999;400:464–468.

8. Morales CP et al. Absence of cancer-associated changes in human fibroblasts immortalized with telomerase. *Nature Genet* 1999;21:115–118.

9. Bishop JM. Cellular oncogenes and retroviruses. *Ann Rev Biochem* 1983;52:301-354.

10. Varmus HE. Form and function of retroviral proviruses. *Science* 1982;216:812–820.

11. Varmus HE. The molecular genetics of cellular oncogenes. *Ann Rev Genet* 1984;18:553–612.

12. Shih C, Padhy LC, Murray M, Weinberg RA. Transforming genes of carcinomas and neuroblastomas introduced into mouse fibroblasts. *Nature* (London) 1981;290:261–264.

13. Schwab M, Alitalo K, Klempnauer KH, et al. J. amplified DNA with limited homology to *myc* cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumor. *Nature* (London) 1983;305:245–248.

14. Hunter T. Oncoprotein networks. *Cell* 1997;88:333–346.

15. Bartek J, Lukas J, Bartkova J. Perspective: defects in cell cycle control and cancer. *J Pathol* 1999;187:95–99.

16. Hall M, Peters G. Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. *Adv Cancer Res* 1996;68:67–108.

17. Evan G, Littlewood T. A matter of life and cell death. *Science* 1998;281:1317–1322.

18. Askenazi A, Dixit VM. Death receptors: signaling and modulation. *Science* 1998;281:1305–1308.

19. Evan GI, Wylie AH, Gilbert CS, et al. Induction of apoptosis in fibroblasts by c-myc protein. *Cell* 1992;69:119–128.

20. Polyak K, Xia Y, Zweier JL, et al. A model for p53-induced apoptosis. *Nature* 1997;389:300–305.

21. Reed JC. Bcl-2 family proteins: regulators of apoptosis and chemoresistance in hematologic malignancies. *Semin Hematol* 1997;34 (Suppl 5):9–19.

22. Schmitt CA, Lowe SW. Apoptosis and therapy. *J Pathol* 1999;187:127–137.

23. Fasano O, Aldrich T, Tamanoi F, et al. Analysis of the transforming potential of the human H-ras gene by random mutagenesis. *Proc Natl Acad Sci* 1984;81:4008–4012.

24. Rabbitts TH. Perspective: chromosomal translocations can affect genes controlling gene expression and differentiation-why are these functions targeted? *J Pathol* 1999;187:39–42.

25. Dockhorn-Dworniczak B, Schafer KL, Dantcheva R, et al. Diagnostic value of the molecular genetic detection of the t(11;22) translocation in Ewing's tumors. *Virchows Arch* 1994;425:107–112.

26. Seeger RC, Brodeur GM, Sather H, et al. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N Engl J Med* 1985;313:1111–1116.

27. Paik S. Clinical significance of erbB-2 (HER-2/neu) protein. *Cancer Invest* 1992;10:575–579.

28. Folkman J. The role of angiogenesis in tumor growth. *Semin Cancer Biol* 1992;3:65–71.

29. Blasi F, Stoppelli MP. Proteases and cancer invasion: from belief to certainty. AACR Meeting on Proteases and Protease Inhibitors in Cancer, Nyborg, Denmark, 14–18 June 1998. *Biochim Biophys Acta* 1998;1423:R35–R44.

30. Fernandez, PL, Melino MJ, Nogales FF, et al. Immunohistochemical profile of basement membrane proteins and 72 kilodalton type IV collagenase in the implantation placental site. An integrated view. *Lab Invest* 1992;66:572–579.

31. Busiek, DF, Ross FP, McDonnell S, et al. The matrix metalloprotease matrilysin (PUMP) is expressed in developing human mononuclear phagocytes. *J Biol Chem* 1992;267:9087–9092.

32. Rodgers, WH, Osteen KG, Matrisian LM, et al. Expression and localization of matrilysin, a matrix metalloproteinase, in human endometrium during the reproductive cycle. *Am J Obstet Gynecol* 1993;168:253–260.

33. Yanagi H, Sasguri Y, Sugama K, et al. Production of tissue collagenase (matrix metalloproteinase 1) by human aortic smooth muscle cells in response to platelet-derived growth factor. *Nature* (London) 1991;352:345–347.

34. Bauer EA, Cooper TW, Huang JS, et al. Stimulation of in vitro human skin collagenase expression by platelet-derived growth factor. *Proc Natl Acad Sci USA* 1985;82:4132–4136.

35. Wolf C, Rouyer N, Lutz Y, et al. Stromelysin 3 belongs to a subgroup of proteinases expressed in breast carcinoma fibroblastic cells and possibly implicated in tumor progression. *Proc Natl Acad Sci USA* 1993;90:1843–1847.

36. Pyke C, Ralfkiaer E, Tryggvason K, et al. Messenger RNA for two type IV collagenases is located in stromal cells in human colon cancer. *Amer J Pathol* 1993;142:359–365.

37. Albini A, Melchioi A, Santi L, et al. Tumor cell invasion inhibited by TIMP-2. *J Natl Cancer Inst* 1991;83:775–779.

38. Azzarn HS, Arand G, Lippman ME, et al. Association of MMP-2 activation potential with metastatic progression in human breast cancer cell lines independent of MMP-2 production. *J Natl Cancer Inst* 1993;85:1758–1764.

39. Vassalli JD, Sappino AP, Belin D. The plasminogen activator/plasmin system. *J Clin Invest* 1991;88:1067–1072.

40. Pyke C, Kristensen P, Ralfkiaer E, et al. Urokinase-type plasminogen activator is expressed in stromal cells and its receptor in cancer cells at invasive foci in human colon adenocarcinomas. *Amer J Pathol* 1991;138:1059–1067.

41. Hart IR. Perspective: tumour spread-the problems of latency. *J Pathol* 1999;187:91–94.

42. Stanbridge EJ. Genetic analysis of human malignancy using somatic cell hybrids and monochromosome transfer. *Cancer Surv* 1988;7:317–324.

43. Goyette MC, Cho K, Fasching CL, et al. Progression of colorectal cancer is associated with multiple tumor suppressor gene defects but inhibition of tumorigenicity is accomplished by correction of any single defect via chromosome transfer. *Mol Cell Biol* 1992; 12:1387–1395.

44. Benedict WF, Murphree AL, Banerjee A, et al. Patient with 13 chromosome deletion: evidence that the retinoblastoma gene is a recessive cancer gene. *Science* 1983;219:973–975.

45. Dryja TP, Cavenee W, White R, et al. Homozygosity of chromosome 13 in retinoblastoma. *N Engl J Med* 1984;310:550–553.

46. Knudson AG. Hereditary cancer, oncogenes, and anti-oncogenes. *Cancer Res* 1985;45:1437–1443.

47. Vogelstein B, Fearon ER, Kern SE, et al. Allelotype of colorectal carcinomas. *Science* 1989;244:207–222.

48. Kinzler KW, Vogelstein B. Gatekeepers and caretakers. *Nature* 1997;386:761–763.

49. Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* 1991;51:3075–3079.

50. Boige V, Laurent-Puig P, Fouchet P, et al. Concerted nonsyntenic allelic losses in hyperploid hepatocellular carcinoma as determined by a high-resolution allelotype. *Cancer Res* 1997;57:1986–1990.

51. Seymour AB, Hruban RH, Redston M, et al. Allelotype of pancreatic adenocarcinoma. *Cancer Res* 1994;54:2761–2764.

52. Radford DM, et al. Allelotyping of ductal carcinoma in situ of the breast: deletion of loci on 8p, 13q, 16q, 17p and 17q. *Cancer Res* 1995;55:3399–3405.

53. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature* 1997;386:623–627.

54. Liggett WH Jr, Sidransky D. Role of the p16 tumor suppressor gene in cancer. *J Clin Oncol* 1998;16:1197–1206.

55. Druck T, Berk L, Huebner K. FHITness and cancer. *Oncol Res* 1998;10:341–345.

56. Bootsma D, Kraemer KH, Cleaver JE, et al. In: Kinzler KW, Vogelstein B, eds. The Genetic Basis of Human Cancer. New York: McGraw-Hill, 1998:245–274.

57. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colorectal cancer. *Cell* 1993;75:1215–1225.

58. Bronner CE, Baker SM, Morrison PT, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* 1994;368:258–261.

59. Li Y and Benezra R. Identification of a human mitotic checkpoint gene: hMAD2. *Sci* 1996;274:246–248.

60. Jin DY, Spencer F, Jeang KT. Human T cell leukemia virus type 1 oncoprotein tax targets the human mitotic checkpoint protein MAD1. *Cell* 1998;93:81–91.

61. Cahill DP, Lengauer C, Yu J, et al. Mutations of mitotic checkpoint genes in human cancers. *Nature* 1998;392:300–303.

62. Taylor SS, McKeon F. Kenetochore localization of murine Bub1 is required for normal mitotic timing and checkpoint response to spindle damage. *Cell* 1997;89:727–735.

63. Rotman G, Shiloh Y. ATM: from gene to function. *Hum Mol Genet* 1998;7:1555–1563.

64. Smith L, Liu SJ, Goodrich L, et al. Duplication of ATR inhibits MyoD, induces aneuploidy and eliminates radiation-induced G1 arrest. *Nature Genet* 1998;19:39–46.

65. Zhang H, Tombline G, Weber BL. BRCA1, BRCA2 and DNA damage response: collision or collusion? *Cell* 1998;92:433–436.

66. Lane D. Awakening angels. *Nature* 1998;394:616–617.

67. Wolf G, Elez R, Doermer A, et al. Prognostic significance of polo-like kinase (PLK) expression in non-small cell lung cancer. *Oncogene* 1997;14:543–549.

68. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* 1998;94:463–470.

69. Livingstone LR, White A, Sprouse J, et al. Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell* 1992;70:923–935.

70. Mao L, Schoenberg MP, Scicchitano M, et al. Molecular detection of primary bladder cancer by microsatellite analysis. *Science* 1996;271:659–662.

71. Steiner G, Schoenberg MP, Linn JF, et al. Detection of bladder cancer recurrence by microsatellite analysis of urine. *Nat Med* 1997;3:621–624.

72. Mourah S, Cussenot O, Vimont V, et al. Assessment of microsatellite instability in urine in the detection of transitional-cell carcinoma of the bladder. *Int J Cancer* 1998;79:629–633.

73. Fitzgerald JM, Ramchurren N, Rieger K, et al. Identification of H-ras mutations in urine sediments complements cytology in the detection of bladder tumors. *J Natl Cancer Inst* 1995;87:129–133.

74. Mao L, Hruban RH, Boyle JO, et al. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. *Cancer Res* 1994;54:1634–1637.

75. Ahrendt SA, Chow JT, Xu LH, et al. Molecular detection of tumor cells in bronchoalveolar lavage fluid from patients with early stage lung cancer. *J Natl Cancer Inst* 1999;91:332–339.

76. Sorensen PH, Liu XF, Delattre O, et al. Reverse transcriptase PCR amplification of EWS/FLI-1 fusion transcripts as a diagnostic test for peripheral primitive neuroectodermal tumors of childhood. *Diagn Mol Pathol* 1993;2:147–157.

77. Sorensen PH, Wu JK, Berean KW, et al. Olfactory neuroblastoma is a peripheral primitive neuroectodermal tumor related to Ewing sarcoma. *Proc Natl Acad Sci* 1996;93:1038–1043.

78. Downing JR, Khandekar A, Shurtleff SA, et al. Multiplex RT-PCR assay for the differential diagnosis of alveolar rhabdomyosarcoma and Ewing's sarcoma. *Am J Pathol* 1995;146:626–634.

79. Gerald WL, Ladanyi M, de Alava E, et al. Clinical, pathologic, and molecular spectrum of tumors associated with t(11;22)(p13;q12): desmoplastic small round-cell tumor and its variants. *J Clin Oncol* 1998;16:3028–3036.

80. Argani P, Zakowski MF, Klimstra DS, et al. Detection of the SYT-SSX chimeric RNA of synovial sarcoma in paraffin-embedded tissue and its application in problematic cases. *Mod Pathol* 1998;11:65–71.

81. Kawai A, Woodruff J, Healey JH, et al. SYT-SSX gene fusion as a determinant of morphology and prognosis in synovial sarcoma. *N Engl J Med* 1998;338:153–160.

82. Ito H, Kyo S, Kanaya T et al. Detection of human telomerase reverse transcriptase messenger RNA in voided urine samples as a useful diagnostic tool for bladder cancer. *Clin Cancer Res* 1998;4:2807–2810.

83. Cuzick J, Szarewski A, Terry G, et al. Human papillomavirus testing in primary cervical screening. *Lancet* 1995;345:1533–1536.

84. Manos MM, Kinney WK, Hurley LB, et al. Identifying women with cervical neoplasia: using human papilloma virus. *JAMA* 1999;281:1605–1610.

85. Barcus ME, Ferreira-Gonzalez A, Buller AM, et al. Genetic Changes in Solid Tumors. *Semin Surg Oncol* 2000; 18(4):358-70.

86. Cairns P, Proctor AJ, Knowles MA. Loss of heterozygosity at the Rb locus is frequent and correlates with muscle invasion in bladder carcinoma. *Oncogene* 1991;6:2305–2309.

87. Zhao J, Richter J, Wagner U, et al. Chromosomal imbalances in noninvasive papillary bladder neoplasms (pTa). *Cancer Res* 1999;59:4658–4661.

88. Silva JM, Dominguez G, Garcia JM, et al. Presence of tumor DNA in plasma of breast cancer patients: clinicopathological correlations. *Cancer Res* 1999;59:3251–3256.

89. Tsukamoto K, Ito N, Yoshimoto M, et al. Allelic loss on chromosome 1p is associated with progression and lymph node metastasis of primary breast carcinoma. *Cancer* 1998;82:317–322.

90. Cheung TH, Chung TK, Poon CS, et al. Allelic loss on chromosome 1 is associated with tumor progression of cervical carcinoma. *Cancer* 1999;86:1294–1298.

91. Jernvall P, Makinen MJ, Karttunen TJ, et al. Loss of heterozygosity at 18q21 is indicative of recurrence and therefore poor prognosis in a subset of colorectal cancers. *Br J Cancer* 1999;79:903–908.

92. Jernvall P, Makinen MJ, Karttunen TJ, et al. Morphological and genetic abnormalities in prediction of recurrence in radically operated colorectal cancer. *Anticancer Res* 1999;19:1357–1362.

93. Minamoto T, Esumi H, Ochiai A, et al. Combined analysis of microsatellite instability and K- *ras* mutation increases detection incidence of normal samples from colorectal cancer patients. *Clin Cancer Res* 1997;3:1413–1417.

94. Liefers GJ, Cleton-Jansen AM, van de Velde CJ, et al. Micrometastases and survival in stage II colorectal cancer. *N Engl J Med* 1998;339:223–228.

95. Liang JT, Chang KJ, Chen JC, et al. Hypermethylation of the p16 gene in sporadic T3N0M0 stage colorectal cancers: association with DNA replication error and shorter survival. *Oncology* 1999;57:149–156.

96. Sugai T, Habano W, Nakamura S, et al. Correlation of histologic morphology and tumor stage with molecular genetic analysis using microdissection in gastric carcinomas. *Diagn Mol Pathol* 1998;7:235–240.

97. Kuniyasu H, Yasui W, Yokozaki H, et al. Frequent loss of heterozygosity of the long arm of chromosome 7 is closely associated with progression of human gastric carcinomas. *Int J Cancer* 1994;59:597–600.

98. Ino Y, Silver JS, Blazejewski L, et al. Common regions of deletion on chromosome 22q12.3-q13.1 and 22q13.2 in human astrocytomas appear related to malignancy grade. *J Neuropathol Exp Neurol* 1999;58:881–885.

99. Lee DJ, Koch WM, Yoo G, et al. Impact of chromosome 14q loss on survival in primary head and neck squamous cell carcinoma. *Clin Cancer Res* 1997;3:501–505.

100. Ransom DT, Barnett TC, Bot J, et al. Loss of heterozygosity on chromosome 2q: possibly a poor prognostic factor in head and neck cancer. *Head Neck* 1998;20:404–410.

101. Blons H, Cabelguenne A, Carnot F, et al. Microsatellite analysis and response to chemotherapy in head-and-neck squamous-cell carcinoma. *Int J Cancer* 1999;84:410–415.

102. Fujiwara Y, Chi DD, Wang H, et al. Plasma DNA microsatellites as tumor-specific markers and indicators of tumor progression in melanoma patients. *Cancer Res* 1999;59:1567–1571.

103. Gallego S, Parareda A, Munell F, et al. Clinical relevance of molecular markers in neuroblastoma: results from a single institution. *Oncol Rep* 1999;6:891–896.

104. Brinkschmidt C, Poremba C, Christiansen H, et al. Comparative genomic hybridization and telomerase activity analysis identify two biologically different groups of 4s neuroblastomas. *Br J Cancer* 1998;77:2223–2229.

105. Tomizawa Y, Adachi J, Kohno T, et al. Prognostic significance of allelic imbalances on chromosome 9p in stage I non-small cell lung carcinoma. *Clin Cancer Res* 1999;5:1139–1146.

106. Tomizawa Y, Nakajima T, Kohno T, et al. Clinicopathological significance of Fhit protein expression in stage I non-small cell lung carcinoma. *Cancer Res* 1998;58:5478–5483.

107. Fong KM, Zimmerman PV, Smith PJ. Correlation of loss of heterozygosity at 11p with tumour progression and survival in non-small cell lung cancer. *Genes Chromosomes Cancer* 1994;10:183–189.

108. Scheurlen WG, Schwabe GC, Joos S, et al. Molecular analysis of childhood primitive neuroectodermal tumor defines markers associated with poor outcome. *J Clin Oncol* 1998;16:2478–2485.

109. Afonso A, Emmert-Buck MR, Duray PH, et al. Loss of heterozygosity on chromosome 13 is associated with advanced stage prostate cancer. *J Urol* 1999;162;922–926.

110. Munier FL, Thonney F, Balmer A, et al. Prognostic factors associated with loss of heterozygosity at the RB1 locus in retinoblastoma. *Ophthalmic Genet* 1997;18:7–12.

111. Tagliabue E, Menard S, Robertson JF, et al. c-erbB-2 expression in primary breast cancer. *Int J Biol Markers* 1999;14:16–26.

112. Shivers SC, Wang X, Li W, et al. Molecular staging of malignant melanoma: correlation with clinical outcome. *JAMA* 1998;280:1410–1415.

113. Bostick PJ, Morton DL, Turner RR, et al. Prognostic significance of occult metastases detected by sentinel lymphadenectomy and reverse transcriptase-polymerase chain reaction in early-stage melanoma patients. *J Clin Oncol* 1999;17:3238–3244.

114. Luketich JD, Kassis ES, Shriver SP, et al. Detection of micrometastases in histologically negative lymph nodes in esophageal cancer. *Ann Thorac Surg* 1998;66:1715–1718.

115. Soria JC, Gauthier LR, Raymond E, et al. Molecular detection of telomerase-positive circulating epithelial cells in metastatic breast cancer patients. *Clin Cancer Res* 1999;5:971–975.

116. Kouraklis G. Progress in cancer gene therapy. *Acta Oncol* 1999;38:675–683.

117. Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* 1999;286:487–491.

118. Baselga J, Tripathy D, Mendelsohn J, et al. Phase II study of weekly intravenous trastuzumab (herceptin) in patients with HER2/ *neu*-overexpressing metastatic breast cancer. *Semin Oncol* 1999;26(4 Suppl 12):78–83.

119. Shak S. Overview of the trastuzumab (Herceptin) anti-HER2 monoclonal antibody clinical program in HER2-overexpressing metastatic breast cancer. Herceptin Multinational Investigator Study Group. *Semin Oncol* 1999;26(4 Suppl 12):71–77.

120. Sliwkowski MX, Lofgren JA, Lewis GD, et al. Nonclinical studies addressing the mechanism of action of trastuzumab (Herceptin). *Semin Oncol* 1999;26(4 Suppl 12):60–70.

121. Pegram MD, Slamon DJ. Combination therapy with trastuzumab (Herceptin) and cisplatin for chemoresistant metastatic breast cancer: evidence for receptor-enhanced chemosensitivity. *Semin Oncol* 1999;26(4 Suppl 12):89–95.

122. Goldenberg MM. Trastuzumab, a recombinant DNA-derived humanized monoclonal antibody, a novel agent for the treatment of metastatic breast cancer. *Clin Ther* 1999;21:309–318.

123. Roh H, Hirose CB, Boswell CB, et al. Synergistic antitumor effects of HER2/neu antisense oligodeoxynucleotides and conventional chemotherapeutic agents. *Surgery* 1999;126:413–421.

124. Adenovirus-mediated delivery of PTEN gene inhibits cell growth by induction of apoptosis in endometrial cancer. *Int J Oncol* 1999;15:1069–1074.

125. Rocco JW, Li D, Liggett WH Jr et al. p16INK4A adenovirus-mediated gene therapy for human head and neck squamous cell cancer. *Clin Cancer Res* 1998;4:1697–1704.

126. Swisher SG, Roth JA, Nemunaitis J, et al. Adenovirus-mediated p53 gene transfer in advanced non-small-cell lung cancer. *J Natl Cancer Inst* 1999;91:763–771.

127. Clayman GL, Frank DK, Bruso PA, et al. Adenovirus-mediated wild-type p53 gene transfer as a surgical adjuvant in advanced head and neck cancers. *Clin Cancer Res* 1999;5:1715–1722.

128. Genetic susceptibility to cancer. *Ann ICRP* 1998;28:1–157.

129. Lynch HT, de la Chapelle A. Genetic susceptibility to non-poltposis colorectal cancer. *J Med Genetics* 1999;36:801–818.

130. Terdiman JP, Conrad PG, Sleisenger MH. Genetic testing in hereditary colorectal cancer: indications and procedures. *American Journal of Gastroenterology* 1999;94:2344–2356.

131. Schaid DJ, McDonnell SK, Blute ML, Thibodeau SN. Evidence for autosomal dominant inheritance of prostate cancer. *Am J Hum Genet* 1998;62:1425–1438.

132. Greene MH. The genetics of hereditary melanoma and nevi: 1998 update. *Cancer* 1999;86(11 Suppl):2464–2477.

133. Santoro M, Melillo RM, Carlomagno F, et al. Different mutations in the *RET* gene cause different human tumoral diseases. *Biochimie* 1999;81:397–402.

134. Eng C. RET proto-oncogene in the development of human cancer. *Journal of Clinical Oncology* 1999;17:380–393.

135. Gualberto A, Aldape K, Kozakiewicz K, et al. An oncogenic form of p53 confers a dominant, gain-of-function phenotype that disrupts spindle checkpoint control. *Proc Natl Acad Sci USA* 1998;95:5166–5171.

136. Birch JM, Blair V, Kelsey AM. Cancer phenotype correlates with constitutional *TP53*genotype in families with the Li-Fraumeni syndrome. *Oncogenes* 1998;17:1061–1068.

137. Sevenet N, Sheridan E, Amram D, et al. Constitutional mutations of the hSNF5/INI1 gene predispose to a variety of cancers. *Am J Hum Genet* 1999;65:1342–1348.

138. Claus EB, Schildkraut JM, Thompson WD, et al. The genetic attributable risk of breast and ovarian cancer. *Cancer* 1996;77:2318–2324.

139. Rebbeck TR. Inherited genetic predisposition in breast cancer: a population-based perspective. *Cancer* 1999;86(11 Suppl):2493–2501.

140. Hughes KS, Papa MZ, Whitney T, et al. Prophylactic mastectomy and inherited predisposition to breast carcinoma. *Cancer* 1999;86(11 Suppl):2502–2516.

141. Easton DF, Steele L, Fields P, et al. Cancer risks in two large breast cancer families linked to *BRCA2* on chromosome 13q12-13. *Am J Hum Genet* 1997;61:120–128.

142. de Wert G. Ethics of predictive DNA-testing for hereditary breast and ovarian cancer. *Patient Educ Couns* 1998;1:43–52.

143. Clayton EW, Steinberg KK, Khoury MJ, et al. Informed consent for genetic research on stored tissue samples. *JAMA* 1995;274:1786–1792.

144. Sobel ME. Ethical issues in molecular pathology: paradigms in flux. *Arch Pathol Lab Med* 1999;123:1076–1078.

10

Molecular Biology of Infectious Diseases

Ramon L. Sandin

Within the field of diagnostic pathology, the areas of clinical microbiology and clinical virology were the first to be strongly influenced by the use of DNA probes and amplification techniques. Numerous reports have been published on the detection of pathogens of different types by DNA probes and amplification techniques, and several examples are provided in Table 10.1, along with references.

TABLE 10.1. SELECTED EXAMPLES OF MICROBES DETECTED WITH THE USE OF DNA PROBES AND AMPLIFICATION TECHNIQUES

Type	Examples (References)
Bacteria	*Legionella spp.* (103,104)
	Group A streptococci (2)
'Higher' bacteria	*Nocardia spp.* (105)
Mycobacteria	*Mycobacterium tuberculosis* (37,106)
	Mycobacterium avium (6)
	Mycobacterium intracellulare (6)
Fastidious or noncultivable organisms	*Coxiella burnetii* (107)
Fungi	*Aspergillus spp.* (108)
	Histoplasma capsulatum (109)
	Blastomyces dermatitidis (110)
	Pneumocystis carinii (111)
	Cryptococcus neoformans (108)
Viruses	*Herpes simplex* (112)
	Human Immunodeficiency Virus (93,94)
	Human Cytomegalovirus (64,65,93)
Parasites	*Toxoplasma gondii* (113)

Prior to the development of synthetic amplification techniques such as the polymerase chain reaction (PCR), DNA probes were prepared either "in-house" or commercially and were used directly on clinical specimens for microbial detection (1, 2). The advent of nucleic acid probe technology offered the promise of speedy and specific direct microbial detection in a clinical specimen that may obviate culture. However, laboratory studies showed that if the number of target molecules in a clinical sample is low, the sensitivity of either "in-house" ("home-brew") or of commercially available nucleic acid probes may be unacceptably low (3). It became clear to those working in the field of nucleic acid probe technology that in order to increase the sensitivity of the methods used for detecting the presence of infectious agents in a clinical specimen, "amplification" of a detectable "entity" may have to occur (4). "Amplification" means making many entities from a few. Such entities may be whole microbes. Amplification of intact microbes via culture has been used in microbiology laboratories for a long period of time. Nucleic acid probes have been used for several years to confirm the nature of a culture-amplified isolate (5). For mycobacterial species, the only commercialized probe assay system today is the AccuProbe System (Gen-Probe Inc., San Diego, California), which is based on detection of ribosomal ribonucleic acid (rRNA) (6) by way of hybridization to single-stranded DNA probes. Instead of the Iodine125-labeled probes used by the earlier Gen-Probe system, AccuProbe uses a chemiluminescent, acridinium ester-labeled probe detected with a luminometer, with a sensitivity of greater or equal to 95% (6). Gen-Probe probes for culture confirmation (7) are available for *Mycobacterium avium* complex, *Mycobacterium tuberculosis* (MTB) complex, *Mycobacterium gordonae*, *Mycobacterium kansasii*, fungi, and bacteria. Probes also are available for direct detection of group A streptococci (2), *Chlamydia trachomatis* (Pace 2 System, Gen-Probe) (8), and *Neisseria gonorrhoeae*. Unfortunately, culture confirmation requires prior growth of the organism in culture either in solid media or in the faster automated broth media (9), where growth still may entail a lengthy period of time, such as several weeks as in MTB, which may not be practical.

TABLE 10.2. CLINICAL SCENARIOS IN INFECTIOUS DISEASES WHERE THE USE OF PCR AND OTHER AMPLIFICATION TECHNIQUES MAY BE ADVANTAGEOUS

Clinical Scenario	Example(s)
Detection of slow-growing organisms	*Mycobacterium tuberculosis*, agents of deep mycoses
Detection of organisms that are highly infectious or dangerous to culture	HIV, *Coxiella burnetii, Coccidioides immitis*
Detection of agents involved in encephalitis where a rapid and sensitive diagnosis may obviate obtaining a brain biopsy	Herpes simplex virus, enterovirus
Detection of microbes for which routine, commercially available culture methods are lacking in diagnostic laboratories	*Treponema pallidum*, 37°C phase of *Coccidioides immitis, Toxoplasma gondii, Pneumocystis carinii*
Detection in asymptomatic newborns where serology may be misleading	Congenital HIV infection
Detection of viruses in invasive specimens (CSF, peripheral blood, bone marrow aspirates, etc.) from patients undergoing bone marrow or solid organ transplantation where greater sensitivity is advantageous	Parvovirus, enterovirus or CMV
Determination of the load of virus or other pathogen for prognostic purposes or for monitoring therapeutic drug effectiveness	HIV quantitation
Determination of the level of expression (i.e., transcription) of a pathogen's DNA by direct measurement of RNA	RT-PCR for CMV, HSV
Detection of antimicrobial resistance genes by amplification of sequences known to mediate such resistance	Isoniazid resistance in *M. tuberculosis*
Reclassification of previously uncultured organisms into the correct position in their genealogical tree	Sequencing of amplified 16s rRNA by RT-PCR or 16s rDNA targets by PCR
Determination of the safety or infectivity of blood products	Presence of HIV or HTLV I/II in blood products

Adapted with permission from Sandin RL, Greene JN. Diagnostic molecular pathology and infectious disease. *Cancer Control*. 1995; 2: 250
Abbreviations: CMV, cytomegalovirus; CSF, cerebrospinal fluid; HIV, Human Immunodeficiency Virus; HSV, herpes simplex virus; HTLV I/II, human T-lymphotropic virus type I and type II; PCR, Polymerase chain reaction; RT, reverse transcription.

If the amplified "entity" is a submicrobial chemical molecule such as DNA or ribonucleic acid (RNA), it may lend itself to a more rapid synthetic amplification process than culture (4). Thus, when the PCR and other non-PCR amplification techniques were discovered and developed, the world of microbiology was boosted even further into a new flurry of activity. Thousands of papers have been published documenting the use of PCR and non-PCR amplification techniques for detection of all types of microorganisms. The extreme sensitivity of these techniques may not be necessary for detection of many of the myriad of organisms on which these techniques have been tested, however. Selected clinical scenarios do exist in clinical microbiology and virology where the use of PCR or non-PCR amplification techniques may be so advantageous that application of such techniques will become the standard-of-care (10). Such is the case already with determination of human immunodeficiency virus

(HIV) viral loads in plasma specimens of acquired immunodeficiency syndrome (AIDS) patients for determination of prognosis and therapeutic drug effectiveness. Table 10.2 summarizes the clinical scenarios where amplification techniques may become indispensable within infectious diseases in the near future.

As technologies such as nucleic acid probes for culture confirmation and direct nucleic acid amplification techniques become more widely used, it becomes easier for agencies such as the Centers for Disease Control and Prevention (CDC) to issue new recommendations about improved turn-around time of cultures and stains, which routine diagnostic laboratories can aspire to fulfill. For example, the CDC recommends that a laboratory deliver results of acid-fast bacilli (AFB) smears within 24 hours of specimen collection, and that growth and identification of MTB in clinical specimens be completed, on the average, within 21 days of specimen collection (11, 12). In those laboratories where such recommendations are met, advanced technologies such as probes and amplification undoubtedly are in use.

AMPLIFICATION TECHNIQUES AT VARIOUS STAGES OF DEVELOPMENT, FDA-APPROVAL STATUS OF VARIOUS COMMERCIAL ASSAYS IN INFECTIOUS DISEASES, AND ASSAYS UNDERGOING DEVELOPMENT

Part of "10 - Molecular Biology of Infectious Diseases"

The amplifiable entities most commonly used in synthetic amplification assays include the following: target DNA as in the PCR; target RNA as in reverse transcriptase PCR (RT-PCR); amplification of a signal or reporter molecule as in the Branched DNA Signal Amplification Assay (Chiron Corporation, Emeryville, California); the ligation product of two or more probes as in the gapped ligase chain reaction (LCR, Abbott Laboratories, Abbott Park, Illinois); target RNA via a transcription-mediated system as in the Gen-Probe TMA systems (Gen-Probe, San Diego, California), and several others. Simplified reviews of these techniques can be found in Wolcott (13) and Sandin (4, 14).

TABLE 10.3. LIST OF SEVERAL COMMERCIAL PCR AND NON-PCR AMPLIFICATION SYSTEMS AT DIFFERENT STAGES OF DEVELOPMENT

Target Amplification Systems

- PCR by Roche under the name Amplicor series (Roche Diagnostic Systems, Branchburg, New Jersey)
- Transcription-mediated amplification systems
  —Examples
  -
    - Gen-Probe rRNA isothermal amplification system (Gen-Probe, San Diego, California) and VIDAS (VITEK, St. Louis, Missouri) adaptation (VIDAS PROBE)
    - NASBA (Nucleic acid sequence-based amplification) isothermal target amplification system by Organon-Teknika (Durham, North Carolina)
    - Strand Displacement Amplification (BDProbe Tec ET) (Becton-Dickinson Diagnostic Instrument Systems, Sparks, Maryland)

Signal Amplification Systems

- Branched DNA Signal Amplification System by Chiron Corporation (Emeryville, California)
- Hybrid Capture System by DiGene Diagnostics, Inc. (Silver Spring, Maryland)

Probe Amplification Systems

- Ligase Chain Reaction by Abbott Corporation (Abbott Park, Illinois)
- Cycling Probe Technology by ID Biomedical (Vancouver, Canada)
- Q-beta Replicase, formerly by Gene-Trak, Inc. (Framingham, Massachusetts)

Several other techniques undergoing research and development

Adapted with permission from Sandin RL. The ABC's of molecular amplification. *Infect Med* 1999; 16:98 and Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In, Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders CO; 1996:617

TABLE 10.4. FDA-APPROVED ASSAYS IN INFECTIOUS DISEASES UTILIZING MOLECULAR AMPLIFICATION TECHNIQUES (updated as of July 1, 1999)

Roche Molecular Systems
A) HIV-1

- Amplicor HIV-1 Monitor 1.0, 400 copies detection limit, microwell plate format (MWP)
- Amplicor HIV-1 Monitor 1.0, Ultrasensitive, 50 copies detection limit, MWP

B) *Mycobacterium*

- Amplicor *Mycobacterium tuberculosis*, MWP

C) *Chlamydia trachomatis*

- Amplicor *C. trachomatis*, MWP format
- COBAS Automated system for *C. trachomatis*

Gen-Probe Transcription-Mediated Amplification Systems (TMAs)
A) *Mycobacterium*

- AMTDT-1 (Amplified *M. tuberculosis* Direct Test, version 1)
- AMTDT-2 (enhanced AMTDT or E-MTD)

B) *Chlamydia trachomatis*

- Amplified *C. trachomatis* Direct Test

Digene Diagnostics Hybrid Capture Systems
A) Human Cytomegalovirus

- CMV Hybrid Capture 1 (tube format), version 1– 5,000 copies/mL of whole blood detection limit, qualitative
- CMV Hybrid Capture I version 2–700 copies/mL of whole blood detection limit, qualitative

B) Human Papillomavirus

- HPV Hybrid Capture I (tube format)
- HPV Hybrid Capture II (MWP)

Abbott Corporation (Ligase Chain Reaction or LCx) Systems
A) *Chlamydia trachomatis*

- LCx *C. trachomatis*, male and female urine and genital swabs

B) *Neisseria gonorrhoeae*

- LCx *N. gonorrhoeae*, male and female urine and genital swabs

Target, probe, and signal amplification systems have undergone intense research and development efforts in the last several years. Table 10.3 offers a list of several commercial PCR and non-PCR amplification systems at different stages of development. Table 10.4 offers a list of Food and Drug Administration (FDA)-approved assays in infectious diseases utilizing several of these molecular amplification techniques, and Table 10.5 offers a list of molecular amplification assays in infectious diseases undergoing research and development efforts by the various manufacturers at the time of this writing. It is expected that many more assays will undergo research and development efforts in the near future. The most widely used synthetic amplification technique is the target amplification system known as PCR. Most in-house or home-brew PCR assays have been developed by academic scientists at their respective institutions. Commercial PCR-based assays by the Roche Molecular Diagnostic Systems (Branchburg, New Jersey) have appeared in the market and will continue to do so in the future. In contrast, most of the research and development on alternate non-PCR amplification techniques that are now becoming available commercially has resided predominantly or exclusively

within industry and companies. Some of these signal, probe, and non-PCR target techniques promise to have some potential advantages over in-house testing by PCR, namely (i) less prone to carryover contamination; (ii) more amenable to quantification; (iii) high specificity; and (iv) an increased ease of automation. The advent of technologies other than in-house PCR testing should allow for strong competition between these techniques to occur in an attempt by companies to secure a place within the diagnostic repertoires of clinical laboratories.

TABLE 10.5. MOLECULAR AMPLIFICATION ASSAYS IN INFECTIOUS DISEASES NOT YET FDA-APPROVED IN THE UNITED STATES, BUT UNDERGOING RESEARCH AND DEVELOPMENT; AVAILABLE FOR USE OUTSIDE THE UNITED STATES ONLY; OR AVAILABLE IN THE UNITED STATES FOR RESEARCH PURPOSES ONLY (I.E., IN CLINICAL TRIALS, OR SUBMITTED TO THE FDA BUT NOT YET APPROVED) (updated as of July 1, 1999)

Roche Molecular Systems

A) HIV-1

- Amplicor HIV-1 Monitor 1.5, with expanded subtype detection capability, MWP
- Amplicor HIV-1 Monitor 1.5, Ultrasensitive
- COBAS automated system for HIV-1
- Qualitative MWP format for HIV-1 DNA amplification

B) *Mycobacterium*

- Amplicor *Mycobacterium avium-complex* (MAC), MWP format
- Amplicor *Mycobacterium* genus-specific system, MWP format
- COBAS automated system for mycobacterial amplification

C) *Chlamydia trachomatis*

- Amplicor *C. trachomatis*/*Neisseria gonorrhoeae*, COBAS and MWP formats

D) Hepatitis C

- Amplicor Hepatitis C, MWP format, qualitative and quantitative
- COBAS automated system for Hepatitis C

E) Others

- Amplicor HTLV I/II, MWP format
- Amplicor CMV, qualitative, MWP format
- COBAS automated system for CMV, quantitative (CMV Monitor)
- Ampli-Prep automated sample preparation system
- Real-time quantitative fluorescent amplification and detection system (TaqMan probes and instrument)

Gen-Probe Transcription-Mediated Amplification Systems (TMAs)

A) *Chlamydia trachomatis*/*Neisseria gonorrhoeae*

- Amplified *C. trachomatis*/*N. gonorrhoeae* Direct test

B) HIV-1

- Amplified HIV-1 Direct test, quantitative

C) Hepatitis C

- Amplified Hepatitis C Direct test, quantitative

D) HIV-1/HCV

- Multiplex amplification of HIV-1 and Hepatitis C, for Blood Bank use ONLY

E) Hepatitis B

- Amplified Hepatitis B Direct test

F) VIDAS PROBE (adaptation of amplification and detection steps of TMA to the VIDAS instrument from Bio-Merieux VITEK)

- *M. tuberculosis*
- *C. trachomatis*
- *C. trachomatis*/*N. gonorrhoeae*
- *N. gonorrhoeae*
- HIV-1 RNA, quantitative

G) TIGRIS Instrument for automation of extraction, amplification and detection steps

H) Atypical pneumonia panel

I) Biochip Technology (Microparticle arrays) adaptation of the TMA technology, in conjunction with Affymetrix, a biochip company.

Digene Diagnostics Hybrid Capture Systems

A) Human Cytomegalovirus

- CMV Hybrid Capture 1 version 2, Quantitative application

B) *Chlamydia trachomatis* and *Neisseria gonorrhoeae*

- *C. trachomatis* Hybrid Capture assay
- *N. gonorrhoeae* Hybrid Capture assay
- *C. trachomatis*/*N. gonorrhoeae* Hybrid Capture assay

C) Other viral assays

- Hepatitis B Hybrid Capture assay
- Herpes simplex virus I/II Hybrid Capture assay
- HIV-1 RNA Hybrid Capture assay
- HIV-1 DNA (provirus) Hybrid Capture assay

Abbott Corporation (Ligase Chain Reaction or LCx) Systems

A) HIV-1

- LCx HIV-1 viral load, quantitative

B) *Mycobacterium*

- LCx *Mycobacterium tuberculosis*, qualitative

C) Other assays

- LCx Hepatitis B assay, quantitative

Organon-Teknika (NASBA) Systems

A) HIV-1

- NucliSens HIV-1 QT assay, quantitative, 400 copies/mL detection limit with 1 mL of plasma, and 4,000 copies/mL detection limit with 100 uL of plasma

B) Human Cytomegalovirus

- NucliSens CMV pp67 mRNA assay

C) Other assays or instruments

- NucliSens Extractor, automated nucleic acid extraction instrument
- NucliSens Isolation kit
- Detection kit for in-house amplification products

Becton-Dickinson Strand Displacement Amplification (BD Probe Tec ET) Systems

A) *Chlamydia trachomatis* and *Neisseria gonorrhoeae*

- BD Probe Tec ET *C. trachomatis*, direct detection from urine and urethral swabs, 510(k) submitted to FDA on 12/98
- BD Probe Tec ET *C. trachomatis* and *N. gonorrhoeae*, differential direct detection from urine and urethral swabs, 510(k) submitted to FDA on 12/98

B) *Mycobacterium*
- BD Probe Tec ET *M. tuberculosis*, direct detection from respiratory specimens, in clinical trials
- BD Probe Tec ET *M. tuberculosis*, for culture confirmation, in clinical trials
- BD Probe Tec ET *M. kansasii*, for culture confirmation, in clinical trials
- BD Probe Tec ET *Mycobacterium avium* complex (MAC), for culture confirmation, in clinical trials

C) Other assays
- Several viral detection and quantification kits

Bayer Diagnostics Branched DNA Signal Amplification Systems (bDNA formerly by Chiron Corporation)

A) HIV-1
- Quantiplex bDNA HIV-1 viral load 1.0 (10,000 copies detection limit)
- Quantiplex bDNA HIV-1 viral load 2.0 (500 copies detection limit)
- Quantiplex bDNA HIV-1 viral load 3.0 (50 copies detection limit)

B) Hepatitis B
- Quantiplex bDNA Hepatitis B viral load 1.0

C) Hepatitis C
- Quantiplex bDNA Hepatitis C viral load 1.0
- Quantiplex bDNA Hepatitis C viral load 2.0
- Quantiplex bDNA Hepatitis C viral load 3.0
- HCV qualitative assay using Gen-Probe TMA technology

Abbreviations: CMV, cytomegalovirus; FDA, Food and Drug Administration; HIV, Human Immunodeficiency Virus; MWP, Microwell Plate; NASBA, Nucleic Acid Sequence Based Amplification.

PCR is the best studied, the most frequently used, and the most widely published amplification technique, almost certainly because it was the first technique developed, and all original PCR assays followed in-house or home-brew protocols. Thus, the specific technical considerations, definitions, and workings of in-house PCR will be presented first in this chapter. Because abundant work has been published on in-house PCR and on stereotypical commercial applications of PCR and non-PCR techniques for detection of MTB in clinical specimens, MTB has been chosen as the prototype organism with which to describe such particulars in this chapter. Another microorganism that also has been the subject of innumerable studies in the development, validation, and availability of commercial amplification techniques, both PCR and non-PCR, is HIV. Unfortunately, given space constraints, all discussion on progress, technical considerations, and results of large studies of the various PCR and non-PCR amplification techniques will be focused on MTB. Because this chapter resides in the Molecular Pathology section of this book, it was intentionally written to provide abundant technical detail on the actual workings and steps involved in “running” a diagnostic molecular assay, and so can serve as a technical manual. *Chlamydia*/*Neisseria*, CMV, and hepatitis C and B also have been the subject of tremendous research and development efforts as documented in the published literature. Various aspects of these and other organisms will be considered as part of Table 10.1, Table 10.2, Table 10.3, Table 10.4, and Table 10.5, and Table 10.18

MYCOBACTERIUM TUBERCULOSIS

Part of "10 - Molecular Biology of Infectious Diseases"

MTB will be used to illustrate the “growing pains” that the field of PCR amplification technology has experienced in developing effective assays, the attendant evolution in procedures, and the problems associated with sensitivity and specificity of first-generation home-brew as well as commercial assays. The importance of this bacillus to our present world cannot be underappreciated. It is estimated that as much as one third of the world's population has been infected with MTB, and there are approximately eight million new cases of tuberculosis annually (15). The incidence of tuberculosis is constantly increasing, and strains of MTB resistant to chemotherapeutic agents have been recovered with increasing frequency. It is for that reason that amplification techniques that can expedite the diagnosis of tuberculosis in clinical specimens are of such crucial importance to clinical medicine.

TABLE 10.18. CPT CODES (1998) INTENDED FOR USE WITH MOLECULAR DIAGNOSTIC TECHNIQUES FOR ANALYSIS OF NUCLEIC ACIDS, WITH EMPHASIS ON THE FIVE INFECTIOUS AGENTS MOST FREQUENTLY AMPLIFIED FROM CLINICAL SPECIMENS

CPT Code	Description
Specific codes for Infectious Disease Identification	
87490	*Chlamydia trachomatis*, direct probe technique
87491	*Chlamydia trachomatis*, amplified probe technique
87492	*Chlamydia trachomatis*, quantification
87495	Cytomegalovirus, direct probe technique
87496	Cytomegalovirus, amplified probe technique
87497	Cytomegalovirus, quantification
87520	Hepatitis C, direct probe technique
87521	Hepatitis C, amplified probe technique
87522	Hepatitis C, quantification
87534	HIV-1, direct probe technique
87535	HIV-1, amplified probe technique
87536	HIV-1, quantification
87537	HIV-2, direct probe technique
87538	HIV-2, amplified probe technique
87539	HIV-2, quantification
87550	Mycobacteria species, direct probe technique
87551	Mycobacteria species, amplified probe technique
87552	Mycobacteria species, quantification
87555	*Mycobacterium tuberculosis*, direct probe technique
87556	*Mycobacterium tuberculosis*, amplified probe technique
87557	*Mycobacterium tuberculosis*, quantification
87560	*Mycobacterium avium-intracellulare*, direct probe technique
87561	*Mycobacterium avium-intracellulare*, amplified probe technique
87562	*Mycobacterium avium-intracellulare*, quantification
General Codes for Infectious Disease Identification	
87797	Infectious agent detection by nucleic acid (DNA or RNA); not otherwise specified, direct probe technique
87798	Not otherwise specified, amplified probe technique
87799	Not otherwise specified, quantification
General codes for nucleic acid amplification	
83890	Molecular diagnostics; molecular isolation or extraction
83892	Enzymatic digestion
83894	Separation (e.g., dot blot, electrophoresis)
83902	Reverse transcription
83912	Interpretation and report

Abbreviations: HIV, human immunodeficiency virus, From CPT Codebook, 1998

Polymerase Chain Reaction, In-House (Home-Brew) Assays, Technical Considerations

Definition and Crucial Steps

PCR can be defined simply as a synthetic method of creating many copies of discrete fragments of DNA present in a clinical specimen, thus facilitating detection of nucleic acids present initially only in very small quantities (i.e., picograms or femtograms) (16, 17). A technical definition of the method would be that it is a semi-automated amplification of DNA *in vitro* by a series of successive incubation steps at different temperatures using a heat-stable DNA-dependent DNA polymerase. There are three steps to amplification by PCR: (i) denaturation; (ii) primer annealing; and (iii) extension. These can be carried out using three different temperatures or the assay may be set up so that two temperatures are used. In step one, the denaturation step, double-stranded DNA is melted or denatured by heating at a high temperature, such as 94°C for a finite period of time, such as 1 minute. In step two, the primer annealing step, primers are allowed to bind to the single-stranded DNA at a lower temperature, such as 52°C for 2 minutes. Primers are crucial to the specificity of the assay because they are small fragments of DNA that flank or delimit the target sequence that is to be amplified and subsequently detected. In step three, the extension step, a new strand is synthesized by the polymerase using single nucleotides as building blocks, starting out at the site of the primer, and "filling in the gap" that exists between the target sequence strand and the primer. An intermediate temperature usually is chosen for step three such as 72°C for a period such as 3 minutes. Steps one to three are called a "cycle." Repetition of cycles yields amplification of the target DNA. Under optimal conditions,10 cycles would lead to approximately 1,000 copies of the target regions and 20 cycles to approximately one million copies. The amplified product subsequently is detected by using the more traditional and common methods for detection, or by resorting to one of the novel ones; both approaches are reviewed later.

Thermocyclers commonly used for PCR utilize metal blocks or water for thermal equilibration, and samples are contained in plastic microcentrifuge tubes or microtiter wells. Average temperature transition rates for heat block thermocyclers are about 1°C per second (18). Thus, a considerable amount of time is spent heating and cooling the sample. A new alternative is the use of a rapid air thermal cycler (Air Thermal Cycler (ATC), Idaho Technology, Idaho Falls, Idaho) (18, 19). It uses high velocity air for heating and cooling. Samples are contained in microcapillary tubes with a high surface area-to-volume ratio. Comparison of air vs. heat block thermal cyclers for smear-positive MTB culture positive specimens showed high sensitivity for both types of thermal cyclers (18); additionally, the air thermal cyclers helped reduce overall cost and total assay time (18, 19).

There are four crucial steps in performing in-house PCR: lysis or extraction of the nucleic acids which one aims to amplify in order to make them accessible to the polymerase; choice of the most appropriate target area of the genome to be amplified; choice of the most appropriate amplification scheme to be used during the process; and determination of the most appropriate method of detection of the amplified product as regards the size and complexity of the laboratory where the work is being performed. Each of these crucial steps now will be examined as it regards MTB. It will become obvious to the reader that because of the enormous variety of methods and schemes available to the user for amplifying this or any other microorganism, comparing results from various studies to one another may be akin to comparing apples to oranges.

Methods of Cell Lysis, Extraction or Purification of Nucleic Acids

The majority of the early studies used the "traditional" and more laborious methods of cell lysis, extraction, and purification of nucleic acids before amplification. This was particularly true in studies through 1992. Since then, more novel and simpler adaptations have been reported in subsequent studies with varying results. The basics of the more traditional method of extraction of nucleic acids in mycobacteriology involves the following steps. Sputum samples are treated with NaOH and N-acetyl-L-cysteine (NALC) for digestion-decontamination. Cell lysis is accomplished by treatment with sodium dodecyl sulfate (SDS) or NaOH, which is followed by several extractions with phenol-chloroform-isoamyl alcohol. The extract may be dialyzed for purity in water and then concentrated by ethanol precipitation followed by dilution in water or buffer prior to being used for amplification by PCR (20, 21, 22).

An early successful and widely used adaptation of this basic traditional method is that of Eisenach et al. (23). In this method, cell lysis occurs by incubation with lysozyme for 2 hours, followed by treatment with SDS, NaOH, and NaCL. The nucleic acid subsequently is purified by passage through powdered glass or by spin column chromatography. Newer, simpler methods have been reported that are expected to challenge the traditional ones in terms of sensitivity, specificity, and speed. Some results are conflicting, however, as will be presented.

TABLE 10.6. MOST COMMON METHODS OF CELL LYSIS, EXTRACTION AND PURIFICATION OF MYCOBACTERIAL NUCLEIC ACIDS FOR "IN-HOUSE" PCR AMPLIFICATION

Method	References
Lysozyme, SDS, NaOH and NaCl	23
Nonionic detergents and proteinase K	24, 25
Boiling with nonionic detergents without proteinase K	24
Freeze-thaw cycles	24
Sonication and boiling	24, 25
Triton X-100 detergent and heat	26
Triton X-100, NaOH, proteinase K, and heat	92
Boiling alone	27
Centrifugation through 50% sterile sucrose solution	28
Guanidinium thiocyanate chaotrope and silica particles	29

Adapted and updated, with permission from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617

Table 10.6 summarizes some of the most common extraction and purification methods for PCR amplification of mycobacteria, with references. Buck et al. (24) evaluated several simple

methods against the more traditional and laborious extraction method. Included among these variations was the use of nonionic detergents and proteinase K for lysis and extraction; boiling samples with nonionic detergents without the use of proteinase K; submitting the samples to eight freeze-thaw cycles; and sonication by flotation of specimens on a dish in water next to a sonicator probe (Sonics and Materials, Danbury, Connecticut) for 30 minutes followed by boiling. These authors report that sonication produced the best results, compared well with the reference method, and was able to allow detection of as little as 10 to 100 organisms. In contrast, Folgueira et al. (25) compared sonication with the use of nonionic detergents and proteinase K and found conflicting results. They found that enzymatic detergent treatment yielded better results than sonication. Clarridge et al. (26) used Triton X-100, a detergent solution, and heat (100°C) without further DNA extraction in treating samples received in a very busy diagnostic laboratory with good results. Kocagoz et al. (27) boiled clinical specimens for 10 minutes and used them directly without further processing. They report a sensitivity of 10 organisms. Stronger bands were obtained after boiling than with the use of lysozyme-proteinase K-SDS lysis followed by phenol-chloroform-ethanol extraction. Victor et al. (28) reported on a simple procedure for the mass screening of sputa. It entailed decontamination of sputa and purification from inhibitors by passing the specimen through a 50% sterile sucrose solution via a 5-minute centrifugation step. They used the pellet straight for PCR with acceptable results. Wilson et al. (29) used guanidinium thiocyanate, a chaotrope, to solubilize specimens and captured the released DNA with silica particles. The DNA was then eluted with water. Sensitivity of PCR with this method was 75%.

Thus, results are very variable and, at times, contradictory. The lack of standardization in extraction methodology for in-house assays may be the culprit. It is recommended that the scientist try one or several of the reported methods, adapt them to his or her particular needs, and perhaps simplify them to suit the constraints of a given laboratory. The author has tried several of the methods and has found modifications of the Eisenach et al. method (23, 30) to be useful. A particular laboratory must evaluate several methods and decide which works best at a particular institution (31).

Most Common Gene Targets for Amplification

A target may be a gene sequence that is present as a single copy in the MTB genome, or one that is present as a repeated sequence. Decisions as to which target to use for amplification carry with it implications regarding the sensitivity of the assay. One also must decide how broad a specificity is desired (i.e., amplification of all mycobacteria, of only organisms within the MTB-complex, or more specifically of only MTB). Table 10.7 summarizes these choices regarding types of targets for amplification and provides various references.

TABLE 10.7. GENE TARGETS FOR PCR AMPLIFICATION OF MYCOBACTERIUM TUBERCULOSIS

Targets	References
A) In terms of copy number	
1. Single copy DNA gene sequences	20, 22, 32, 33
2. Repeated DNA sequences	21, 23, 35, 36, 39, 114
3. Ribosomal RNA sequences via RT-PCR	47, 48
B) In terms of specificity	
1. Specific for the entire genus *Mycobacterium*	22, 50, 51, 60, 68, 115, 116, 117
2. Specific for the MTB-complex (MTB, *M. bovis, M. bovis* BCG strain, *M. microtii*, and *M. africanum*)	23, 34, 35, 36, 51, 52,68, 118
3. Specific for the species *M. tuberculosis*	34, 42, 51, 53, 54, 55,56, 58
4. Specific for the *M. bovis* BCG vaccine strain	61

Adapted and updated, with permission, from Sandin RL. Polymerase Chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617

The use of single-copy genes as targets is more risky than the use of targets that constitute repeated DNA or RNA sequences because the sensitivity of the assay is innately reduced in the former when compared with amplification of targets that are preamplified in the organism. Additionally, a mutation in the single-copy target gene may lead to false-negative results. Brisson-Noel et al. (20) used a single-copy sequence as the target in their assay, the gene coding for a part of the 65 kd antigen present in organisms of the MTB-complex (22). Because the sensitivity of this assay was not very high, this same group subsequently improved it by using a two-step nested PCR approach, an approach to amplification, which will be discussed in a subsequent section. These authors used 50 cycles of amplification for each of the two rounds of amplification. Hermans et al. (32) used a 158 base pair single-copy fragment located near a repetitive sequence. Five of 34 isolates of MTB from cultures appeared not to have the target copy and did not amplify, thus producing false negatives. Miyazaki et al. (33) showed increased sensitivity by using a nested PCR approach for amplification of a single-copy gene from the 38-kd protein gene of mycobacteria.

A more common approach is the use of repeated DNA or RNA sequences (34). Eisenach et al. (23, 35) and Thierry et al. (36) have used the IS6110 sequence as a target for amplification. This is a repetitive DNA sequence of 1361 base pairs in length that has been described, albeit with controversy, to be specific for the MTB-complex and present in high copy number in MTB (10 to 20 copies) but in low copy number in *M. bovis* (1 to 5 copies). Because there are several copies of this sequence per organism, it serves as a “preamplified” sequence, thus increasing the sensitivity of a PCR assay (34). An additional advantage of using repeated sequences is that deletion or mutation of one of the sequences should still allow for others to be intact for amplification. The IS6110 has been perhaps the favorite, or at least the most frequently utilized, target sequence for in-house amplification of MTB in reports to date in the world literature (34, 37, 38). However, it must be mentioned that von Soolingen et al. (39) reported on one strain of MTB that was exceptional in that it lacked the IS6110 sequence. This is felt to be rare but suggested to the authors that the widely accepted practice of using this element for amplification may have to be revisited. Yet

others have challenged the specificity of this sequence and the notion that it is found solely in the MTB-complex organisms (40, 41, 42). The heated debate that has ensued has produced a flurry of recent publications where groups of investigators have challenged one another's assertions (38, 40, 41, 42, 43, 44 and 45). It now appears that there are selected regions of the IS6110 that are MTB-complex-specific targets (38, 43), while others may not be, and that care must be taken in choosing which regions to use for amplification assays. Homology to other species outside of the MTB-complex may be restricted to the central 181-base pair region of IS6110 (43). The use of nucleotides 762 to 865 and 970 to 1026 of IS6110 has been reported as specific to the MTB-complex (38, 43). Table 10.8 summarizes the disadvantages raised in the literature as it regards the use of IS6110 for PCR amplification of mycobacteria.

TABLE 10.8. DISADVANTAGES RECENTLY RAISED AGAINST THE USE OF THE REPEATED SEQUENCE IS6110

Disadvantage	References
Producing possible false negative results	
• Rare MTB strains (originally from South East Asia and India) may lack IS6110	39, 119
• A few strains have only one copy of IS6110	120
Producing possible false positive results	
• Homology has been reported to exist with mycobacteria other than MTB	40, 41, 42, 44, 45

Abbreviations: MTB, *Mycobacterium tuberculosis*

Investigators also have resorted to the use of ribosomal RNA sequences as targets that can be amplified via the RT-PCR. There are advantages to the use of RNA as a target. 16S rRNA sequences are found in large copy numbers (1,000 to 10,000 molecules/cell) in MTB and thus would serve as a preamplified target for PCR amplification. Some stretches of rRNA are highly variable among species and would serve for the delineation of species-specific amplification schemes, whereas other stretches are conserved and would serve as targets for genus-specific amplification (46). rRNA sequences also serve as a yardstick for the evolutionary stage of a particular microorganism regarding other organisms and can be used for such studies. Thus, Boddinghaus et al. (47) amplified conserved regions of mycobacterial nucleic acids via RT-PCR and then determined the specific species amplified with the use of probes to species-specific variable regions. A similar approach was used by Patel et al. (46), but amplification was of the 16S rRNA gene (i.e., 16S rDNA) via PCR, rather than 16S rRNA via RT-PCR. Kirschner et al. (48) followed amplification with sequencing of the amplified product to determine which species was amplified. Messenger RNA (mRNA) is extremely short-lived, averaging only a few minutes, and has been used in RT-PCR formats to detect viable mycobacteria (49), whereas DNA targets do not distinguish between viable and nonviable organisms.

Primers also can be designed to suit the desired degree or breadth of specificity of a given investigator. Some workers have set out to amplify any organism within the genus *Mycobacterium* present in a clinical sample, while others set out to amplify only those within the MTB-complex, whereas still others have designed primers only for MTB. Targets such as the DNA J gene are genus specific (22, 50) and must be followed by species-specific probes. Sequences such as the IS6110 (23, 35, 36, 51) and others (52) are repeated sequences reported to be specific to the MTB-complex. One group from South America (53) emphasized their need to distinguish between human tuberculosis of bovine origin (i.e., as a result of *M. bovis*) from tuberculosis resulting from MTB. This group designed primers homologous to sequences within the mtp40 region (53), which is reported to be specific for MTB, and helps distinguish this species from other closely related species within the MTB-complex (53 to 56) (see Table 10.7). It is part of the mpcA gene, which encodes a phospholipase C protein (57). A rapid method of differentiating MTB from *M. bovis* would be helpful clinically because of the intrinsic resistance of *M. bovis* to pyrazinamide, and epidemiologically because risk factors for *M. bovis* infection may be different from those for MTB infections (58). Thus, the discovery of the usefulness of the mtp40 gene sequences for such purposes was welcome. Unfortunately, the mtp40 region is yet another sequence that has been the subject of controversy, this time as it regards its sensitivity. The mtp40 gene is not present in all strains of MTB and, hence, may not be useful for differentiating MTB from *M. bovis* in all cases of interest (57, 58). A different scheme recently suggested for the differentiation of MTB from *M. bovis* is an oxyR allele-specific amplification method (59).

Additional recent studies continue to shed light on the availability of mycobacterial sequences of various specificities to suit the needs of those using amplification schemes. The 32-kD protein gene described recently is found in all mycobacterial species, contains diversity between different species, and only some intraspecies variability has been detected (60). Thus, it has been suggested as a good target for identification of mycobacteria by PCR-based sequencing. Other recent studies have shed light on sequences specific for *M. bovis* and for the vaccine strain *M. bovis* Bacillus Calmette-Guerin (BCG), which can be used in PCR-based assays (61). The BCG is an attenuated derivative of a virulent strain of *M. bovis*. BCG has been used as a vaccine against MTB, as cancer immunotherapy, and as a recombinant vehicle for multivalent vaccines against other infectious diseases (61). BCG can cause disease in humans, especially those with cellular immunodeficiencies (62). Therefore, the ability to rapidly and specifically identify BCG is clinically important. The genomic region designated RD1 was found to be present in all virulent *M. bovis* and MTB strains tested, but deleted from all BCG strains evaluated (61). A multiplex PCR with primers to the MTB complex and to the RD1 region would be useful in identifying sequences from the *M. bovis* BCG strain present in a clinical specimen. A comprehensive table of sequences available for PCR amplification of mycobacterial organisms can be found in reference 51.

TABLE 10.9. ADVANTAGES AND DISADVANTAGES OF THE TRADITIONAL OR "STRAIGHT" PCR APPROACH IN AMPLIFICATION

Advantages	
	Simple
	Cheaper than more laborious approaches
	Fast
	Routinely accepted
	Has less contamination risks than when amplicon transfer is required for amplification
Disadvantages	
	Less sensitive than approaches using a second round of amplification
	Possible nonspecific banding ("dirty background") after 40–50 cycles if hot-start is not used (see text)

Reproduced with permission from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617
Abbreviations: PCR, Polymerase chain reaction

Technical Approaches or Schemes Used for Amplification ("Straight" PCR, Nested PCR, RT-PCR)

Most published studies have used the traditional or "straight" PCR (SP) approach to amplification for 30 to 40 cycles to amplify MTB. Table 10.9 summarizes the advantages and

disadvantages of this approach. This most popular and simple approach entails the addition of amplification reagents to PCR tubes from the start and the placement of tubes in the temperature cycling heating blocks for amplification to occur.

In order to increase the sensitivity of assays, particularly those that use single-copy target sequences, other investigators have resorted to the use of one or more of the following alternate modalities: two-step (two-tube) nested PCR (TSNP) (63, 64, 65 and 66); one-step (one-tube) nested PCR (OSNP) (29, 30); reamplification with the same primer pair (67); or RT-PCR (47, 48 and 49). Multiplex PCR (5, 68), or the addition of several primer pairs concurrently, is yet another choice to consider when amplification of more than one mycobacterial species is desired, as in cases of mixed infections. The advantages and disadvantages of TSNP are summarized in Table 10.10, and of the OSNP in Table 10.11.

TABLE 10.10. ADVANTAGES AND DISADVANTAGES OF THE TWO-TUBE OR TWO-STEP NESTED PCR (TSNP)

Advantages	
	Increased sensitivity
	Addition of new enzyme and reagents allows for longer maintenance of the reaction in the “exponential” phase
	Target DNA used in nested (second) step is the amplicon from the first round, which is many fold greater in number than original target in clinical sample
	Smaller size of final amplicon produced by second round of amplification increases efficiency of reaction
	Inhibitors of PCR in clinical sample are decreased during second round due to dilution as new reagents are added during the second step
	Increased specificity
	Decrease in nonspecific banding (“dirty background”) over traditional PCR following 40 or more cycles of amplification. These bands are a result of coamplification of genomic sequences (reference 63)
	Decreased false positives from concatenation of fragmented DNA (reference 64)
Disadvantages	
	Greater expense than straight PCR as twice as much enzyme and reagents are used
	Extra manipulations and length of the assay as cycling blocks cannot be programmed for all cycles from the start
	Increased risk of contamination unless three rooms are used for setup instead of two (pre-PCR, PCR, and Internest transfer rooms)

Adapted and updated, with permission, from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617
Abbreviations: PCR, Polymerase chain reaction

TABLE 10.11. ADVANTAGES AND DISADVANTAGES OF THE ONE-STEP (ONE-TUBE) NESTED PCR (OSNP)

Advantages	
	Less expensive than TSNP
	Faster than TSNP as there is no internest transfer step and instrument can be programmed from the start
	Less contamination risks than TSNP and less manipulations
	Addition of twice the amount of enzyme from the start may help approximate the level of sensitivity of the TSNP
Disadvantages	
	No added advantages of diluting out inhibitors
	May never approximate the extreme sensitivity of the TSNP

Reproduced with permission from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617

In the TSNP approach, tubes are kept in the cycling heating blocks for 20 or 30 cycles and the reaction then is stopped. An aliquot of the “amplicon” (amplified sequence) then is transferred to a new tube to which are added new enzyme and reagents and a new primer pair that is different from the outer primer pair used during the first round. This inner (nested) primer pair will code for a shorter amplicon during the second round of amplification, and this shorter amplicon is contained within the stretch of DNA amplified during the first round and thus, “nested” within it. Reaction tubes are placed on the cycling heating blocks and allowed to proceed for an additional 20 or 30 cycles, after which detection of the amplified product occurs. Because of the reasons stated in Table 10.10, the sensitivity of this assay tends to be higher than that of SP. A major disadvantage of this system is that there is an increased risk of contamination from the added step of transferring amplicon before the second round of amplification, and most diagnostic laboratories do not like to use this approach. An additional room is required for transfer of amplicon between steps (69).

The OSNP approach is a compromise between SP and TSNP. It entails the addition of two primer pairs (i.e., outer and inner primer pairs) to tubes from the start. There is no need to transfer aliquots before a second round of amplification, thus decreasing the potential for amplicon carryover contamination as with the TSNP approach. Primer pairs have to be designed carefully such that the optimum annealing temperature of the inner primers is different from that of the outer primers. During the first 20 or 30 cycles of amplification, inner primers are functionally inactive at annealing temperatures that are optimal for the outer primers. Inner primers come into play once the reaction parameters change at the end of the first round of amplification. The cycling heating blocks are programmed so that the switch occurs automatically during the last of the 20 or 30 cycles of amplification, and thus the instrument can operate uninterrupted during the night or while the technologist is performing other duties.

Using the enclosed tables that summarize the advantages and disadvantages of each of these three amplification schemes (Table 10.9, Table 10.10 and Table 10.11) the laboratorian may decide which system is optimal to perform in-house PCR for MTB. The use of RT-PCR as an additional and useful technical approach to amplification was discussed in an earlier section.

Methods of Detection of the Amplified Product

Some of the most popular techniques used to mediate detection of mycobacterial nucleic acids amplified by in-house PCR assays include use of ethidium bromide for size fractionation; hybridization with probes via Southern hybridization, in solution hybridization, or dot blotting; and sequencing of the amplicon. More novel methods include use of digoxigenin, chemiluminescence, and colorimetry. A comprehensive list with several references is included as Table 10.12. Some of the earlier studies, and subsequent studies using repeated sequences as targets, relied only on ethidium bromide for detection of the amplified product (70). Ethidium bromide is a dye that makes nucleic acids detectable by fluorescence upon intercalation between the bases of the nucleotides. It allows for size fractionation of the amplicon (i.e., determination of the size of the amplified nucleic acid compared with controls of known size). The sensitivity of an assay may be increased by using a nucleic acid probe for detection of the amplicon instead of, or in addition to, the use of ethidium bromide. Nucleic acid probes traditionally have been labeled radioactively with p32 that has served as the “reporter” molecule in many early studies. The use of such a second method of detection also enhances assay specificity. Binding or hybridization of such as probe to the amplicon may be performed via a dot blot format, a Southern hybridization format (20), or an in-solution hybridization format (65).

TABLE 10.12. METHODS OF DETECTION OF AMPLIFIED MYCOBACTERIAL NUCLEIC ACIDS

Method	References
Ethidium bromide for size fractionation	70,88
Probe confirmation using:	
Southern hybridization with p32 labeled probe	20
Southern hybridization with enzyme-labeled oligonucleotides	88
In-solution hybridization	65
Dot-blot hybridization with enzyme-labeled oligonucleotide	88
Sequencing of the amplified product	48,121,122
Determination of RFLPs with restriction endonucleases	66,72,73,74,117,123,124
Digoxigenin-mediated detection	29,32,71
Chemiluminescence	71
Colorimetry	29,116
FITC-conjugated capture probes captured by anti-FITC coated beads, followed by alkaline phosphatase-conjugated reporter probes and a chromogenic substrate	92

Abbreviations: FITC, fluorescein isothiocyanate. RFLPs, restriction fragment length polymorphism
Adapted and updated, with permission, from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1966: 617

The use of digoxigenin as part of the scheme of detection is becoming more popular (32, 71). Fiss et al. (71) used digoxigenin-labeled probes in a reverse dot blot format with good results. In the last step of the reaction, an antidigoxigenin antibody was added that was linked to alkaline phosphatase. Upon the addition of a substrate for this enzyme that chemiluminesces, detection of the original amplicon can be mediated via chemiluminescence. This group consistently detected 100 mycobacterial colony forming units (CFUs). Wilson et al. (29) used digoxigenin- and biotin-labeled primers in a colorimetric enzyme immunoassay format. Following a first round of amplification with unlabeled outer primers, inner (nested) primers that were labeled with digoxigenin and biotin, respectively, were added for a second round of amplification. Thus, the amplified product contained both digoxigenin and biotin. Microtiter wells coated with avidin served for attachment of the amplicon via the biotin moiety whereas an antibody/alkaline phosphatase conjugate reacted with the digoxigenin moiety of the amplicon. Addition of substrate allowed for colorimetric detection and quantification of the amplified product.

Kirschner et al. (48) amplified 16S rRNA sequences, sequenced the amplicon, and was able to discriminate among species of mycobacteria based on hypervariable regions within the amplified gene. Others have treated the amplicon with restriction endonucleases and evaluated the resulting restriction fragment length polymorphism (RFLP) patterns (72, 73). The greatest advancements in detection methods entail the use of nonradioactive techniques, automation, and systems that allow detection of amplicons from multiple samples at the same time (ie., microtiter plates), such as are becoming common in the commercially available amplification kits.

Commercial software programs exist that computerize the analysis of restriction fragment length polymorphisms (74). One such package is the GelCompar (Applied Maths, Kortryk, Belgium) and sold in the United States as Molecular Analyst Fingerprinting Plus (Bio-Rad Laboratories, Hercules, California). Another is the BioImage System (BioImage Corporation, Ann Arbor, Michigan). Both systems were compared (74) and found useful, however, they both require the user to make critical decisions in the analysis, to check that artifacts are not being detected and to visually verify that the system is finding all bands in each lane. Commercial sequencing software packages also exist. One such system sequences amplicons from the 16S rRNA gene (Micro Seq 16S rRNA Gene kit, Perkin-Elmer Applied Biosystems, Foster City, California).

Another aspect of the PCR procedure that has hampered its potential for an even wider range of applications, but which also appears to be under evaluation for commercialization, is the need for standardized, quality-controlled frozen or refrigerated ingredients. The availability of ready-to-use freeze-dried PCR mixes of proven quality would greatly facilitate implementation in laboratories with relatively limited facilities and who choose to use home-brew PCR. The development of one such stabilized, freeze-dried PCR mix for detection of mycobacteria was recently described by Klatser et al. (75). Reaction mixes that were off-the-shelf,

preoptimized, premixed, predispensed and freeze-dried were stable for 1 year at 20°C and for up to 3 months at 37°C. All that was required was reconstitution of the mix and addition of the template DNA. Recently, a commercial freeze-dried basic PCR mix became available (Ready-to-Go PCR beads, Pharmacia Biotech) (75). Likewise, frozen standards for use as controls are being evaluated at the present time by commercial companies.

Sensitivity and Specificity of Published Assays, and Results of Large Published Studies on In-House PCR for MTB

Most published studies show proof of the capacity of in-house systems to detect between five and 100 CFUs of mycobacteria in sensitivity studies performed using dilutions of cultural isolates. There is variability between published studies, however, because the laboratories reporting such results differ in terms of target sequences, schemes of amplification, and detection methods used in their in-house PCR amplification assays. The highest sensitivity has been reported using schemes where the target sequence is either a repeated sequence, involves nested PCR, or the assay entails the use of amplification of preamplified rRNA via RT-PCR. One group that reported detection of three to 60 mycobacteria (22) improved the sensitivity of their assay with a nested PCR scheme using the same single copy gene target sequence (i.e., the 65 kd-antigen) (67). Other groups have reported detection of fewer than 10 organisms using a repetitive sequence as target (52), of 1 fg of DNA from culture using ethidium bromide only and 40 cycles of non-nested PCR (70), and of 1 fg of DNA using a repetitive insertion sequence (35) and 30 cycles of PCR from culture. Miyazaki et al. (33) increased the sensitivity of their assay 1,000-fold by a nested PCR approach, 35 cycles per run, down to detection of 10 fg of DNA, and Boddinghaus et al. (47) reported detection down to 0.1 fg of purified nucleic acid from cultures using RT-PCR of rRNA and a p32-labeled probe. This overwhelming variability in detection of amplified products by these investigators must be attributed to the enormous diversity in methodology used by each of the reporting laboratories.

In general, however, when actual clinical specimens are run and the results are compared with those of culture from the same specimens, the sensitivities of assays are usually lower than when dilutions from culture isolates are prepared and subsequently amplified (76). That is not surprising given that there are several known, as well as many uncharacterized, inhibitors of the Taq polymerase. These include, among others, hemoglobin (77), phenol, SDS, anticoagulants such as heparin and EDTA (77), and polyanetholsulfonate (77). In one study of amplification for *Mycoplasma pneumoniae* (78), PCR inhibition was observed in throat swabs submitted in routine bacteriological transport media, where it was the agar that was shown to be inhibitory. In other studies, inhibition was noted by calcium alginate and aluminum swab shafts (79), and by mucolytic agents (80). PCR application to fecal specimens has been very limited because of the presence of abundant and unknown inhibitors in such specimens. Recent manipulations, treatments, and columns have yielded significant improvements in this arena (81). Some of the measures reported to nullify the presence of inhibitors in clinical samples are included in Table 10.13. Some measures also have been reported to help increase the specificity of the assays by decreasing the rate of false positives, and some of these include: probe confirmation of the amplicon and use of “hot start” techniques to reduce mispriming at low temperatures. Measures are also known that will decrease or prevent the rate of PCR carry-over contamination of aerosolized amplicons from one clinical sample to another, and a comprehensive list is included in Table 10.14. Some of these interventions, such as the d-UNG, would have been helpful in the earliest studies in the literature reporting on the use of actual clinical specimens, such as that of Shankar et al. (76). This group reported on the use of PCR to diagnose meningitis as a result of MTB in cerebrospinal fluid. Although PCR was found to be more sensitive than culture in highly probable cases (15 of 20 versus 4 of 20, respectively), they reported on the possibility of cross-contamination of some of their negative controls and carryover abrogation techniques were not available at the time of that study.

TABLE 10.13. SEVERAL MEASURES USED TO NULLIFY OR DETECT THE PRESENCE OF INHIBITORS OF PCR AMPLIFICATION THAT MAY BE PRESENT IN CLINICAL SPECIMENS

Measure	Reference
Centrifugation through a sucrose gradient	28
Use of guanidinium thiocyanate/silica particles as the lysis and capture methods	29
Two-step nested PCR	65
Sample dilution	31
Use of internal control primers for myosin, actin, etc.	30
Use as internal control of a sequence that coamplifies with the target sequence, but with a product of different size	23
Column chromatography	23,30
SDS-NaOH as lysis pretreatment	125

Abbreviations: PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.
Adapted and updated, with permission, from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WS Saunders Co; 1996: 617

TABLE 10.14. SEVERAL PRECAUTIONS AND MEASURES KNOWN TO DECREASE OR PREVENT PCR CARRYOVER CONTAMINATION OF AEROSOLIZED AMPLICONS FROM ONE CLINICAL SAMPLE TO ANOTHER

Preparation and use of master mixes, separate pipettes and rooms for pre- and post-PCR work, single use frozen reagent aliquots
Use of cotton-plugged pipette tips
Use of commercially available enzyme systems to inactivate prior contaminating amplicons d-UNG (uracyl DNA-N-glycosylase enzyme) with dUTP nucleotides
Photochemical modification of contaminating DNA with psoralen or isopsoralen
Treating the reaction mixture with multiple restriction enzymes, DNAase, bleach or hydroxylamine
Autoclaving
Exposing the reaction mixture to ultraviolet light

Abbreviations: PCR, polymerase chain reaction

Several large studies have been published using in-house PCR assays for detection of MTB in actual clinical specimens. The great majority use respiratory specimens, especially sputa, for the evaluation. Of note are studies conducted by Clarridge et al. (26), Kirschner et al. (82), Forbes et al. (83) and Nolte et al. (84). The specifics on sensitivity and specificity, as well as other important details, from each of these studies are found in Table 10.15. Although some of the earlier studies utilized culture alone as the gold standard for the study, others used an expanded gold standard (EGS), whereby culture alone was not the only criterion. Any questionable or incompatible result where there was a discrepancy between culture and amplification, was submitted to a clinical investigation consisting of chart review of the patient's history and physical examination, evaluation of roentgenogram results, as well as any other prior laboratory studies. This evaluation helped adjudicate true positivity to either culture or amplification, with results from the other method being considered falsely negative.

TABLE 10.15. SENSITIVITY AND SPECIFICITY OF LARGE PUBLISHED STUDIES UTILIZING IN-HOUSE PCR AMPLIFICATION FOR DETECTION OF MYCOBACTERIUM TUBERCULOSIS IN RESPIRATORY SPECIMENS

Reference	Total # of Specimens	# of MTB Positive Specimens	Gold Standard	PCR Sensitivity For all Specimens (%)	PCR Sensitivity For Smear Negatives (%)	PCR Specificity (%)	Smear Sensitivity (%)	Culture Sensitivity (%)	Target Used
26	5,000	218	C	83.5	62	99	66(A)	100	IS6110
82	8,000	729	EGS	84.5	72.8	99.5	11(A)	77.5	16S rRNA Gene
83	727	75	EGS	87.2	-	97.7	69.3(A)	89.3	IS6110 & PAB gene
84	313	124	C	91	57	100	88.7	100	IS6110

Abbreviations: A, auramine stain; C, culture; EGS, expanded gold standard; K, Kinyoun stain; MTB, *Mycobacterium tuberculosis*; PCR, Polymerase chain reaction

Clarridge et al. (26) reported on the large-scale use of PCR in a routine mycobacteriology laboratory that processed over 5,000 specimens. Of 623 cultures positive for AFB, 218 were positive for MTB. PCR was positive in 181 cases (85%). They report the presence of inhibitors in 8% of clinical specimens. When compared with culture, PCR had a sensitivity of 83.5%, a specificity of 99%, and a positive predictive value of 94.2%. If results from multiple specimens on the same patient were considered, no patient who had three or more sputa tested in their laboratory was misdiagnosed, a comment echoed by several authors in the other large published studies. This group concluded that such results are encouraging for the use of PCR for the rapid identification of the most infective cases of tuberculosis. Kirschner et al. (82) offered an additional suggestion: in paucibacillary disease (i.e., where the specimens are smear-negative), do not report out a negative amplification result until three specimens have been evaluated. A summary and “take-home” figures on sensitivity and specificity from all large studies conducted using in-house PCR for MTB detection in actual clinical respiratory specimens, can be found in Table 10.16. It is obvious that it is the low sensitivities achieved while testing smear-negative specimens that will prove to be the Achilles' tendon of all molecular diagnostic mycobacteriological methods (85). It is suggested that the reader compare these figures with those to be presented later in this chapter from large studies using commercially available kits.

TABLE 10.16. SUMMARY OF RESULTS FROM LARGE PUBLISHED STUDIES ON USE OF IN-HOUSE PCR FOR DETECTION OF *MYCOBACTERIUM TUBERCULOSIS* IN CLINICAL RESPIRATORY SPECIMENS

Average Sensitivity (all specimens):	83–90%
Range of study sensitivities:	70–100%
Average Sensitivity for Smear (+):	>90%
Average Sensitivity for Smear (–):	50–80%
Average Specificity:	>90%

In-house PCR assays applied to clinical specimens other than those from the respiratory tree, such as pericardial tissue in cases of tuberculous pericarditis in HIV-positive patients, also show promise and merit further development (86). Likewise, there are published schemes on PCR detection of MTB in formalin-fixed, paraffin-embedded tissues using primer pairs that yield high sensitivities (34).

A word of caution must be raised in regards to the results from an important and somewhat discouraging report on the actual status of sensitivity and specificity of in-house studies as they have been performed in many laboratories around the world (87, 88 and 89). Noordhoek et al. (87, 88) recruited seven laboratories from different countries and continents to participate in a blinded study of 200 sputa, saliva, and water samples spiked with known numbers of *M. bovis* BCG cells or with no organisms. The IS6110 repeated sequence was used as target. Each laboratory was instructed to use its own protocol for pretreatment, extraction, detection, and amplification of samples, as they would if these were specimens coming from patients admitted to their institution. The results were surprising. As expected, there was an enormous variability in the methods used at all steps in the process, from pretreatment and lysis of specimens through detection of amplicon. The reported sensitivities ranged widely. In samples with as many as 1,000 microbes, the reports of positive results by PCR ranged from 2% to 90% of samples. In the interest of fairness, it must be mentioned that *M. bovis* BCG cells, sent to participants in lieu of MTB for safety purposes in shipping, usually contain fewer molecules of IS6110 than MTB (1 to 5, versus 10 to 20). Sensitivities with actual MTB-containing samples theoretically should have been higher than those reported in the study. However, specificities also varied widely. False-positive results ranged from 3% to 20% with one extreme value of 77%. It was concluded in this study that implementation of an effective system for monitoring the sensitivity and specificity of in-house PCR assays would be required before the technique can be used reliably. It also was suggested that more controls be added throughout the entire process, from collection of samples through detection of amplified product. This study supported the conclusion of the U.S. Public Health Service at

the time that with such a current state of technology, reliable detection of MTB still depended on conventional laboratory methods.

In 1996, a subsequent study along the same lines was undertaken once again by Noordhoek et al. (90) in order to shed light further on the reliability of nucleic acid amplification and detection of MTB in the world. This time, it was an international collaborative quality control study among 30 laboratories in 18 countries. Blinded panels of 20 sputum samples were prepared, containing zero, 100, or 1,000 mycobacterial cells. Each laboratory was asked to detect MTB by their routine method of nucleic acid amplification. This is an important difference from the earlier study where participants were asked to use the repetitive IS6110 sequence of MTB as the target for amplification. Part of the large variation in sensitivity and specificity of this earlier study may have been because of the lack of familiarity of some investigators with the use of the target IS6110. Therefore, this second study allowed participating laboratories to use their own protocol for amplification and detection of MTB without any restriction on the method or the nature of the target. Some of the laboratories used commercial amplification kits, such as those that will be discussed subsequently in this chapter. Only 20 blinded samples were provided this time, instead of 200, so as not to disrupt the routine workloads of the participating laboratories.

Only five laboratories correctly identified the presence or absence of mycobacterial DNA in all 20 samples. Seven laboratories detected mycobacterial DNA in all positive samples, and 13 laboratories correctly reported the absence of DNA in the negative samples. Lack of specificity was more of a problem than lack of sensitivity. Of interest, no two laboratories used exactly the same procedure for liquefaction and decontamination, pretreatment for lysis of MTB, DNA extraction, analysis of PCR products, or monitoring for the presence of inhibitors of the polymerase. Eight participants used commercially available kits, 19 participants used IS6110 as target. There was no correlation between correct test results and the use of any particular method. Seventeen laboratories used dUTP/dUNG to avoid contamination with amplicon, however, eight of these reported false-positive results among the BCG-negative samples.

The results of this second study show that nucleic acid amplification methods for MTB detection may be unreliable, even when laboratories use the method routinely applied in their laboratory. Even well-standardized reagents for amplification and detection such as those present in commercially available kits do not ensure reliable detection. This may be because of the fact that these kits do not include standardized procedures and reagents for liquefaction and decontamination of specimens. Pipetting, mixing, and centrifugation are critical to avoid loss of DNA and to circumvent cross-contamination. Appropriate positive and negative controls should be used in each run so that the whole procedure from sample lysis to detection undergoes monitoring. The authors recommend the use of negative sputum samples to which a known, small number of mycobacteria have been added by spiking. With controls such as these, the entire procedure from pretreatment to analysis is monitored for sensitivity. Because mycobacteria tend to clump, however, control reagents are not easy to prepare, and should only be prepared under standardized conditions and eventually should become available commercially. Internal controls to monitor for inhibition of the polymerase in the individual samples should also be used, which could consist of a modified target that is added to the sample. To conclude on this second study (90), the outcome of the study underlines the need for reference reagents and standardized operating procedures that would enable experienced technicians to perform reliable quality control assessment of amplification methods. Reference reagents should be used to monitor the performance of the whole assay, including pretreatment of clinical samples.

Notwithstanding any discouraging results gleaned from the two previous studies, a large number of reputable laboratories around the world perform in-house PCR tests for MTB with dependable results. There are many measures that can be taken to enhance the legitimacy of in-house PCR assays for MTB, beyond those stipulated on Table 10.13 and Table 10.14. At the level of one's own local hospital or institution, a laboratory must institute the regular use of high-positive, low-positive, and negative controls with each run. There should be stringent carryover abrogation techniques instituted in the laboratory. Before using an assay with patient samples, it must be validated via a sensitivity study using dilutions from culture, a specificity study using species other than MTB, a study of actual clinical samples compared with culture, and additional studies to optimize the reaction with the use of adjuncts (Table 10.17) and known abrogators of inhibitors. At the international level, the development of standardization attempts such as stereotypical protocols published by the National Committee for Clinical Laboratory Standards (NCCLS) for molecular microbiology must continue to be

encouraged. Enrollment in proficiency testing, laboratory inspections, laboratory surveys, and other policing modalities such as those performed by the College of American Pathologists is fundamental. It may well be, however, that the ever-increasing commercial availability of stereotypical PCR and non-PCR amplification kits will greatly standardize the field by allowing techniques that may be less prone to contamination with amplicon, more amenable to quantification, or more sensitive than the first-generation-in-house assays that have been reported in the literature to date. Developed countries are starting to employ such commercially available kits with great speed at present. It remains to be determined, however, if the high costs of such kits will allow the developing world access to such urgently needed technology, or if the poorest of countries may have to continue to rely only on potentially cheaper in-house PCR assays. In the worse of scenarios, some countries may be incapable of resorting to the use of amplification techniques at all.

TABLE 10.17. SEVERAL APPROACHES TAKEN TO OPTIMIZE ASSAY SENSITIVITY AND BAND INTENSITY OF AN IN-HOUSE PCR ASSAY FOR DETECTION OF MYCOBACTERIUM TUBERCULOSIS UTILIZING GELS FOR SIZE FRACTIONATION OF THE AMPLICON (REFERENCE 30)

Incorporation of 5% formamide during hybridization step
Optimization of hybridization temperature
Incorporation of 15% glycerol during the amplification step
Increase to 2X of Taq polymerase quantity
Use of thin-walled tubes for amplification
Evaluation of combinations of water and buffers from different commercial sources
Use of fresh primer aliquots

Reproduced with permission from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617

Possible Explanations for the Attainment of Less Than 100% Sensitivity in Amplification Assays, Particularly with Smear-Negative Specimens

If PCR amplification does offer the potential for detection of as little as a single organism present in a clinical specimen, why is less than 100% sensitivity attained in clinical studies? There are at least six possible explanations to this phenomenon.

Almost uniformly in the large published studies, a much greater portion of the treated sputum or clinical sample was utilized for culture of the specimen in liquid and/or solid media, and the remainder was used for PCR amplification. This is so because the gold standard for identification has been culture, and because these were actual clinical specimens, they deserved to be tested by the routinely used and accepted diagnostic methods, which must take precedence over techniques under evaluation. Thus, from a 1-mL volume of treated sputum, as much as 66% (26), 80% (82), or a greater fraction was used for culture in these studies, and left-over fractions were utilized for PCR.

It is well known that the distribution of acid-fast bacilli within a clinical specimen tends to be uneven because of clumping. This could be worse in paucibacillary specimens where clumping of very few organisms could create great specimen heterogeneity. Sampling error thus becomes a major problem with these specimens, and the testing method that is alloted the greatest amount of sample will produce the best results. Because a smaller fraction of the specimens was tested by PCR, the sensitivity of PCR may have been artefactually low. This intra-specimen variation in distribution because of clumping also could affect specimens collected on different days (i.e., inter-specimen variation), so that the same phenomenon could be occurring with all specimens tested on a given patient, with the same result: culture is favored to receive the few cultivable clumps of organisms, while PCR receives a fraction for amplification that may contain no organisms.

A fraction of the extracted or exposed mycobacterial DNA also can get lost during the extraction procedure, a factor that gains importance with paucibacillary specimens. This variable would not be expected to affect culture, as extraction of nucleic acids is not a part of cultural procedures.

Perhaps of greatest importance is the fact that there are inhibitors of PCR present in clinical samples, a phenomenon alluded to earlier in this chapter, and whose presence could endanger PCR sensitivity particularly for paucibacillary specimens. Large studies report that on average, 3% to 16% of all clinical specimens possess inhibitors of the polymerase (91). Measures such as the use of internal control sequences, or of spin columns to remove proteic inhibitors, or the use of PCR amplification of samples previously submitted to broth enrichment, may be necessary in order to get rid of amplification inhibitors and increase the sensitivity of PCR. Some of these measures will be the subject of further comment in the following section of this chapter. Internal controls are of immense utility in PCR and are sequences that are amplified either concurrently with the target sequence (i.e., in a multiplex style) or in a separate aliquot of the lysed specimen (92), and are used to evaluate the presence of inhibitory substances in the sample. These controls should always be run as part of a clinical assay. They may consist of 100 to 200 base pair sequences found within normal cellular gene sequences and expected to be present in all specimens whether or not the target (i.e., mycobacterial) sequence is present. Examples of internal controls include myosin heavy chain gene (93, 94), b-globin, and lipoprotein, as well as other genomic endogenous sequences. Such sequences help monitor the integrity of the nucleic acid target. A disadvantage of the use of endogenous sequences is that at times these may not accurately reflect amplification of the primary target because of differences in the primer sequences, size of the amplicon, and the relative amount of the two targets (95). Other internal controls that are not normal cellular gene sequences also have been described, which serve their purpose well, such as sequences within the 16S to 23S rRNA intergenic space regions of MTB (92). Alternatively, another approach consists of the use of synthetic genetically engineered internal control sequences with primer binding regions identical to those of the target sequence and which, thus, coamplify with it. The internal sequence may be similar to the target sequence in length and base composition but contains a unique probe-binding region that differentiates the internal control from amplified microbial target nucleic acid (95). Such a system has been described for use in conjunction with the Roche Cobas Amplicor, and microwell Amplicor systems (Roche Molecular Systems, Branchburg, New Jersey) (95), which will be described later. The internal control also may be made longer or shorter than the target sequence so as to render it detectable by size fractionation, if a gel system is being utilized for detection of the amplified product.

A sixth and final possible explanation for the low sensitivities attained with in-house PCR assays is offered, and becomes manifest when workers try to boost their systems by adding more lysed specimen to the amplification tubes in an attempt to increase the amount of MTB DNA placed in PCR reaction tubes. Addition of more lysed specimen will result in the addition of mostly genomic DNA, always present in much greater amounts in a specimen than mycobacterial DNA. Given the stoichiometric ratios that must be met within reaction volumes of 100µL or less, too much genomic DNA may inhibit PCR instead of boosting its sensitivity. There are attempts at enriching for specific MTB DNA and using such enriched fractions for amplification, and one such method is described next.

Trends in the Field for Increasing Sensitivity of In-House Assays, Especially for Smear-Negative Specimens.

Several trends are being followed in the field of amplification in an attempt to increase sensitivity when testing smear-negative specimens. Two such trends will be examined in this section.

MTB can be allowed to undergo cultural amplification in broth (i.e., as in the 12B bottles of the Bactec 460 system) up to a low growth index (GI), such as between 10 and 50 (91, 96). Instead of waiting until a considerably higher index is produced, with its attendant delay in diagnosis, to be followed by the use of Gen-Probe Accuprobe DNA probes for culture confirmation, an aliquot of broth preamplified to a low GI then is submitted to PCR amplification. This approach saves time over culture confirmation with probes and increases sensitivity over the use of PCR directly from the original smear-negative clinical specimen. It could be argued that this abrogates the promise from amplification techniques of allowing diagnosis while obviating the need for culture, but with sensitivities as low as those obtained with smear-negative specimens in the previously reported studies, the boost that such an approach can offer can not be dismissed. Success with a similar but expanded approach was described by Sanguinetti et al. (97). These investigators applied PCR amplification of 16S rRNA gene sequences and a subsequent reverse cross-hybridization assay with 14-C species-specific probes, to the identification of mycobacteria grown in liquid medium positive by the Bactec 460 TB and the Bactec 9000 MB systems (Becton Dickinson, Franklin Lakes, New Jersey). Commercially-available nucleic acid amplification systems are also being used to amplify MTB from positive broth cultures (77, 98, 99).

A second approach described is the enrichment and capture of MTB DNA (100, 101a) extracted from a clinical specimen and concentrated as the lysed specimen runs through a specific chromatography column, with the added benefit of removal of genomic DNA and protein inhibitors (100). Such columns can be coated internally with anti-MTB probes that specifically bind MTB DNA subsequently eluted from the column and used for PCR amplification (i.e., “Sequence-capture PCR”)(100). Although these systems do increase assay sensitivity, they are more labor intensive and require more modifications to the in-house procedure than systems without such enrichment attempts. Perhaps eventually commercial kits will include such modifications for amplification from smear-negative specimens. Alternatively, as described with studies using *M. ulcerans* (101b), immunomagnetic separation was utilized to isolate the intact organism from a heterogeneous sample, prior to lysis of the microbe's nucleic acid in preparation for PCR amplification. The technique uses monodisperse magnetic beads coated with antibodies against surface epitopes of the bacterium of interest.

CPT Codes Intended for Use with Molecular Diagnostic Techniques for Analysis of Nucleic Acids

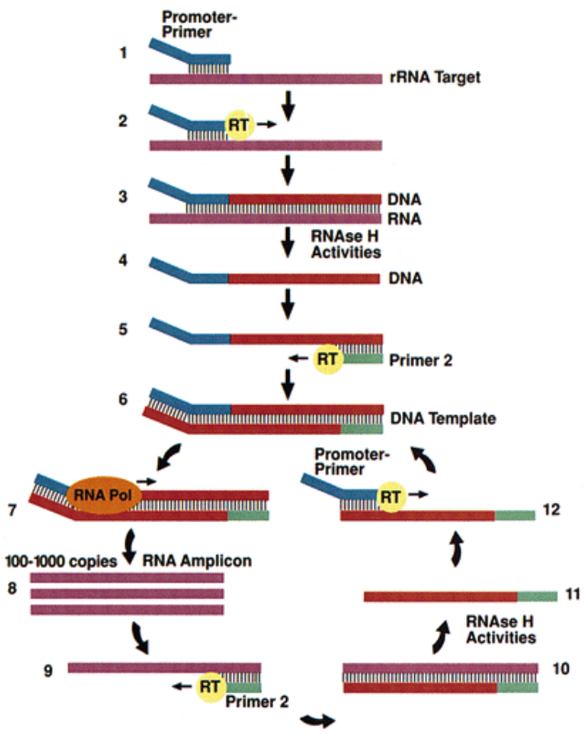


FIGURE 10.1. Steps in the transcription-mediated amplification cycle (TMA): (i) Promoter-primer binds to rRNA target; (ii) reverse transcriptase (RT) creates DNA copy of rRNA target; (iii) RNA:DNA duplex; (iv) RNAse H activities of RT degrades the rRNA; (v) primer 2 binds to the DNA and RT creates a new DNA copy; (vi) double-stranded DNA template with a promoter sequence; (vii) RNA polymerase (RP) initiates transcription of RNA from DNA template; (viii) 100 to 1,000 copies of RNA amplicon produced; (ix) primer 2 binds to each RNA amplicon and RT creates a DNA copy; (x) RNA:DNA duplex; (xi) RNAse H activities of RT degrades the rRNA; (xii) promoter-primer binds to the newly synthesized DNA. RT creates a double-stranded DNA and the autocatalytic cycle repeats resulting in a billion-fold amplification. Reproduced with permission from the Gen-Probe company.

There are several Current Procedural Terminology (CPT) codes described specifically for use whenever molecular diagnostic techniques are utilized to analyze nucleic acids. Whereas the earlier series of codes 83890 through 83912 was used for testing of infectious diseases, the CPT 1998 manual offers a new and expanded list of codes for testing of microorganisms. The specific assay run should be reported as precisely as possible by using the most specific code available. If there is no specific agent code for the particular organism being tested, then the general methodology code for infectious diseases should be used (i.e., 87797, 87798, 87799). Table 10.18 summarizes the specific codes applicable to detection by nucleic acid probes and/or amplification techniques of the five microorganisms most frequently amplified from clinical specimens (i.e., MTB, HIV, *Chlamydia*, CMV, and

hepatitis C). It also lists general codes for use when the organism in question has yet to have a specific code allotted to it, as well as other general codes that apply to any molecular diagnostic procedure. In cases where a labor-intensive in-house assay with multiple steps is used for analysis of an organism and no specific microbial code is available, then each procedure could be coded separately. For example, the steps of isolation of DNA, restriction endonuclease digestion, electrophoresis, and nucleic acid amplification could be coded as 83890, 83892, 83894, and 87798.

Gen-Probe TMA Commercially-Available System (rRNA Transcription-Mediated Amplification System) Coupled with the Hybridization-Protection Chemiluminescent Detection System for Mycobacterium Tuberculosis.

Features

The original Gen-Probe (San Diego, California) amplification system for MTB is called the AMTDT-1 (Amplified MTB Direct Test, Version 1). It was the first amplification assay for MTB to be approved by the FDA for use in the United States. It received approval in December 1995 for use with smear-positive respiratory specimens collected from previously untreated patients. Such restricted use was no doubt from the paucity of data at the time on its performance with nonrespiratory specimens such as CSF, pleural and synovial fluid, etc., to the lack of data on how long sputa from positive patients would remain amplification-positive for MTB after appropriate therapy had been instituted (12), and as will be presented later in this chapter, to the low sensitivities achieved when smear-negative specimens were tested. The test also has been marketed in Europe, Canada, Japan, Australia, and other Pacific-rim countries. Development of amplified tests for other microbes using the same technology is underway.

A transcription-mediated amplification system (TMA) differs from PCR, most notably, in the type of amplicon produced. In PCR, DNA is produced in large quantities from a DNA template. In RT-PCR, RNA templates are converted to single-stranded DNA templates that serve as substrate for the production of multiple DNA particles during amplification. In TMA, ribosomal RNA substrates are converted to single-stranded DNA substrates that are then transcribed into an RNA amplicon (Fig. 10.1). TMA is an isothermal, autocatalytic, homogeneous system. It is isothermal in that a single temperature is used during the amplification step, and homogeneous in that it does not use a solid substrate or precipitin reaction to separate the bound product (ligand-ligate) from free product; this occurs in the solution phase. Following a 2-hour incubation step at 42C, abundant RNA is produced that subsequently is detected by the Gen-Probe chemiluminescence-labeled DNA probes using the Hybridization Protection Assay (HPA) (Fig. 10.2). TMA assays provide several advantages: (i) there is abundant rRNA inside the MTB (2,000 to 5,000 molecules) so that the target is “preamplified” and thus has the potential for an increased sensitivity; (ii) a single temperature is used during the amplification step; (iii) a single-tube format is used whereby reagents are added to the tube but never removed from it, and detection occurs in the tube that minimizes the risk of carryover contamination; (iv) results are available on the same day (4 to 5 hours); (v) there is no nucleic acid purification step, just sonication; (vi) the HPA format has been well tested for several years when used for culture confirmation; (vii) RNA is more labile outside the tube than DNA, which reduces carryover risks; and (viii) system picks up MTB-complex members. A summary of the general features of AMTDT-1 can be found in Table 10.19.

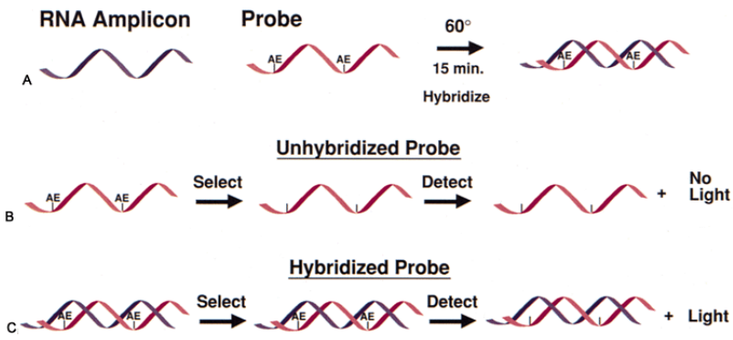


FIGURE 10.2. Detection of amplicon with DNA probes and the Hybridization Protection Assay (HPA) technique. (**A**) Acridinium ester (AE)-labeled DNA probes are added and allowed to hybridize to specific target sequences within the amplicon produced in the TMA reaction. (**B**) Separation of hybridized from unhybridized probes is done by the addition of selection reagent that hydrolyzes the AE on the unhybridized probes. No light is emitted in the luminometer from the unhybridized probes. (**C**) The AE on the hybridized probes is protected within the double helix and is not hydrolyzed by the selection reagent. Light is emitted and detected by the luminometer. Reproduced with permission from the Gen-Probe company.

TABLE 10.19. FEATURES OF THE GEN-PROBE AMTDT-1 (AMPLIFIED *MYCOBACTERIUM TUBERCULOSIS* DIRECT TEST, VERSION 1)

Type of amplification	Target
Target	16S rRNA (repeated sequences, very “preamplified”) of MTB complex
Amplification method	TMA (Transcription-mediated Amplification)
Enzymes	Reverse transcriptase, T7 RNA polymerase
Sample volume	50 µL of treated specimen, with N-acetyl-L-cysteine/NaOH suggested as liquefier
Sample lysis	Sonication and beads
Format	Single tube
Thermal conditions	Isothermal reaction
Amplified product	RNA
Detection	Chemiluminescence (Hybridization-Protection Assay)
Amplicon containment	Procedural (no removal); RNA is labile in environment
Instruments	Luminometer, heating blocks
Assay time	4–5 hours (amplification step at 42C is 2 h.)
Kit format	50 tests. Shelf-life is 12 months with unreconstituted reagent vial, and 1 month once reconstituted
Laboratory space requirements	Single room
Training	Offered by company

Abbreviations: MTB, *Mycobacterium tuberculosis*

The main steps involved with this system are enumerated in Table 10.20. Sputum samples or bronchoalveolar lavages are treated for digestion decontamination and concentration and a sample placed in lysing tubes containing glass beads. Specimen lysis is mediated via sonication, which breaks cell walls and liberates rRNA. The mycobacterial 16rRNA appears unaffected by the routine decontamination and lysis procedures used in most laboratories that entail the use of N-acetyl-L-cysteine/NaOH, as opposed to mRNA whose levels may be reduced significantly (102). Following lysis, an amplification reagent is added to each

reaction tube and overlaid with oil. Various DNA primers are included in this reagent. Aliquots from the lysed specimens then are added to the reaction tubes with the amplification reagent in it. Incubation occurs at 95°C for 15 minutes in a dry heating block and that is followed by cooling at 42°C for 5 minutes, which allows the primers to anneal to the regions that flank the target. Then, the enzyme mix is added. There are two enzymes added together: reverse transcriptase (RT) and T7 RNA polymerase (RP). Nucleotides (dNTPs and NTPs) also are included. This mixture is incubated at 42°C for 2 hours, during which time the RT will create complementary DNA (cDNA), and the RP will transcribe that to multiple RNA amplicons. Subsequently, a termination reagent is added to each tube to destroy any cDNA template.

TABLE 10.20. MAIN STEPS INVOLVED IN THE GEN-PROBE AMPLIFIED *MYCOBACTERIUM TUBERCULOSIS* DIRECT TEST, VERSION 1 (AMTDT-1)

1)Specimen lysis

2)Addition of an “amplification reagent”

3)Incubation

4)Addition of an “enzyme mix”

5)Incubation

6)Addition of a “termination reagent”

7)Detection by the Hybridization Protection Assay

Adapted and updated, with permission, from Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia, PA: WB Saunders Co; 1996: 617

Detection occurs via hybridization with a DNA probe. Reconstituted probe reagent now is added to the reaction tubes and incubation occurs at 60°C for 15 minutes. A selection reagent then is added for 10 minutes at 60°C to allow for the selective detection of hybridized amplicon. Acridinium ester, the reporter molecule on the probe, will undergo hydrolysis only when present on unhybridized probes. This is the essence of the hybridization protection assay (HPA). Upon the formation of a hybrid between the DNA probe and the RNA amplicon complementary to it, the acridinium ester label on the probe is

protected against alkali treatment by the formation of the double-stranded hybrid (DNA-RNA helix). This selection step occurs in solution in a single phase. Chemiluminescence is achieved by excitation of the acridinium ester molecule with the addition of H2O2/OH-. Detection is accomplished by reading chemiluminescence in a luminometer. A large signal-to-noise ratio is achieved. In regards to hardware, only a sonicator, a luminometer, heating blocks, and water baths are required.

Large Published Studies

The AMTDT-1 has been evaluated in at least seven large published studies, using predominantly respiratory specimens, the majority of which were sputa (126, 127, 128, 129, 130, 131 and 132). The cutoff for positivity used in these studies was approximately 30,000 relative light units (RLUs). The specifics for each study are summarized in Table 10.21. It can be concluded from such large studies that the sensitivity of AMTDT-1 for all specimens ranges from 70% to 98% with an average of 86% to 90%, while that of smear-negative specimens ranges from 70% to 80%, with some selected lower results. Specificities were high, ranging from 97% to 100%, with an average of 98%. Once again, it is the smear-negative specimen that challenges amplification systems in their inability to detect very low levels of mycobacteria. The majority of the studies concluded once again that the use of three or more specimens per patient leads to an accurate diagnosis.

TABLE 10.21. SENSITIVITY, SPECIFICITY AND OTHER DATA FROM THE SEVEN LARGE PUBLISHED STUDIES ON THE AMTDT-1 USING PREDOMINANTLY RESPIRATORY SPECIMENS, MOSTLY SPUTA

Reference	Total # Specimens	# MTB Positive Specimens	Gold Standard	AMTDT-1 Sensitivity (%) for all Specimens	AMTDT-1 Sensitivity (%) for Smear Negatives	AMTDT-1 Specificity (%)	Smear Sensitivity (%)	Culture Sensitivity (%)
126	135	37	EGS	91.9	70	100	71.9 (K)	96.9
127	758	119	EGS	82	71.8	99	53 (A)	88
128	750	156	EGS	91	72.5	98.5	67.3 (A)	91
129	938	111	EGS	97.8	80	97.1	Not reported (A)	89
130	617	21	C	71.4	14 (1/7)	99	66.7 (A)	100[a]
131	760	175 (both hospitals, H and L)	C	H 65, L 93	48 (both hospitals)	H99, L99	H 54 (A) L 87 (A)	100
132	412	61	C	98.4	Not reported	98.9	Not reported	Not reported

[a] Only specimens positive by culture were considered true positives, so the sensitivity of culture as gold standard was 100%. Yet, 6 specimens from patients with tuberculosis and on anti-TB medication were AMTDT-1 positive/culture negative but were disregarded when calculating sensitivities.
Abbreviations: A, auramine; C, culture; EGS, expanded gold standard; K, Kinyoun; MTB, *Mycobacterium tuberculosis*.

Issues to be Addressed and Attempts at Addressing Them

The issue of potential for inhibition during amplification of clinical samples when a target amplification system such as TMA is used, is always of concern and should be addressed by the incorporation of an internal control in the system. The response thus far has been that because the amplicon in TMAs is RNA, which is very labile outside of the tube, this should not be of concern. The need for full automation of the system is being addressed in two ways (see Table 10.5): by way of the TIGRIS instrument now under development, and by a collaboration with the VITEK company (Shelton, Connecticut) to produce cartridges that allow the use of the TMA technology within the format of the VIDAS instrument (i.e., VIDAS PROBE). In terms of inoculum size, it has been suggested to the company that increasing sample inoculum from 50 to 500 µL would increase sensitivity, and the enhanced version of the kit has taken that into account, as will be described later. The kit needs adaptation to nonrespiratory samples, and the packaging format should be revisited to include kits with 50 tests or less that would allow small laboratories access to the technology before the kit shelf-life expires. Alternatively, the shelf-life of the reagent vial following reconstitution should be increased.

Bodmer et al. (133) increased the volume of decontaminated respiratory specimen used from 50 to 500 µL with an increase in sensitivity from 71.4% to 83.3%. This modification was adapted into the enhanced second version of the kit, known as the AMTDT-2. Four studies reported on the adaptation of the procedure for use with CSF and other nonrespiratory specimens, with high sensitivities (125, 134, 135 and 136), and the results are presented in Table 10.22. It is precisely the extrapulmonary presence of tuberculosis (i.e., pleuritis, lymph nodes, CSF, etc.) that requires a rapid and very sensitive laboratory diagnostic method, because the traditional techniques have well-known limitations, mainly because of the small amount of bacteria present in these specimen types. In fact, direct CSF smears in tuberculous meningitis, although virtually diagnostic if the organism is found, usually are positive in less than 10% of cases, while cultures require up to 8 weeks and also are often negative (137, 138 and 139).

TABLE 10.22. EVALUATION OF THE AMTDT-1 WITH A VARIETY OF NONRESPIRATORY SPECIMENS, INCLUDING CSF

Reference	No. of Specimens	Type of Specimens	Sensitivity (%) for all Specimens	Specificity (%) for all Specimens
134	322	Variety*	93.1	97.7
135	294	Variety*	83.9	99.6
136	224	Variety*	85.7	100
125	136	121 peripheral bloods, 15 bone marrows	94.3	100

* Variety of specimens includes CSF, urine, lymph nodes, various body fluids or aspirates, biopsy tissues, and pus.
Abbreviations: CSF, cerebrospinal fluid

Pfyffer et al. (134) increased the amount of sample to 500 µL, pretreated specimens with SDS-NaOH and increased the amplification time from 2 to 3 hours. Sensitivity increased significantly, especially for CSF where they consistently detected 100 cells per milliliter. SDS denatures potential protein inhibitors of amplification, but must be removed by sufficient washings prior to running the assay.

Ehlers (135) treated specimens with NaLC-NaOH with no

adaptations of the stereotypical procedure. Substances inhibitory to amplification were found in 7.5% of all samples, and they tried various commercial extraction columns and resins but without any consistent ability to eliminate inhibitors.

Gamboa et al. (136) also pretreated specimens with SDS-NaOH and found it to be mandatory in obtaining consistent and reproducible results. When this group increased the specimen size to 500 µL concurrent to the use of SDS, the sensitivity of the study increased to 95.6%. The same group then reported in another paper (125) the use of 10% SDS for the preparation of peripheral blood and bone marrow aspirates from AIDS patients with high sensitivities. Once again, SDS appears to abrogate inhibitors present in clinical specimens.

The enhanced version of the Gen-Probe kit for MTB recently was FDA-approved (see Table 10.4), and is called the AMTDT-2 or E-MTD for enhanced version MTD. At least one large published study and several abstracts are available on its performance (98, 140, 141).

Some changes in the kit format include: increase in sample volume size from 50 to 450 µL, reduction in incubation time of the 42°C amplification step from 2 hours to 30 minutes, and elimination of the termination reaction. The assay turn-around time is decreased from 5 hours to 3.5 hours. Gamboa et al. (140) compared AMTDT versions 1 and 2 in respiratory and nonrespiratory specimens and the summary is found in Table 10.23. All samples were pretreated with SDS-NaOH or 10% SDS. Overall results of both versions were concordant for 97% of samples. Version 2 appeared to be more sensitive with smear negative specimens, but was more susceptible to inhibitory substances. The authors suggest that internal controls incorporated into the kits by the company would increase the user's confidence in the reliability of negative results.

TABLE 10.23. COMPARISON OF AMTDT VERSIONS 1 AND 2 WITH RESPIRATORY AND NONRESPIRATORY SPECIMENS

		Version 1				Version 2			
Type of Specimen	No. of Specimens	Sensitivity (%) of all Specimens	Sensitivity of Smear Positives	Sensitivity of Smear Negatives	Specificity (%)	Sensitivity (%) of all Specimens	Sensitivity of Smear Positives	Sensitivity of Smear Negatives	Specificity (%)
Respiratory	410	83	100	72.3	100	94.7	100	83	100
Non-respiratory	272	83	100	76.6	100	86.8	100	88.9	100

Gamboa F, Fernandez G, Padilla E, et al. Comparative evaluation of initial and new versions of the Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test for direct detection of MTB in respiratory and non-respiratory specimens. *J Clin Microbiol* 1998; 36: 684

VIDAS Adaptation (VIDAS PROBE)

The Gen-Probe company is in partnership with the Bio-Merieux Vitek company to adapt the amplification and detection steps of the TMA technology to the automated VIDAS, an immunoassay instrument. According to Table 10.5, assays undergoing research and development include the following targets: MTB, *C. trachomatis*, *C. trachomatis*/*N. gonorrhoeae*, *N. gonorrhoeae*, and HIV-1 quantitative. This adaptation will allow use of the TMA technology particularly by small laboratories with small specimen volumes, while the company continues to develop its TIGRIS instrument for larger laboratories with high throughput that require full automation of the extraction as well as the amplification and detection steps.

Polymerase Chain Reaction, Commercial, Roche Amplicor for M. Tuberculosis

Features

Roche has developed a series of PCR-based amplification assays under the trade name Amplicor. Those kits, which are FDA-approved, are included in Table 10.4, and the many more that are undergoing clinical trials, research and development, or awaiting FDA approval, are included in Table 10.5. The MWP format for MTB is FDA-approved, while the automated COBAS format is pending approval. Like the Gen-Probe assay, the Roche kit has been approved for use with respiratory samples, which are smear positive, from previously untreated patients.

The basic format of the Amplicor series includes a specimen collection and transport kit, specimen preparation kit, amplification kit with prepackaged components (Fig. 10.3), and a detection kit. The system includes the AmpErase (uracil-N-glycosylase) method of abrogating amplicon carryover contamination as part of the “master mix” (Fig. 10.4). dUTP is present also in the mix instead of dTTP. Amplicons from previous runs will contain dU instead of dT. AmpErase will destroy uracil-containing amplicons before thermal cycling during the first denaturation cycle, but the enzyme itself is inactivated at cycling temperatures so it cannot destroy new genuine amplicon present in an uncontaminated specimen. Thus, new amplicons are unaffected before detection. The UNG activity is limited to amplicons below a particular size and does not work for prior dT-containing contaminating amplicons. The “master mix” includes

the following reagents: Taq polymerase, AmpErase, dATP, dCTP, dGTP, dUTP, and biotinylated primers (Fig. 10.3). Because primers have biotin, the amplicon will be biotinylated and captured in microtiter wells that contain the probe. Avidin-horseradish peroxidase conjugates are added, substrates are added subsequently, color is produced, and absorbances are determined and read in an enzyme immunoassay plate reader. Thus, the format appears to be user-friendly and amenable to use for detection of amplicon from multiple clinical samples at the same time. A complete summary of the most salient features of the system is found in Table 10.24.

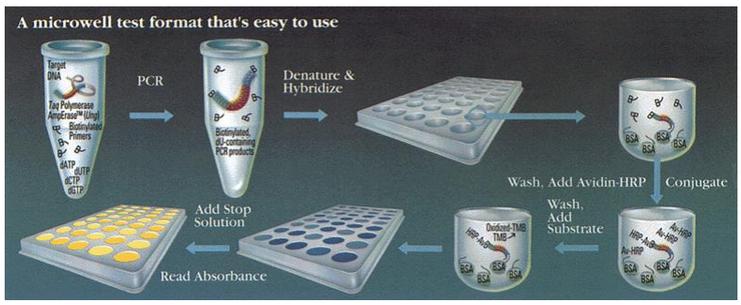


FIGURE 10.3. The microwell-plate format (MWP) of the Roche Amplicor kits allows for detection of amplicons produced by PCR from master mixed reagents provided with the kits. Reproduced with permission from Roche Diagnostics Systems, Inc.

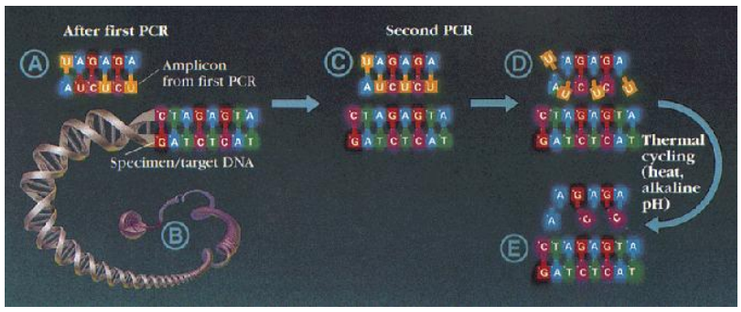


FIGURE 10.4. To control carryover contamination, the Roche kits contain AmpErase (uracil-N-glycosylase) as part of the master mix. This enzyme catalyzes the destruction of uracil-containing DNA strands prior to thermal cycling. Contaminating amplicons are broken apart during the first denaturation cycle to render them inactive, but the enzyme is itself inactivated at cycling temperatures, so newly formed amplicons are unaffected prior to detection. Reproduced with permission from Roche Diagnostic Systems, Inc.

TABLE 10.24. SUMMARY OF FEATURES OF THE ROCHE AMPLICOR MWP FORMAT FOR *MYCOBACTERIUM TUBERCULOSIS*

Type of Amplification	Target
Target	DNA (16s rRNA gene, single copy gene of MTB complex)
Amplification method	PCR
Enzymes	Taq polymerase
Sample volume	100 µL of treated specimen, with NaLC-NaOH as suggested liquefier
Sample lysis	60C, NaOH, Triton X
Format	Microtiter plates
Thermal conditions	Three-temperature setup
Amplified product	DNA
Detection	EIA
Amplicon containment	Uracyl-N-glycosylase (AmpErase)
Instruments	Thermal cycler, etc.
Assay time	6 h
Kit format	96 tests, shelf-life 3 months
Laboratory space requirements	Pre- and postamplification rooms
Training	Offered by the company

Abbreviations: EIA, enzyme immunoassay; MWP, microwell plate; PCR, polymerase chain reaction

Large Published Studies

Twelve large studies have been published that evaluate the MTB MWP kit with real clinical specimens, the majority of which are respiratory samples, mostly sputum (142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152 and 153). The specific details from each study are summarized in Table 10.25. General conclusions that can be drawn from that table are the following: the sensitivity for all specimens ranges from 58% to 95%, with the average being about 75% to 80%; specificities range from 96% to 100%, with an average of 99%; sensitivity of smear-negative specimens ranges from 50% to 70% with selected studies

showing lower values; and most studies concluded that the use of three or more positive specimens leads to an accurate diagnosis. Two studies comparing the Roche MWP MTB format with the Gen-Probe AMTDT version 1.0 (159, 160) showed similar sensitivities and specificities for both assays.

TABLE 10.25. LARGE PUBLISHED STUDIES ON THE ROCHE AMPLICOR *MYCOBACTERIUM TUBERCULOSIS* MWP FORMAT ON CLINICAL SPECIMENS

Reference	Total No. Specimens	No. of MTB Positive Specimens	Gold Standard	Sensitivity (%) for all Specimens	Sensitivity (%) for Smear Negatives	Specificity (%)	Smear Sensitivity (%)	Culture Sensitivity (%)
142	985	55	EGS	66.7	55.3	99.6	25 (K)	60.7
143	535	26	EGS	58	46	99.1	22 (A)	56
144	532	87	EGS	95	50 (3/6)	96	93 (A)	96
145	1,009	186	EGS	85	66	99.6	51 (A)	87
146	504	27	EGS	70.4	60	98	52.4 (K)	88.9
147	2,073	184	C	86	74	98	58	not reported
148	1,480	100	EGS	79	53	99	56 (A)	84.7
149	956	61	EGS	79.4	40	99.6	67 (A)	91
150	784	114	EGS	83.3	not reported	100	not reported (K)	71.6
151	7,194	654	EGS	86.2	60.9	99.4	64.2 (A)	77
152	656	41	EGS	68	46.1	97.4	not reported (K)	not reported
153	113	42	EGS	73.8	not reported	97.2	not reported	64.3

Abbreviations: A, auramine; C, culture; EGS, expanded gold standard; K, Kinyoun.
Reference 150 also evaluated nonrespiratory samples, but these were not included in the calculations presented.

Issues to be Addressed and Attempts at Addressing Them

Several issues of importance that are undergoing scrutiny by the company will be mentioned, and subsequent paragraphs will address the actions being taken by the company. The need for an Internal Control (IC) to test for potential inhibitors in clinical specimens has been addressed by the creation of an IC for use with COBAS and the newest versions of the MWP assays such as those for MTB, *C. trachomatis*, *N. gonorrhoeae*, and hepatitis C virus. The need for advanced automation of the amplification and detection steps is being addressed with the COBAS Amplicor instrument. The company may consider increasing the sample inoculum size from 100µL to 200 or 500 µL with a possible attendant increase in sensitivity. An adaptation of the MTB system for use with nonrespiratory samples is a must, and a revisiting of the packaging format to suit the needs of smaller laboratories also would be advantageous. Second- and third-generation kits most likely will address all of these issues in the future.

An article has been published describing the creation of an internal control (IC) for routine use in combination with the COBAS Amplicor MTB, *C. trachomatis*, *N. gonorrhoeae*, and hepatitis C virus (95). It contains primer-binding regions that are identical to those of the target sequence and contain a unique probe-binding region that differentiates the IC from amplified target nucleic acid. Only 20 copies of the IC are introduced into each test sample, which helps determine inhibition that may occur at the crucial limit of test sensitivity. Inhibition rates of 5% to 9% were exhibited in the COBAS assays for these four organisms, and test sensitivities were 1% to 6% greater with the use of the IC. When introduced into an unprocessed specimen, the IC can help monitor nucleic acid recovery during specimen preparation. Of note, both target and IC DNA will draw from the same limited and common pool of reagents in the master mix, during amplification. Thus, IC amplification may be suppressed because of competition in samples containing large amounts of target DNA, while it may be preferentially amplified in samples containing extremely low levels of target microbial sequences. In general, IC should be included with every specimen that is run, but alternate schemes may be envisioned, such as its selective use in specimens that test negative for the primary target, or its restricted use with specimens most likely to show inhibition such as those with traces of blood, urogenital specimens from pregnant women, etc. In the case of MTB, it may be used when the

probability of infection is high and the consequence of a false-negative result is severe. Such is the case when testing smear-positive specimens where a clinical suspicion for tuberculosis is great, yet a false-negative test result would lead the clinician to conclude that the patient has a nontuberculous mycobacterial infection (95). Commercially available ligase chain reaction and transcription-mediated amplification systems do not have such internal controls as have been described here.

The MTB MWP format was evaluated with CSF in diagnosing tuberculous meningitis (137).

The assay had a sensitivity of 60% and a specificity of 100%, which may not be great, but still the test was more sensitive than stain and culture. The kit also has been evaluated for MTB detection from Bactec 12B bottles or MGIT tubes after giving the specimen, usually a smear-negative specimen, a culture preamplification “boost” prior to submission of an aliquot to PCR amplification. Smith et al. (77) tested 12B broth cultures with a GI of 20 or more with the MTB MWP kit, while testing those broth cultures reaching a GI of 300 or more with DNA-probes. The sensitivity and specificity of PCR were 93% and 100%. The mean time from specimen inoculation to detection and identification was 16 days for PCR and 28 days for the DNA probe assay. Ichiyama et al. (99) tested positive MGIT broths (Mycobacterium Growth Indicator Tubes, Becton Dickinson) with both PCR and DNA-probes and found the sensitivities to be 98% for PCR and 77% with the probes. PCR assays specific to the entire genus *Mycobacterium* and useful in the detection of a wide range of mycobacterial species in clinical samples have been described (154). Genus-specific primers and probes amplify, capture, and detect amplicons, while products then can be hybridized to species-specific probes for more selective identification.

The automated COBAS Amplicor PCR system will automate the amplification and detection steps of the PCR reaction, providing containment of the detection components. Manual sample preparation for the DNA extraction step is still required, but perhaps will be carried out in the future by a separate instrument. COBAS allows for the simultaneous amplification of different infectious agents, given that it contains several thermal cyclers, each of which can be programmed to amplify a different target at its required parameters, and there will be coamplification with an internal control. Biotinylated amplicons are captured onto magnetic particles that are coated with specific probes, and the bound products are detected colorimetrically. The is no evidence of instrument carryover between high positive specimens and negative samples placed in tandem. There was 100% correlation between samples run by COBAS for MTB and those run with the MWP assay format (155). The amount of sediment used for MTB in COBAS is the same as with the MWP format (i.e., 100 µL of pretreated respiratory specimen). The actual hands-on time for the MTB assay has been reported to be 4.4 minutes (155).

At least four large studies have compared COBAS MTB with culture (15, 156, 157 and 158), and the comparison of all studies is included in Table 10.26. Bodmer et al. (156) tested COBAS in a population with high prevalence for tuberculosis. There was inhibition of amplification in 4.7% of samples that included several smear-positive specimens, which suggested to them that such a rate of inhibition justifies the systematic inclusion of the internal amplification control. Yuen et al. (157) compared COBAS with LCX from Abbott and found it to be more sensitive than LCX for all respiratory specimens as well as for smear-negative ones. Rajalahti et al. (15) reported an increase in sensitivity when three consecutive sputa were received per patient. Finally, Eing et al. (158) published the first report on the COBAS MTB assay applied to both respiratory and nonrespiratory specimens. With a modified protocol for DNA extraction (i.e., washing plus sonication), sensitivities with smear-negative gastric aspirates were superior to those of sputa. They also suggest that three or more smear-negative sputum specimens be tested per patient prior to accepting a negative result as a true negative.

TABLE 10.26. COMPARISON OF THE COBAS *MYCOBACTERIUM TUBERCULOSIS* ASSAY WITH CULTURE OR WITH THE MWP FORMAT FOR DETECTION OF *MYCOBACTERIUM TUBERCULOSIS* IN CLINICAL SPECIMENS

Reference	No. of Specimens	Type of Specimens	Gold Standard	Sensitivity (%) for all Specimens	Sensitivity (%) for Smear Positives	Sensitivity (%) for Smear Negatives	Specificity (%)	Culture Sensitivity	Correlation with MWP Format
155	230	respiratory	-	-	-	-	-	-	100%
156	1,012	respiratory	C	92.6	-	-	99.6	Gold standard (100%)	-
157	204	respiratory	C	84.3	-	57.9	100	Gold standard (100%)	-
15	324	respiratory	C	83	-	68	99	Gold standard (100%)	-
158	1,681	respiratory and non-respiratory	EGS	66.3	96.4	45.4	99.7	-	-

Abbreviations: C, culture; EGS, expanded gold standard (see text); MWP, microwell plate.

CONCLUSIONS ON THE MOLECULAR BIOLOGY OF INFECTIOUS DISEASES

Part of "10 - Molecular Biology of Infectious Diseases"

There should be no doubt in the reader's mind that the use of DNA probes and molecular amplification techniques in infectious

diseases has revolutionized the field and is here to stay. In-house assays are ever-present around the world, but it is the presence of the technology in stereotypical kit forms from companies that has begun to increase the accessibility of this technology to the small microbiology and virology laboratory. A look at Table 10.3, Table 10.4 and Table 10.5 should excite the reader as to what is currently available and what is upcoming in terms of robust assays in kit format.

The majority of scenarios that will require the use of amplification have been stated in Table 10.2. However, no one should think that these techniques will completely free the laboratory from establishing labor-intensive culture procedures for a myriad applications. Using an assay that detects MTB-complex as an example, all non-TB mycobacteria present in the sample would be missed, members within the MTB-complex would not be distinguished from one another, and viable organisms would still be required in culture for susceptibility testing, despite progress in the molecular understanding and rapid detection of resistance against antituberculosis agents with amplification techniques. Lastly, positive signals from amplification of dead or dying treated organisms, or of saprophytic microbes that could be present in the absence of pathology, will always require clinical correlation for an accurate interpretation, and the eventual use of amplification kits for performing the “test of cure” for several conditions will require stringent evaluation as part of large clinical trials.

REFERENCES

1. Sandin RL, Greene JN. Diagnostic molecular pathology and infectious disease. *Cancer Control* 1995;2:250.

2. Heiter BJ, Borbeau PP. Comparison of the Gen-Probe Group A Streptococcus Direct Test with culture and a rapid streptococcal antigen detection assay for the diagnosis of streptococcal pharyngitis. *J Clin Microbiol* 1993;31:2070.

3. Enns RK. DNA probes: an overview and comparison with current methods. *Lab Med* 1988;19:295.

4. Sandin RL. The ABC's of molecular amplification. *Infect Med* 1999;16:98.

5. Cousins D, Francis B, Dawson D. Multiplex PCR provides a low cost alternative to DNA probe methods for rapid identification of *Mycobacterium avium* and *M. intracellulare*. *J Clin Microbiol* 1996;34:2331.

6. Devallois A, Picardeau M, Goh KS, et al. Comparative evaluation of PCR and commercial DNA probes for detection and identification to species level of *Mycobacterium avium* and *M. intracellulare*. *J Clin Microbiol* 1996;34:2756.

7. Goto M, Oka S, Okuzumi K, et al. Evaluation of acridinium-ester labelled DNA probes for identification of MTB and MAI complex in culture. *J Clin Microbiol* 1991;29:2473.

8. Beebe JL, Sharpton TR, Zanto SN, et al. Performance characteristics of the Gen-Probe Probe Competition Assay used as a supplementary test for the Gen-Probe PACE2 and 2A Assays for detection of Chlamydia trachomatis. *J Clin Miocrobiol* 1997;35:477.

9. Labombardi V, Carter L, Massarella S. Use of nucleic acid probes to identify mycobacteria directly from DIFCO ESP-Myco Bottles. *J Clin Microbiol* 1997;35:1002.

10. Sandin RL, Rinaldi M. Special considerations for the clinical microbiology laboratory in the diagnosis of infections in the cancer patient. *Infect Dis Clin NA* 1996;10:413.

11. Tenover FC, Crawford JT, Huebner RFE, et al. The resurgence of tuberculosis: is your lab ready? *J Clin Microbiol* 1993;31:767.

12. Hall GS. Newer methodologies in isolation and identification of *Mycobacterium* species. *Inf Dis Clin Pract* 1997;6:369.

13. Wolcott M. Advances in nucleic acid-based detection methods. *Clin Microbiol Rev* 1992;5:370.

14. Sandin RL. Laboratory technology. Non-PCR amplification techniques. *Adv Admin Labor* 1997;6:69.

15. Rajalahti I, Vuorinen P, Nieminen MM, et al. Detection of MTB complex in sputum specimens by the Automated Roche Cobas Amplicor MTB test. *J Clin Microbiol* 1998;36:975.

16. Alkan S, Diaz JI, Sandin RL. Techniques of diagnostic molecular pathology. *Cancer Control* 1995;2:149.

17. Oste C. Polymerase chain reaction. *Biotechniques* 1988;6:162.

18. Chapin K, Lauderdale T. Evaluation of a rapid air thermal cycler for detection of MTB. *J Clin Microbiol* 1997;35:2157.

19. Kearns AM, Freeman R, Steward M. Evaluation of a rapid air thermal cycler for detection of MTB. *J Clin Microbiol* 1998;36:604.

20. Brisson-Noel A, Lecossier D, Nassif X, et al. Rapid diagnosis of tuberculosis by amplification of mycobacterial DNA in clinical samples. *Lancet* 1989;ii:1069.

21. Eisenach KD, Crawford JT, Bates JH. Repetitive DNA sequences as probes for *Mycobacterium tuberculosis*. *J Clin Microbiol* 1988;26:2240.

22. Hance AJ, Grandchamp B, Levy-Frebault V, et al. Detection and identification of mycobacteria by amplification of mycobacterial DNA. *Molec Microbiol* 1989;3:843.

23. Eisenach KD. PCR detection of Mycobacterium tuberculosis. In: Persing DH, Smith TF, Tenover F, White TJ, eds. *Diagnostic Molecular Microbiology. Principles and Applications*. Washington, DC: Am Soc Microbiol, 1993.

24. Buck G, O'Hara L, Summersgill J. Rapid, simple method for treating clinical specimens containing *Mycobacterium tuberculosis* to remove DNA for polymerase chain reaction. *J Clin Microbiol* 1992;30:1331.

25. Folgueira L, Delgado R, Palenque E, et al. Detection of *Mycobacterium tuberculosis* DNA in clinical samples by using a simple lysis method and polymerase chain reaction. *J Clin Microbiol* 1993;31:1019.

26. Clarridge III J, Shawar R, Shinnick P, et al. Large scale use of polymerase chain reaction for detection of *Mycobacterium tuberculosis* in a routine mycobacteriology laboratory. *J Clin Microbiol* 1993;31:2041.

27. Kocagoz T, Yilmaz E, Ozkara S, et al. Detection of *Mycobacterium tuberculosis* in sputum samples by polymerase chain reaction using a simplified procedure. *J Clin Microbiol* 1993;31:1435.

28. Victor T, duToit R, vanHelden P. Purification of sputum samples through sucrose improves detection of *Mycobacterium tuberculosis* by polymerase chain reaction. *J Clin Microbiol* 1992;30:1514.

29. Wilson SM, McNerney R, Nye P, et al. Progress towards a simplified polymerase chain reaction and its application to diagnosis of tuberculosis. *J Clin Microbiol* 1993;31:776.

30. Werking C, Sandin RL, Hearn C. Optimization of an in-house one-tube nested PCR assay for detection of M. tuberculosis. *Clin Infect Dis* 1994;19:566.

31. Greenfield L, White TJ. Sample preparation methods. In: Persing DH, Smith TF, Tenover F, White TJ, eds. *Diagnostic Molecular Microbiology. Principles and Applications*. Washington, DC: Am Soc Microbiol, 1993.

32. Hermans P, Schuitema A, vanSoolingen D, et al. Specific detection of *Mycobacterium tuberculosis* complex strains by polymerase chain reaction. *J Clin Microbiol* 1990;28:1204.

33. Miyazaki Y, Koga H, Kohno S, et al. Nested polymerase chain reaction for detection of *Mycobacterium tuberculosis* in clinical samples. *J Clin Microbiol* 1993;31:2228.

34. Marchetti G, Gori A, Catozzi L, et al. Evaluation of PCR in detection of MTB from formalin-fixed, paraffin-embedded tissues: comparison of four amplification assays. *J Clin Microbiol* 1998;36:1512.

35. Eisenach KD, Cane MD, Bates J, et al. Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J Infect Dis* 1990;161:977.

36. Thierry D, Brisson-Noel A, Vincent V, et al. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS6110, and its application in diagnosis. *J Clin Microbiol* 1990;28:2668.

37. Sandin RL. Polymerase chain reaction and other amplification techniques in mycobacteriology. In: Heifitz, L. *Clinics in Laboratory Medicine. Clinical Mycobacteriology*. Philadelphia: WB Saunders Co, 1996:617.

38. Hellyer TJ, DesJardin LE, Beggs ML, et al. IS6110 homologs are present in multiple copies in mycobacteria other than tuberculosis-causing mycobacteria. Letter to the Editor. *J Clin Microbiol* 1998;36:853.

39. VonSoolingen D, deHaas P, Hermans P, et al. Comparison of various repetitive DNA element as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1993;31:1987.

40. Kent L, McHugh TD, Billington O, et al. Demonstration of homology between IS6110 of *Mycobacterium tuberculosis* and DNAs of other *Mycobacterium*. *J Clin Microbiol* 1995;33:2290.

41. Mulcahy GM, Kaminski ZC, Albanese EA, et al. IS6110-based PCR methods for detection of Mycobacterium tuberculosis. *J Clin Microbiol* 1996;34:1348.

42. Liebana E, Aranaz A, Francis B, et al. Assessment of genetic markers for species differentiation within the *Mycobacterium tuberculosis* complex. *J Clin Microbiol* 1996; 34:933.

43. Hellyer TJ, DesJardin LE, Assaf MK, et al. Specificity of IS6110-based amplification assays for MTB-complex. *J Clin Microbiol* 1996;34:2843.

44. Gillespie SH, McHugh TD, Newport LE. Specificity of IS6110-based amplification assays for MTB complex. Letter to the Editor. *J Clin Microbiol* 1997;35:799.

45. McHugh TD, Newport LE, Gillespie SH. IS6110 Homologs are present in multiple copies in mycobacteria other than tuberculosis-causing mycobacteria. *J Clin Microbiol* 1997;35:1769.

46. Patel S, Yates M, Saunders NA. PCR-enzyme-linked immunosorbent assay and partial rRNA gene sequencing: a rational approach to identifying mycobacteria. *J Clin Microbiol* 1997;35:2375.

47. Boddinghaus B, Rogall T, Flohr T, et al. Detection and identification of mycobacteria by amplification of rRNA. *J Clin Microbiol* 1990;28:1751.

48. Kirschner P, Springer B, Vogel U, et al. Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a two year experience in a clinical laboratory. *J Clin Microbiol* 1993;31:2882.

49. Jou N-T, Yoshimori RB, Mason GR, et al. Single-tube, nested, RT-PCR for detection of viable MTB. *J Clin Microbiol* 1997;35:1161.

50. Takewaki S, Okuzumi K, Ishiko H, et al. Genus specific polymerase chain reaction for the mycobacterial DNA J gene and species specific oligonucleotide probes. *J Clin Microbiol* 1993;31:446.

51. Hawkey PM. The role of PCR in the diagnosis of mycobacterial infections. *Rev Med Microbiol* 1994;5:21.

52. DeWit D, Steyn L, Shoemaker S, et al. Direct detection of *Mycobacterium tuberculosis* in clinical specimens by DNA amplification. *J Clin Microbiol* 1990;28:2437.

53. Del Portillo P, Murillo L, Elkin M. Amplification of a species specific DNA fragment of *Mycobacterium tuberculosis* and its possible use in diagnosis. *J Clin Microbiol* 1991;29:2163.

54. Del Portillo P, Thomas MC, Martinez E, et al. Multiprimer PCR system for differential identification of mycobacteria in clinical samples. *J Clin Microbiol* 1996;34:324.

55. Herrera EA, Segovia M. Evaluation of mtp-40 genomic fragment amplification for specific detection of *Mycobacterium tuberculosis* in clinical specimens. *J Clin Microbiol* 1996;34:1108.

56. Gori A, Franzetti F, Marchetti G, et al. Specific detection of MTB by mtp40 nested PCR. *J Clin Microbiol* 1996;34:2866.

57. Vera-Cabrera L, Howard ST, Laszlo A, et al. Analysis of genetic polymorphism in the phospholipase region of MTB. *J Clin Microbiol* 1997;35:1190.

58. Weil A, Plikaytis BB, Butler WB, et al. The mtp40 gene is not present in all strains of MTB. *J Clin Microbiol* 1996;34:2309.

59. Espinosa LE, Galan JC, Gutierrez M, et al. Allele-specific PCR based detection method based on pncA and oxyR sequences for distinguishing M. bovis from MTB: intraspecific *M. bovis* pncA sequence polymorphism. *J Clin Microbiol* 1998;36:239.

60. Soini H, Viljanen MK. Diversity of the 32-kilodalton protein gene may form a basis for species determination of potentially pathogenic mycobacterial species. *J Clin Microbiol* 1997;35:769.

61. Talbot EA, Williams DL, Frothingham R. PCR identification of Mycobacterium bovis BCG. *J Clin Microbiol* 1997;35:566.

62. Lotte A, Wasz-Hockert O, Poisson N, et al. BCG complications: estimates for the risks among vaccinated subjects and statistical analysis of their main characteristics. *Adv Tuberc Res* 1984;21:107.

63. Albert J, Fenyo EM. Simple sensitivity and specific detection of human immunodeficiency virus type I in clinical specimens by polymerase chain reaction with nested primers. *J Clin Microbiol* 1990;28:1560.

64. Porter-Jordan K, Rosenberg E, Keiser J, et al. Nested polymerase chain reaction assay for the detection of cytomegalovirus overcomes false positives caused by contamination with fragmented DNA. *J Med Virol* 1990;30:85.

65. Sandin RL, Rodriguez ER, Rosenberg E, et al. Comparison of sensitivity for human cytomegalovirus of the polymerase chain reaction, traditional tube culture and shell vial assay by sequential dilutions of infected cell lines. *J Virol Meth* 1991;32:181.

66. Ros Bascunana C, Belak K. Detection and identification of mycobacteria in formalin-fixed, paraffin-embedded tissues by nested PCR and restriction enzyme analysis. *J Clin Microbiol* 1996;34:2351.

67. Pierre C, Lecossier D, Boussougant Y, et al. Use of reamplification protocol improves sensitivity of detection of *Mycobacterium tuberculosis* in clinical samples by amplification of DNA. *J Clin Microbiol* 1991;29:712.

68. Kox LF, Jansen HM, Kuijper S, et al. Multiplex PCR assay for immediate identification of the infecting species in patients with mycobacterial disease. *J Clin Microbiol* 1997;35:1492.

69. Garrett CT, Rodriguez ER, Comerford J, et al. Establishing a molecular diagnostics laboratory. *Adv Pathol Lab Med* 1992;5:35.

70. Patel RJ, Fries J, Piessens W, et al. Sequence analysis and amplification by polymerase chain reaction of a cloned DNA fragment for identification of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1990;28:513.

71. Fiss E, Chebab F, Brooks G. DNA amplification and reverse dot plot hybridization for detection and identification of mycobacteria to the species level in the clinical laboratory. *J Clin Microbiol* 1992;30:1220.

72. Telenti A, Marchesi F, Balz M, et al. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *J Clin Microbiol* 1993;31:175.

73. Vaneechoutee M, DeBeenhouwer H, Claeys G, et al. Identification of *Mycobacterium* sp. by using amplified ribosomal DNA restriction analysis. *J Clin Microbiol* 1993;31:2061.

74. Gerner-Smidt P, Graves LM, Hunter S, et al. Computerized analysis of restriction fragment length polymorphism patterns: comparative evaluation of two commercial software packages. *J Clin Microbiol* 1998;36:1318.

75. Klatsen PR, Kuijper S, vanIngen CW, et al. Stabilized, freeze-dried PCR mix for detection of mycobacteria. *J Clin Microbiol* 1998;36:1798.

76. Shankar P, Manjunath N, Mohan K, et al. Rapid diagnosis of tuberculous meningitis by polymerase chain reaction. *Lancet* 1991;337:5.

77. Smith MB, Bergmann JS, Woods GL. Detection of MTB in Bactec 12B broth cultures by the Roche Amplicor PCR Assay. *J Clin Microbiol* 1997;35:900.

78. Gibb AP, Wong S. Inhibition of PCR by agar from bacteriological transport media. *J Clin Microbiol* 1998;36:275.

79. Wadowsky RM, Laus S, Libert T, et al. Inhibition of PCR-based assay for Bordetella pertussis by using calcum alginate fiber and aluminum shaft components of a nasopharyngeal swab. *J Clin Microbiol* 1994;32:1054.

80. Deneer HG, Knight I. Inhibition of the polymerase chain reaction by mucolytic agents. *Clin Chem* 1994;40:171.

81. Lou Q, Chong SKF, Fitzgerald JF, et al. Rapid and effective method for preparation of fecal specimens for PCR assays. *J Clin Microbiol* 1997;35:281.

82. Kirschner P, Rosenau J, Springer B, et al. Diagnosis of mycobacterial infections by nucleic acid amplification: 18-month prospective study. *J Clin Microbiol* 1996;34:304.

83. Forbes BA, Hicks KES. Direct detection of Mycobacterium tuberculosis in respiratory specimens in a clinical laboratory by polymerase chain reaction. *J Clin Microbiol* 1993;31:1688.

84. Nolte FS, Metchock B, McGowan JE Jr, et al. Direct detection of *Mycobacterium tuberculosis* in sputum by polymerase chain reaction and DNA hybridization. *J Clin Microbiol* 1993;31:1777.

85. Afghani B, Stutman HR. Diagnosis of tuberculosis: can the polymerase chain reaction replace acid-fast bacilli smear and culture? *J Infect Dis* 1995;172:903.

86. Cegielski JP, Devlin BH, Morris AJ, et al. Comparison of PCR, culture and histopathology for diagnosis of tuberculosis pericarditis. *J Clin Microbiol* 1997;35:3254.

87. Noordhoek GT, van Embden J, Kolk A. Questionable reliability of the polymerase chain reaction in the detection of *Mycobacterium tuberculosis*. *N Engl J Med* 1993;329:2036.

88. Noordhoek GT, Kolk AH, Bjune G, et al. Sensitivity and specificity of PCR for detection of Mycobacterium tuberculosis: a blind comparison study among seven laboratories. *J Clin Microbiol* 1994;32:277.

89. Macher A, Goosby E. PCR and the misdiagnosis of active tuberculosis. *N Engl J Med* 1995;332:128.

90. Noordhoek GT, van Embden JDA, Kolk AHJ. Reliability of nucleic acid amplification for detection of MTB: an international collaborative quality control study among 30 laboratories. *J Clin Microbiol* 1996;34:2522.

91. Forbes BA. Current and future applications of mycobacterial amplification assays. *Clin Microbiol Newsltr* 1995;17:145.

92. Maher M, Glennon M, Martinazzo G, et al. Evaluation of a novel PCR-based diagnostic assay for detection of Mycobacterium tuberculosis in sputum samples. *J Clin Microbiol* 1996;34:2307.

93. Sandin RL, Chan MM, Rodriguez ER, et al. Comparison of HIV-1 and human CMV serostatus in children, with nested-primer PCR amplification and flow cytometric phenotype analysis of peripheral blood mononuclear cell-enriched fractions. *Pediatr AIDS and HIV Infection: Fetus to Adolescent* 1992;3:53.

94. Rodriguez ER, Nasim S, Hsia J, et al. Cardiac myocytes and dendritic cells harbor HIV in infected patients with and without cardiac dysfunction: detection by multiplex, nested, PCR in individually microdissected cells from right ventricular endomyocardial biopsy tissue. *Am J Cardiol* 1991;68:1511.

95. Rosenstraus M, Wang Z, Chang S-Y, et al. An internal control for routine diagnostic PCR: design, properties and effect on clinical performances. *J Clin Microbiol* 1998;36:191.

96. Forbes BA, Hicks KE. Ability of PCR assay to identify *Mycobacterium tuberculosis* in BACTEC 12B vials. *J Clin Microbiol* 1994;32:1725.

97. Sanguinetti M, Posteraro B, Ardito F, et al. Routine use of PCR-reverse cross-blot hybridization assay for rapid identification of Mycobacterium species growing in liquid media. *J Clin Microbiol* 1998;36:1530.

98. Bergmann JS, Yuoh G, Woods GL. Detection of MTB complex in Bactec 12B and AccuMed ESP II broth cultures by the Gen Probe enhanced AMTDT (E-MTD) test. *Abstr 98th Ann Mtg Am Soc Microbiol* 1998.

99. Ichiyami S, Iinuma Y, Yamori S, et al. Mycobacterium growth indicator tube testing in conjunction with the AccuProbe or the Amplicor-PCR assay for detecting and identifying mycobacteria from sputum samples. *J Clin Microbiol* 1997;35:2022.

100. Mangiapan G, Vokurka M, Schouls L, et al. Sequence capture-PCR improves detection of mycobacterial DNA in clinical specimens. *J Clin Microbiol* 1996;34:1209.

101a. Mazurek GH, Reddy V, Murphy D, et al. Detection of *Mycobacterium tuberculosis* in cerebrospinal fluid following immunomagnetic enrichment. *J Clin Microbiol* 1996;34:450.

101b. Roberts B, Hirst R. Immunomagnetic separation and PCR for detection of *Mycobacterium ulcerans*. *J Clin Microbiol* 1997;35:2709.

102. DesJardin LE, Perkins MD, Teixeira L, et al. Alkaline decontamination of sputum specimens adversely affects stability of mycobacterial mRNA. *J Clin Microbiol* 1996;34:2435.

103. Winn W. *Legionella* and the clinical microbiologist. *Infect Dis Clin NA* 1993;7:377.

104. Fain JS, Bryan RN, Cheng L, et al. Rapid diagnosis of Legionella infection by a nonisotopic in situ hybridization method. *Am J Clin Pathol* 1991;95:719.

105. Provost F, Laurent F, Camacho LR, et al. Molecular study of persistence of Nocardia asteroides and *Nocardia otitidiscaviarum* strains in patients with long-term nocardiosis. *J Clin Microbiol* 1997;35:1157.

106. Sandin RL. Amplification for *Mycobacterium* spp. with special emphasis on *M. tuberculosis*. *JAMA SE Asia* 1994;10:26.

107. Fournier P-E, Marrie TJ, Raoult D. Diagnosis of Q fever. *J Clin Microbiol* 1998;36:1823.

108. Mitchell TG, Sandin RL, Bowman BH, et al. Molecular mycology: DNA probes and applications of PCR technology. *J Med Vet Mycol* 1994;32:251.

109. Sandin RL, Isada CM, Hall GS, et al. Aberrant *Histoplasma capsulatum*: confirmation of identity by a chemiluminescence-labelled DNA probe. *Diagn Microbiol Infec Dis* 1993;17:235.

110. Sandin RL, Hall GS, Longworth DL, et al. Unpredictability of commercially-available exoantigen culture confirmation tests in confirming the identity of 5 Blastomyces dermatitidis isolates. *Am J Clin Pathol* 1993;99:542.

111. Ribes JA, Limper AH, Espy MJ, et al. PCR detection of *Pneumocystis carinii* in bronchoalveolar lavage specimens: analysis of sensitivity and specificity. *J Clin Microbiol* 1997;35:830.

112. Mitchell PS, Espy MJ, Smith TF, et al. Laboratory diagnosis of CNS infections with HSV by PCR performed with CSF specimens. *J Clin Microbiol* 1997;35:2873.

113. Franzen C, Altfeld M, Hegener P, et al. Limited value of PCR for detection of *Toxoplasma gondii* in blood from HIV-infected patients. *J Clin Microbiol* 35:2639.

114. Folgueira L, Delgado R, Palenque E, et al. Rapid diagnosis of Mycobacterium tuberculosis bacteremia by PCR. *J Clin Microbiol* 1997;34:512.

115. Kox LFF, van Leeuwen J, Knijper S, et al. PCR assay based on DNA coding for 16S rRNA for detection and identification of mycobacteria in clinical samples. *J Clin Microbiol* 1995;33:3225.

116. Tevere VJ, Hewitt PL, Dare A, et al. Detection of *Mycobacterium tuberculosis* by PCR amplification with pan-*Mycobacterium* primers and hybridization to an *M. tuberculosis* -specific probe. *J Clin Microbiol* 1996;34:918.

117. Devallois A, Goh KS, Rastogi N. Rapid identification of mycobacteria to species level by PCR-restriction fragment length polymorphism analysis of the hsp 65 gene and proposition of an algorithm to differentiate 34 mycobacterial species. *J Clin Microbiol* 1997;35:2969.

118. Magdalena J, Vachee A, Supply P, et al. Identification of a new DNA region specific for members of MTB-complex. *J Clin Microbiol* 1998;36:937.

119. Yuen LKW, Ross BC, Jackson KM, et al. Characterization of Mycobacterium tuberculosis strains from Vietnamese patients by Southern blot hybridization. *J Clin Microbiol* 1993;31:1615.

120. Fumokong NG, Tang TH, Al-Maamary S, et al. Insertion sequence typing of *M. tuberculosis*: characterization of a widespread sub-type with single copy of IS6110. *Tubercle Lung Dis* 1994;75:435.

121. Springer B, Stockman L, Teschner K, et al. Two-laboratory collaborative study on identification of mycobacteria: molecular versus phenotypic methods. *J Clin Microbiol* 1996;34:296.

122. Avaniss-Aghajani E, Jones K, Holtzman A, et al. Molecular techniques for rapid identification of mycobacteria. *J Clin Microbiol* 1996;34:98.

123. Dobner P, Feldmann K, Rifai M, et al. Rapid identification of mycobacterial species by PCR amplification of hypervariable 16S rRNA gene promoter region. *J Clin Microbiol* 1996;34:866.

124. Frothingham R. Differentiation of strains in Mycobacterium tuberculosis complex by DNA sequence polymorphisms, including rapid identification of *M. bovis* BCG. *J Clin Microbiol* 1995;33:840.

125. Gamboa F, Manterola J, Lonca J, et al. Detection and identification of mycobacteria by amplification of RNA and DNA in pre-treated blood and bone marrow aspirates by a simple lysis method. *J Clin Microbiol* 1997;35:2124.

126. Abe C, Hirano K, Wada M, et al. Detection of *Mycobacterium tuberculosis* in clinical specimens by polymerase chain reaction and Gen-Probe amplified *Mycobacterium tuberculosis* direct test. *J Clin Microbiol* 1993;31:3270.

127. Jonas V, Alden M, Curry J, et al. Detection and identification of Mycobacterium tuberculosis directly from sputum sediments by amplification of rRNA. *J Clin Microbiol* 1993;31:2410.

128. Miller N, Hernandez S, Cleary T. Evaluation of Gen-Probe amplified *Mycobacterium tuberculosis* direct test and PCR for direct detection of *Mycobacterium tuberculosis* in clinical specimens. *J Clin Microbiol* 1994;32:393.

129. Pfyffer G, Kissling P, Wirth R, et al. Direct detection of Mycobacterium tuberculosis complex in respiratory specimens by a target amplified test system. *J Clin Microbiol* 1994;32:918.

130. Bodmer T, Gurtner A, Schopper K, et al. Screening of respiratory tract specimens for the presence of *Mycobacterium tuberculosis* by using the Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test. *J Clin Microbiol* 1994;32:1483.

131. LaRocco MT, Wanger A, Ocera H, et al. Evaluation of a commercial rRNA amplification assay for direct detection of *Mycobacterium tuberculosis* in processed sputum. *Eur J Clin Microbiol Infec Dis* 1994;13:726.

132. Vlaspolder F, Singer P, Roggeveen C. Diagnostic value of an amplification method (Gen-Probe) compared with that of culture for diagnosis of tuberculosis. *J Clin Microbiol* 1995;33:2699.

133. Bodmer T, Mockl E, Muhlemann K, et al. Improved performance of Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test when 500 instead of 50 microliters of decontaminated sediment is used. *J Clin Microbiol* 1996;34:222.

134. Pfyffer GE, Kissling P, Jahn EMI, et al. Diagnostic performance of amplified *Mycobacterium tuberculosis* direct test with cerebrospinal fluid, other nonrespiratory, and respiratory specimens. *J Clin Microbiol* 1996;34:834.

135. Ehlers S, Ignatius R, Regnath T, et al. Diagnoses of extrapulmonary TB by Gen-Probe AMTD. *J Clin Microbiol* 1996;34:2275.

136. Gamboa F, Manterola JM, Vinado B, et al. Direct detection of MTB complex in non-respiratory specimens by Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test. *J Clin Microbiol* 1997;35:307.

137. Bonington A, Strang JIG, Klapper, et al. Use of Roche Amplicor MTB PCR in early diagnosis of tuberculous meningitis. *J Clin Microbiol* 1998;36:1251.

138. Kilpatrick ME, Girgis NI, Yassin MW, et al. Tuberculous meningitis—clinical and laboratory review of 100 patients. *J Hyg Camb* 1986;96:231.

139. Teoh R, Humphries M. Tuberculous meningitis. In: Lambert HP, ed. Kass Handbook of Infectious Diseases: Infections of the Central Nervous System. Philadelphia: BC Decker, 1991:189.

140. Gamboa F, Fernandez G, Padilla E, et al. Comparative evaluation of initial and new versions of the Gen-Probe Amplified Mycobacterium Tuberculosis Direct Test for direct detection of MTB in respiratory and non-respiratory specimens. *J Clin Microbiol* 1998;36:684.

141. Yuoh G, Fish G, Woods GL. Detection of MTB complex in respiratory specimens from prison inmates by Gen-Probe enhanced AMTDT (E-MTD). Abstr 98th Ann Mtg Am Soc Microbiol, 1998.

142. D'Amato RF, Wallman AA, Hochstein LH, et al. Rapid diagnosis of pulmonary tuberculosis by using Roche AMPLICOR *Mycobacterium tuberculosis* PCR test. *J Clin Microbiol* 1995;33:1832.

143. Chin DP, Yajko DM, Hadley WK, et al. Clinical utility of a commercial test based on the polymerase chain reaction for detecting Mycobacterium tuberculosis in respiratory specimens. *Am J Respir Crit Care Med* 1995;151:1872.

144. Gleason Beavis K, Lichty MB, Jungkind DL, et al. Evaluation of Amplicor PCR for direct detection of Mycobacterium tuberculosis from sputum specimens. *J Clin Microbiol* 1995;33:2582.

145. Moore DF, Curry JI. Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by Amplicor PCR. *J Clin Microbiol* 1995;33:2686.

146. Schirm J, Oostendorp LAB, Mulder JG. Comparison of Amplicor, in-house PCR, and conventional culture for detection of *Mycobacterium tuberculosis* in clinical samples. *J Clin Microbiol* 1995;33:3221.

147. Carpentier E, Drouillard B, Dailloux M, et al. Diagnosis of tuberculosis by Amplicor *Mycobacterium tuberculosis* test: a multicenter study. *J Clin Microbiol* 1995;33:3106.

148. Wobeser WL, Krajden M, Conly J, et al. Evaluation of Roche Amplicor PCR assay for *Mycobacterium tuberculosis*. *J Clin Microbiol* 1996;34:134.

149. Bergmann JS, Woods GL. Clinical evaluation of the Roche AMPLICOR PCR Mycobacterium tuberculosis test for detection of *M. tuberculosis* in respiratory specimens. *J Clin Microbiol* 1996;34:1083.

150. Devallois A, Legrand E, Rastogi N, et al. Evaluation of Amplicor MTB test as adjunct to smears and culture for direct detection of *Mycobacterium tuberculosis* in the French Caribbean. *J Clin Microbiol* 1996;34:1065.

151. Bennedsen J, Ostergaard Thomsen V, Pfyffer GE, et al. Utility of PCR in diagnosing pulmonary tuberculosis. *J Clin Microbiol* 1996;34:1407.

152. Cartuyvels R, deRidder C, Jonckheere S, et al. Prospective clinical evaluation of Amplicor MTB PCR test as a screening method in a low-prevalence population. *J Clin Microbiol* 1996;34:2001.

153. Huang T-S, Liu Y-C, Lin H-H, et al. Comparison of the Roche Amplicor Mycobacterium assay and Digene Sharp Signal System with in-house PCR and culture for detection of MTB in respiratory specimens. *J Clin Microbiol* 1996;34:3092.

154. Stauffer F, Haber H, Rieger A, et al. Genus level identification of mycobacteria from clinical specimens by using an essay-to-handle Mycobacterium-specific PCR assay. *J Clin Microbiol* 1998;36:614.

155. Jungkind D, DiReuzo S, Beavis KG, et al. Evaluation of Automated COBAS Amplicor PCR System for detection of several infectious agents and its impact on laboratory management. *J Clin Microbiol* 1996;34:2778.

156. Bodmer T, Gurtner A, Scholkmann M, et al. Evaluation of the COBAS Amplicor MTB System. *J Clin Microbiol* 1997;35:1604.

157. Yuen K-Y, Yam W-C, Wong L-P, et al. Comparison of two automated DNA amplification systems with a manual one-tube nested PCR assay for diagnosis of pulmonary tuberculosis. *J Clin Microbiol* 1997;35:1385.

158. Eing BR, Becker A, Johns A, et al. Comparison of Roche Cobas Amplicor MTB Assay with in-house PCR and culture for detection of MTB. *J Clin Microbiol* 1998;36:2023.

159. Vuorinen P, Miettinen A, Vuento R, et al. Direct detection of *Mycobacterium tuberculosis* complex in respiratory specimens by Gen-Probe Amplified *Mycobacterium Tuberculosis* Direct Test and Roche Amplicor *Mycobacterium Tuberculosis* Test. *J Clin Microbiol* 1995;33:1856.

160. Ichiyama S, Iinuma Y, Tawada Y, et al. Evaluation of Gen-Probe Amplified *Mycobacterium Tuberculosis* Direct Test and Roche PCR-Microwell plate hybridization method (AMPLICOR MYCOBACTERIUM) for direct detection of mycobacteria. *J Clin Microbiol* 1996;34:130.

11

Clinical Applications of Molecular Biology Hematopoietic Disorders

Serhan Alkan

Curtis A. Hanson

Advances in molecular biology of hematologic neoplasms have led to the development of diagnostic tools routinely used for detection and evaluation of treatment efficiencies. As the knowledge and use of more molecular-based diagnostic testing evolves in clinical laboratories, the technology will eventually become automated and a standard practice in diagnostic laboratories for clinical evaluation of various disorders. We are already witnessing the use of semi-automated molecular based assays for quantification and processing of complex information by microchip technology to evaluate diseases with underlying genetic causes. These changes, occurring during the last couple of years, are establishing a new standard in patient management. The molecular techniques that are particularly related to diagnostic testing have been providing laboratorians and physicians with highly sensitive, accurate, and reproducible results in a very fast and efficient manner.

These technologies are providing new molecular approaches to assess the genome at all levels, including the ability to determine the presence, or absence of individual genes, their base sequences, and structural alterations in chromosomes. As correlations between genetic anomalies and diseases are established, there will be future application of novel technologies. The development of clinical diagnostic products will provide information critical to the evaluation and management of various disorders, including lymphomas, leukemias, other cancers, as well as other genetic diseases. For example, there is emergence of microchip technology, in which either oligonucleotides or complementary deoriboxynucleic acid (cDNAs) are spotted onto solid supports and subsequently hybridized with genes obtained from patient samples. This powerful technology has the potential to design various chips for disease-specific screening tools e.g., acute leukemia microchip including common fusion genes generated by translocations. It is likely that chip technology will lead to refinements of hematologic neoplasms and help characterize molecular events associated with particular tumors (1, 2).

As molecular biologic techniques become further embedded into clinical laboratories, it is necessary for pathologists to understand both the technology and the clinical applications of these diagnostic modalities. To accomplish this, we aim to provide in this chapter information for commonly used techniques in diagnostic hematopathology including immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangement, detection of translocations that has become a diagnostically important tool in laboratory evaluation of hematopoietic neoplasms, and review common genetic abnormalities seen in various hematologic neoplasms. This section is essentially divided into several parts to review common abnormalities seen in lymphomas, leukemias, and nonneoplastic hematopoietic disorders in which the latter group includes several important recently described molecular abnormalities in diseases related to coagulation and anemias. The genetics of commonly seen anemias such as hemoglobinopathies are covered elsewhere in this book hence are not discussed here.

- HEMATOLOGIC NEOPLASMS
- IMMUNOGLOBULIN AND T-CELL RECEPTOR STRUCTURE AND FUNCTION
- NON-NEOPLASTIC DISORDERS
- GENETIC FACTORS ASSOCIATED WITH COAGULATION DISORDERS

HEMATOLOGIC NEOPLASMS

Part of "11 - Clinical Applications of Molecular Biology Hematopoietic Disorders"

The most important and commonly used tests used in evaluation of lymphoid neoplasms include immunoglobulin and T-cell receptor gene rearrangement by using Southern blot analysis and polymerase chain reaction (PCR) amplication. Molecular gene rearrangement studies have had the biggest diagnostic impact in non-Hodgkin's lymphoma. In particular, these studies have been used in (i) identifying clonal proliferations, especially in peripheral T-cell lymphoma; (ii) confirming the cellular lineage of a lymphoproliferation; (iii) analyzing for recurrent disease; and (iv) evaluating “atypical” hyperplasias for the presence or absence of a clonal population when routine morphology or immunologic methods are equivocal.

IMMUNOGLOBULIN AND T-CELL RECEPTOR STRUCTURE AND FUNCTION

Part of "11 - Clinical Applications of Molecular Biology Hematopoietic Disorders"

Both the immunoglobulin and T-cell receptors are heterodimers composed of two distinct disulfide-linked chains. These receptors are part of the so-called immunoglobulin supergene family. This family of related genes all has a similar molecular structure and encode their genetic material through a process of DNA recombination before a functional protein can be produced. The immunoglobulin genes (heavy chain, kappa, and lambda light chain genes) and the T-cell receptor genes (alpha, beta, gamma, and delta chain genes) each are composed of three or four basic

gene regions: variable (V), diversity (d), joining (J), and constant (C) regions (Fig. 11.1). As will be discussed later, one segment from each of these regions recombines to encode for a particular immunoglobulin or T-cell receptor protein. Both receptor proteins are composed of a constant region and a variable region. The constant region confers a common characteristic structure, and the variable region is the site responsible for unique antigen recognition. The vast diversity of variable region structure is attributable to gene rearrangement, a process that is replayed uniquely in every developing B- and T-lymphocyte. To fully appreciate the diagnostic and therapeutic impact of molecular techniques in the clinical laboratory, one must first understand this gene rearrangement process.

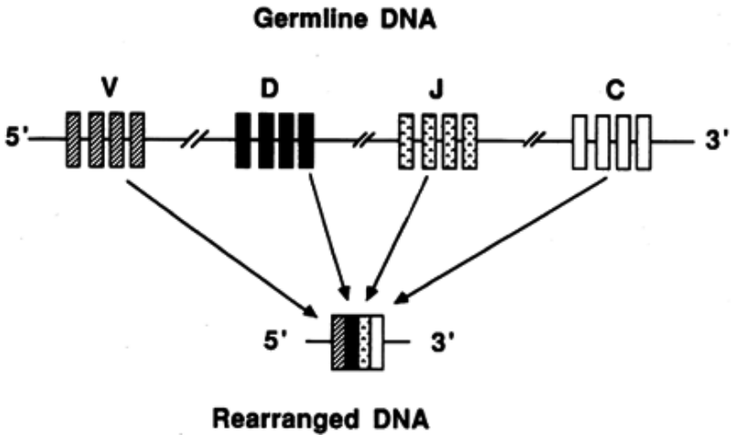


FIGURE 11.1. General structure of the immunoglobulin and T-cell receptor genes. One gene segment from each of the variable (V), diverse (d), joining (J), and constant (C) regions from the germline DNA recombine during the rearrangement process to form a unique gene sequence.

Immunoglobulin Genes

Construction of an intact immunoglobulin molecule is the result of rearrangements involving multiple regions within each of the immunoglobulin heavy and light chain genes (3, 4). Although the immunoglobulin heavy and light chain genes are located at different chromosomal loci, all undergo a similar rearrangement process of their different gene regions. The immunoglobulin heavy chain gene is located on chromosome 14q32, the kappa light chain gene on chromosome 2p12, and the lambda light chain gene on chromosome 22q11. The first rearrangement step in the early B-precursor cell begins with the heavy chain gene locus. The heavy chain gene contains greater than one hundred variable (V) region genes, approximately 30 D region genes, and six J region genes that create more than 18,000 possibilities for generations of immunoglobulin H rearrangement pattern. The nine constant (C) region genes in this locus encode the immunoglobulin constant region and correspond to the particular immunoglobulin isotype: Cμ(IgM), Cδ(IgD), Cγ1-4(IgG1-4), Cα1-2(IgA1-2), or CεIgE). Heavy chain rearrangement initially begins with one D region gene recombining with one J region gene, followed by recombination with a V region gene (Fig. 11.2). The resulting rearranged DNA is transcribed and undergoes ribonucleic acid (RNA) splicing, which involves removing any intervening intron material and splicing the V, D, J, and Cμ exons together (Fig. 11.2). The resulting messenger RNA (mRNA) product can then be translated into an immunoglobulin heavy chain molecule. If, however, the initial D-J or V-D-J rearrangement mistakenly encodes for a termination or nonsense codon, then transcription or translation of the gene becomes “non-productive.” Subsequently, the same recombination process will begin on the second heavy chain gene located on the other chromosome 14. If the second attempt at heavy chain rearrangement is unsuccessful, the B-precursor cell will die.

The next step in the rearrangement process occurs in the late stage of B-cell development is the rearrangement of one of the kappa immunoglobulin light chain genes, located on chromosome 2p12. The kappa light chain gene structure differs slightly from the heavy chain gene: first, the kappa gene lacks any known D regions, and second, there is only one C region gene (Fig. 11.3). If the initial kappa V-J rearrangement is a “productive” rearrangement, the transcription and translation process will lead to the formation of a complete immunoglobulin light chain. This light chain then will combine with the heavy chain to form an intact immunoglobulin molecule. If, however, the first V-J rearrangement is nonproductive, another V-J rearrangement is attempted on the second kappa light chain gene on the other chromosome 2p12. If both kappa rearrangements are nonproductive, the light chain gene rearrangement process begins anew with the lambda light chain gene on chromosome 22q11. Like the kappa gene, the lambda light chain gene structure lacks any known D regions. In contrast to the kappa gene, the lambda gene includes six C regions, each having a high degree of homology to one another (Fig. 11.3). If the light chain gene rearrangement process fails with both lambda genes, immunoglobulin differentiation cannot proceed and the cell is destined to die in the bone marrow. The sequential rearrangement of kappa before lambda accounts for the normal two-to-one ratio of kappa-to-lambda positive B-cells in peripheral blood. After transcription, the rearranged V-D-J or V-J segments undergo RNA splicing, juxtaposing the V-D-J or V-J segments with the C region. The Cμ gene is initially juxtaposed with the rearranged V-D-J sequence, because it is the first 5′ constant region exon. Only Cμ will be spliced into this V-D-J sequence, and the remaining constant regions will be “ignored” during transcription. Therefore, the μ heavy chain constant region, i.e., IgM, will be the initial immunoglobulin expressed by the B-lymphocyte.

Isotype Switching

The process by which B-cells acquire other surface immunoglobulin isotypes is called isotype switching. Although these cells have different heavy chain constant regions, they still maintain the same antigen binding specificity because there has been no change in the variable region of the antibody. One mechanism of isotype switching involves the transcription of a long V-D-J-Cμ-Cδ transcript; this results in the formation of a mixture of V-D-J-Cμ and V-D-J-Cδ mRNA and, consequently IgM and IgD immunoglobulin. Another mechanism of isotype switching involves deletion of intervening C regions between the V-D-J rearrangement and the C region gene of choice.

T-Cell Receptor Genes

The development of the mature T-cell begins as a bone marrow-derived thymocyte precursor, which migrates to the thymus where subsequent maturation occurs. In the thymus, T-cells that express the preferred T-cell receptors with appropriate MHC molecule are selected for further differentiation, and potential “self” reacting T-cells are removed.

All four T-cell receptor genes (alpha, beta, gamma, and delta) undergo a recombination process involving V, D, J, and C regions, similar to the immunoglobulin heavy and light chain genes. However, there are some important differences that require understanding in order to detect T-cell receptor gene rearrangement.

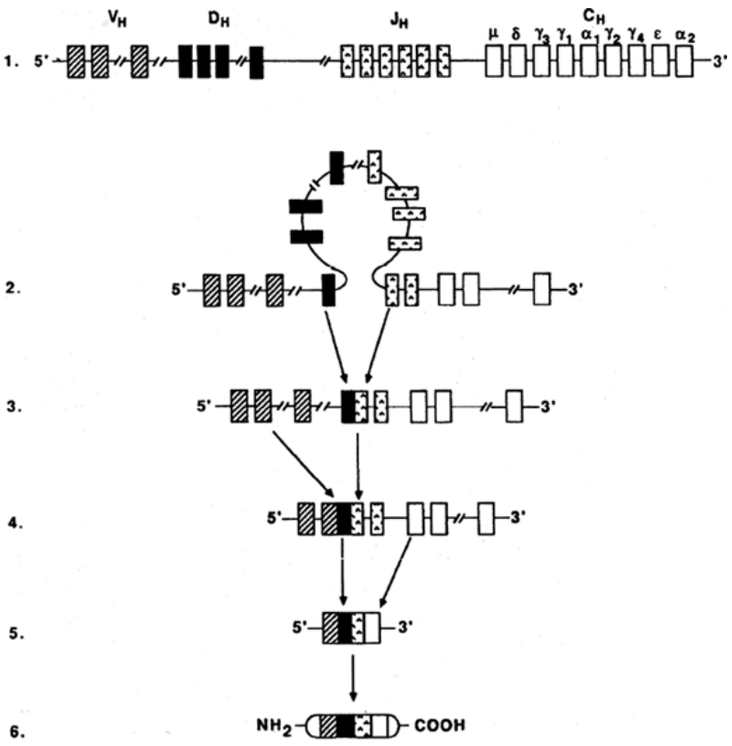


FIGURE 11.2. Rearrangement of the immunoglobulin heavy chain gene. The hatched boxes [variable (V)], solid boxes [diversity (d)], stippled boxes [joining (J)], and open boxes [constant (C)] represent the exons of the respective gene regions. 1) germline configurations; 2) the initial step in D-J rearrangement with “looping out” of intervening DNA, and 3) eventual D-J recombination; 4) V-D-J rearrangement; 5) RNA transcription and splicing step, and 6) immunoglobulin heavy chain production.

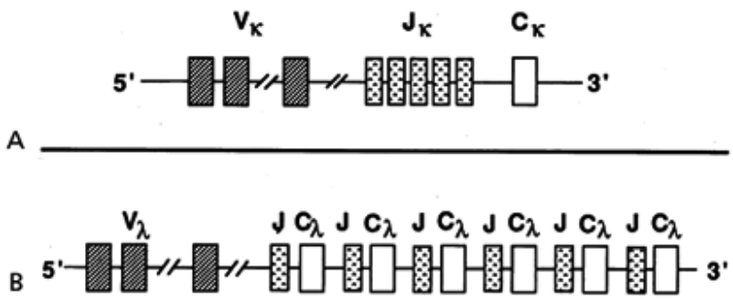


FIGURE 11.3. Structure of the immunoglobulin kappa (**A**) and lambda (**B**) light chain genes. See Figure 11.1 and Figure 11.2 for general explanation of figure characters.

There are at least two types of T-cell receptors expressed by T-lymphocytes. The most common T-cell receptor is composed of alpha (α) and beta (ß) chains, while the second type of T-cell receptor is composed of a gamma (γ)/delta(δ) heterodimer, α/ß positive T-cells predominate within the peripheral blood and lymph nodes, while γ/δ-positive T-cells predominate in epithelial and epidermal locations (5). This differential localization of the T-cell population suggests that each type of T-cell receptor serves a different role in the immune system. Both the α and δ chain genes are located on chromosome 14q11, with the δ chain gene located

entirely within the α gene. The ß chain gene is located on chromosome 7q34, and the γ chain gene is located on chromosome 7p15. The ß and δ genes consist of V, D, J, and C regions, while the α and γ genes are composed of V, J, and C regions. Each of these T-cell receptor genes undergoes recombination in a fashion similar to the immunoglobulin genes, leading to the production of an intact, heterodimeric, T-cell receptor protein (Fig. 11.4).

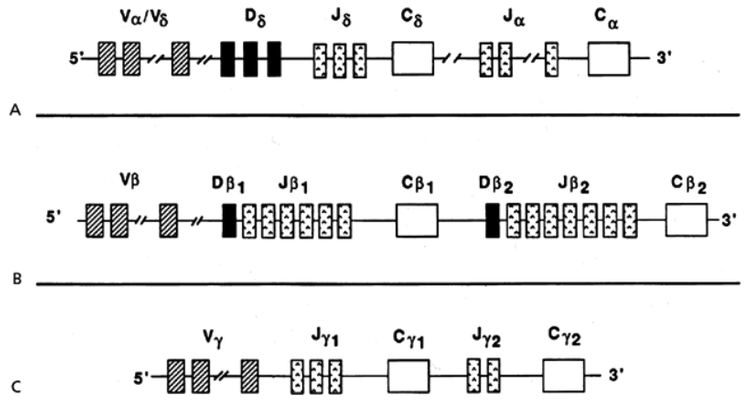


FIGURE 11.4. Structure of the T-cell receptor genes. See Figure 11.1 and Figure 11.3 for explanation of Figure characters. (**A**) shows the structure of the Tα/δ gene on chromosome 14q11. (**B**) shows the germline configuration of the Tß gene. (**C**) shows the Tγ gene structure.

T-α Gene

The α gene consists of approximately 100 V regions and over 100 J regions, which are spread over a 1,000 kb portion of chromosome 14 (Fig. 11.4). The large gene size and the numerous J regions have contributed to the difficulty in studying the α chain gene. The V-J recombination subsequently rearranges with a single α constant region, Cα, to form a complete α transcript.

T-ß Gene

In contrast to the other members of the immunoglobulin supergene family, the ß chain gene consists of two separate loci, $ß_1$ and $ß_2$, each composed of its own D, J, and C regions (Fig. 11.4). The $ß_1$ and $ß_2$ regions recombine only within their own groups, i.e., $Dß_1$-$Jß_1$-$Cß_1$ or $Dß_2$-$Jß_2$-$Cß_2$. The two constant regions, $Cß_1$ and $Cß_2$, are highly homologous, suggesting that the constant region of the ß receptor protein is functionally similar regardless of whether $Cß_1$ or $Cß_2$ is expressed. Rearrangement of Vß within the $ß_2$ locus is associated with deletion of the $Dß_1$-$Jß_1$-$Cß_1$ locus. As with the immunoglobulin family, if the initial V-D-J rearrangement of the ß chain gene is unsuccessful, an attempt will be made to rearrange the ß chain gene on the other chromosomal allele.

T-γ Gene

The γ chain gene was initially discovered in a small subset of circulating T-cells that were CD3 positive, but that lacked CD4, CD8, and the α/ß T-cell receptor. The CD3-associated T-cell receptor from this group of cells was found to have some similarities to the ß chain gene, but offered less diversity, suggesting that it was based on a more primitive recognition system. The γ gene contains 11 V regions and only two J loci, which leads to a limited diversity in this gene subset (Fig. 11.4). Thus, a polyclonal mixture of T-lymphocytes can demonstrate what appear to be oligoclonal rearrangements of the Tγ gene, as only seven or eight possible gene combinations can form. Consequently, gene rearrangement analysis of the Tγ gene must be interpreted with great caution.

T-δ Gene

The δ gene has a unique structure, as it is located entirely within the α chain gene; the V, D, and J regions of the δ locus are situated between the V and J regions of the α gene (Fig. 11.4). Similar to the γ chain gene, only 10 Vδ regions have been identified. The δ gene locus contains only one C region, which is structurally similar to Cα. As should be obvious from the genomic structure of the α/δ locus, a V-Jα rearrangement will effectively delete the entire δ gene. Thus, it would be impossible to express the δ receptor in a cell that has a functional α/ß receptor. Because of their location and similar structure, it has been hypothesized that the α gene evolved as a result of a δ gene duplication.

Clonality

Clonality is an essential concept to understand before interpreting Southern blot hybridization studies. A clonal process is defined as a proliferation of daughter cells originating from a single precursor cell, each having identical phenotypes and growth characteristics as that original precursor cell. Normal B- and T-lymphocytes are polyclonal proliferations, with each lymphocyte having its own unique structure and functional characteristics. Theoretically, each lymphocyte within an individual should have

its own unique molecular structure and, thus, be capable of responding to a different antigenic stimulation.

As stated earlier in the immunoglobulin and T-cell receptor section, there are literally millions of different potential recombinations that can occur among the V, D, and J regions in a gene. In addition to these random associations among V-D-J, N-region diversification and somatic gene mutations can occur that further add to the tremendous diversity of the system. N-region diversification is a terminal deoxynucleotidyl transferase (TdT)-dependent process that results in additions or deletions of nucleotides at the V-J or V-D-J junctions.

Laboratory Analysis of DNA

Commonly used molecular assays used for DNA extraction, restriction endonuclease digestion, and Southern blot transfer and hybridization studies previously were discussed. Two important concepts that must be thoroughly understood before interpreting gene rearrangement studies are: (i) probe structure and location, and (ii) the restriction enzyme map of the gene.

The most commonly used immunoglobulin heavy chain gene probe is complementary to the J_H and Jk region. Thus, when a clonal V-D-J rearrangement has occurred, the J_H probe will bind to the segment: the J_H probe will recognize any or all of the six J_H exons.

The T-cell receptor gene probes work in a fashion similar to their immunoglobulin gene equivalents. Both Jß and Cß probes have been used in the evaluation of the ß chain gene. Because the two Jß regions and the two Cß regions each have a high degree of homology, the Jß and Cß probes will bind to both Jß and both Cß regions, respectively.

Immunoglobulin Gene Rearrangement

A knowledge and understanding of the restriction enzyme map of the desired gene is essential before a Southern blot study can be interpreted. These maps localize various restriction enzyme sites relative to the gene of interest as well as to each other. The whole concept of gene rearrangement refers to the process by which a restriction enzyme site is “rearranged” relative to its previous location, thus changing the size and blot location of a particular hybridization fragment. For example, the immunoglobulin heavy chain gene contains a *Bam*HI restriction enzyme site, located 5′ upstream from the joining region, and another BamHI site, located 3′ downstream from the Cμ constant region. This results in a germline 17 kb fragment after digestion with *Bam*HI. During the process of D-J or V-D-J rearrangement, the recombination into the J region removes the intervening sequences 5′ to the J region area, thus eliminating the initial 5′ BamHI restriction site (Fig. 11.5). Consequently, the germline 17 kb fragment seen with *Bam*HI-cut DNA then would be rearranged to either a larger or smaller fragment, depending on where the next, random 5′ *Bam*HI site is located. Because of its different size, the rearranged fragment will migrate to a different location in the blot and is thus referred to as a rearranged immunoglobulin heavy chain gene. Southern blot analysis of clonal rearrangements of either the immunoglobulin or T-cell receptor genes is capable of detecting a clonal proliferation, if the specimen being studied contains five percent monoclonal cells.

PCR Detection of Monoclonality

As discussed earlier, diversity of the antigen receptor gene is accomplished by random selection of one of the V, D, and J region genes. Because of this random selection of the V region genes, PCR methods for amplification of immunoglobulin heavy chain genes is less sensitive in determination of clonality compared to Southern blotting because it is technically very difficult to generate primers to bind efficiently to all of the V region sequences. In contrast, PCR is more sensitive for detection of monoclonal population if primer used for the V consensus region amplification anneals successfully during the PCR amplification. To be able to increase the sensitivity by PCR amplification, multiple V region primers are designed to bind to the consensus region of immunoglobulin genes, which are referred to as framework (FR) regions. The most common targets include FR3 and FR1. The use of consensus primers allows greater sensitivity for the priming of the many V region genes. In polyclonal proliferation, PCR amplification of Ig genes using consensus sequence primers will produce no prominent band because all of the lymphocytes amplified by PCR represent different length of segments of DNA amplified using V (FR3 or FR1) and J_H primers. In the presence of a monoclonal population, however, these will be a prominent single band visualized by electrophoresis and ethidium bromide staining. There is a significant difference in evaluation of monoclonality between the agarose gel and polyacrylamide gel electrophoresis (PAGE) analysis. PAGE allows a much better separation of the monoclonal proliferation and is recommended for monoclonality assessment. Although gene rearrangement studies are quite useful for assessment of lymphoma or leukemia, there are important considerations in the utility of this assay. Gene rearrangements can be seen in acute myelogenous leukemia and myelodysplastic disorder; hence it should not be used for differential diagnosis between myeloid and lymphoid leukemic process. Also, some of the high-grade lymphoid tumors have promiscuous gene rearrangement (6, 7). Another important aspect of the PCR analysis is the consideration of clonal evolution that describes development of a separate clone in a previously established tumoral clone and giving rise to different rearrangement of the antigen receptor genes (8). This is believed to be secondary to continuous active state of recombinase, which is giving rise to new rearrangements.

T-Cell Receptor Gene Rearrangement

In the T-cell receptor gene system, a similar rearrangement process occurs. The most commonly used gene probe for T-cell receptor gene rearrangement studies are those complementary to the T-cell receptor ß chain gene. In this gene, there is a *Bam*HI restriction enzyme site 5′ of the $Jß_1$ region and 3′ of $Cß_2$: this *Bam*HI fragment, therefore, surrounds both Jß-Cß loci. Similar to that described with the immunoglobulin heavy chain gene, the 5′ upstream *Bam*H1 enzyme site is eliminated during the process of either a D-J or V-D-J rearrangement. Consequently, the germline fragment seen with *Bam*HI-cut DNA would be rearranged

to either a larger or smaller fragment. Although this enzyme would appear ideal, the large size of the germline band (23 kb) seen with *Bam*HI may be difficult to distinguish from rearranged fragments. Because of this interpretative difficulty, it is essential that multiple restriction enzymes be used when evaluating the T-ß chain gene. Most laboratories use *Bam*HI, *Eco*RI, and *Hind*III restriction enzymes when evaluating for T-cell receptor gene rearrangements.

The *Eco*RI restriction enzyme sites for the T-cell receptor ß chain gene are located 5′ of the JB_1 region, 3′ of the CB_1 region, and both 5′ and 3′ of CB_2 (Fig. 11.6A); this will yield a germline configuration of 11 kb and 4 kb bands when probed with a Tß constant region probe (Fig. 11.6B). When either a DB_1-JB_1 or Vß-DB_1-JB_1 rearrangement has occurred, the most 5ß *Eco*RI restriction site is removed, thus changing the size of the germline 11 kb band (Fig. 11.6C). The 4-kb band is not changed, as the *Eco*RI sites around (two are not disrupted following either DB_1-JB_1 or V-DB_1-JB_1 rearrangement).

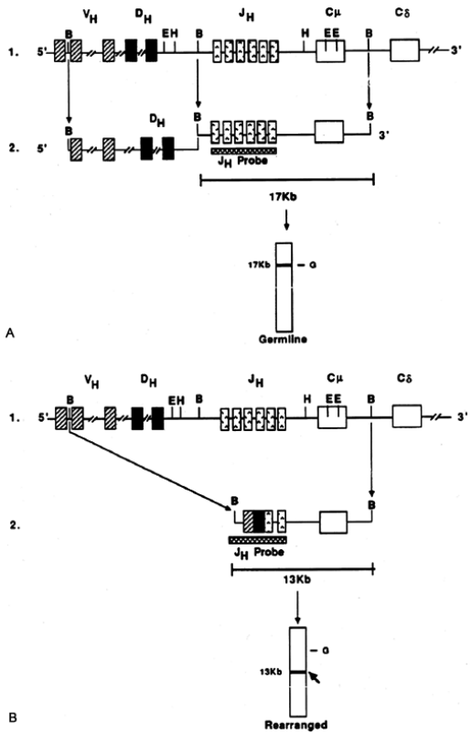


FIGURE 11.5. Restriction enzyme map of the immunoglobulin heavy chain gene. The thin vertical lines represent recognition sites for *Bam*HI (*B*), EcoRI (*E*), or *Hind*III (H). See Figure 11.1 and Figure 11.3 for an explanation of other figure characters. (**A**) The heavily stippled box represents the J_H probe. When cut with *Bam*HI, the unrearranged immunoglobulin heavy chain gene leads to a germline 17 kb band. (**B**) Following V-D-J rearrangement, the original 5′ *Bam*HI site disappears and is replaced by the next random *Bam*HI site, leading in this example to a rearranged 13 kb fragment as detected by the J_H gene probe. Other rearrangements might yield *Bam*HI fragments of smaller or larger size.

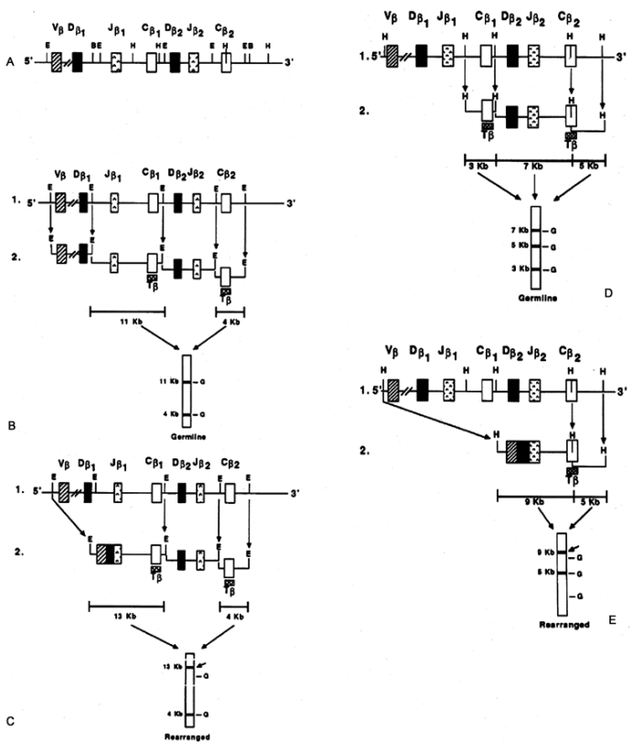


FIGURE 11.6. Tß chain rearrangement analysis (**A**) and by EcoRI (**B, C**) and HindIII (**D, E**) are shown. Restriction enzyme map of the Tß chain gene. See Figure 11.1 and Figure 11.4 for explanation of figure characters. (**B**) The heavy stippled box represents the Tß constant region probe. The germline configuration shows two bands following *Eco*RI (E) enzymatic digestion, as the probe will bind to both Cß regions. (**C**) Following Vß-DB_1-JB_1 rearrangement, the 5′ *Eco*RI site disappears and is replaced by the next random *Eco*RI site. This leads to a rearranged 13 kb fragment (arrow). The 4 kb fragment of CB_2 has not been affected by this rearrangement process. (**D**) The germline configuration shows three bands following *Hind*III (H) enzymatic digestion, as the probe will bind to both Cß regions. (**E**) Following the Vß-DB_2-JB_2 rearrangement, the HindIII site 3′ of CB_1 is deleted and is replaced by the next random *Hind*III site. This leads to a rearranged fragment of 9 kb size (arrow) and deletion of the 3-kb band. The 5-kb fragment of CB_2 has not been affected by this rearrangement process.

The limitation of using *Eco*RI for T-cell receptor ß chain analysis is its inability to detect rearrangements involving the ß2 locus. When a DB_2-JB_2 or Vß-DB_2-JB_2 rearrangement has occurred, the ß1 locus and, thus, the *Eco*RI restriction sites around the ß1 locus, is deleted, including the germline 11 kb band. However, this D-J or V-D-J rearrangement into the ß2 locus will not disrupt either of the *Eco*RI restriction sites surrounding the CB_2 region. Thus, with a ß2 rearrangement, the 4-kb band is not changed and one is left with an allelic deletion of the 11 kb band. This deletion will not be evident with rearrangement studies if the other allele remains in germline configuration or if there are significant numbers of non-T-lymphocytes in the specimen being analyzed, which could contribute to the germline band.

To overcome this limitation, *Hind*III usually is used as an additional restriction enzyme for T-cell receptor ß chain gene analysis (Fig. 11.6D). The *Hind*III restriction enzyme sites are located immediately 5′ and 3′ of Cß1, within the CB_2 segment, and 3′ of CB_2; this will yield germline bands of 3 kb, 5 kb, and

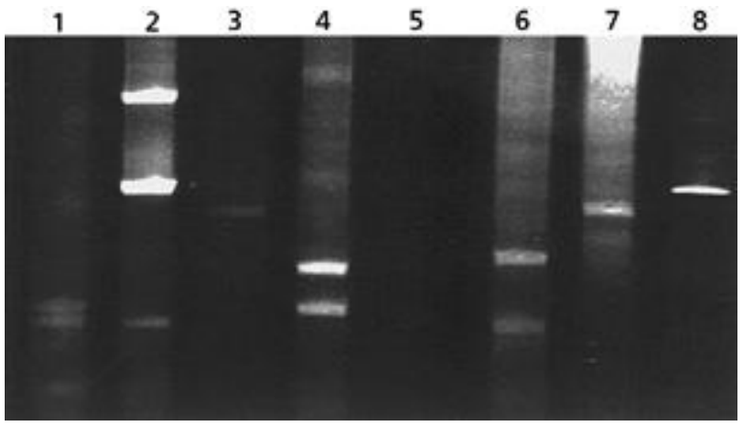


FIGURE 11.7. Temporal temperature gradient electrophoresis (TTGE) analysis of various lymphoid disorders. DNA is polymerase chain reaction (PCR) amplified by GC-clamped primers. Small aliquots of amplicons then were separated by a denaturing polyacrylamide gel and gradients of increasing temperature. Distinct bands detected under ultraviolet light demonstrate evidence for monoclonality. Sample in lane 5 is polyclonal while others are example of T-cell neoplasm.

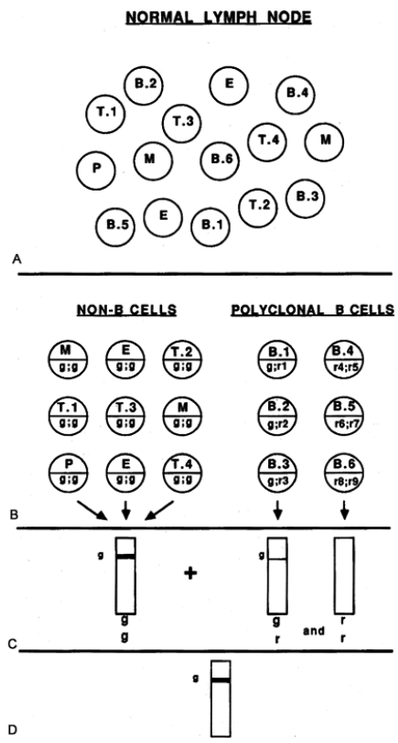


FIGURE 11.8. Schematic drawing of a normal lymph node and resulting immunoglobulin gene rearrangement studies. (A) A normal lymph node consists of a mixture of polyclonal B-cells (B.1-B.6), polyclonal T-cells (T.1-T.4), macrophage/histiocytes (M), polymorphonuclear granulocytes (P), and endothelial cells (E). (B) Schematic depiction of gene rearrangement analysis applied to cells from a normal lymph node. The non-B-lymphocytes are shown on the left-hand side and consist of T-cells, macrophages, neutrophils, and endothelial cells. These cells retain their immunoglobulin genes in germline configuration. The polyclonal B-cells, depicted on the right-hand side, consist of a mixture of B-cells (B.1-B.3) that retain one immunoglobulin gene allele in germline configuration while rearranging the other allele in a unique recombination (r1-r3). The remaining B-cells (B.4-B.6) have attempted to rearrange both immunoglobulin alleles (r4-r9), with only one allele per cell exhibiting a "productive rearrangement." (C) This schematic of a Southern blot audioradiograph shows that the non-B-cells contribute to the formation of one germline band because both alleles remain in germline configuration. Some polyclonal B-cells have one germline allele and one uniquely rearranged allele; the germline alleles will contribute to the formation of a germline band on Southern blotting, while the single-copy rearranged alleles are undetectable by this method. The remaining B-cells with two rearranged alleles also exhibit no detectable bands, as each rearrangement again is a unique recombination. (D) Final Southern blot resulting from the combination of the non-B-cell component and polyclonal B-cells, therefore, leads to a single germline configuration with no rearranged bands identified.

7 kb when probed with a TB constant region probe. If a rearrangement into the B1 locus has occurred, none of the *Hind*III restriction enzyme sites are affected, thus giving a germline configuration on Southern blot. However, rearrangement into the B2 locus will remove the *Hind*III site between CB_1 and DB_2, thus changing the size of the germline 7 kb band (Fig. 11.6E). The 3 kb band will be deleted along with the B1 locus; the 5 kb band will be unchanged as the *Hind*III sites surrounding CB_2 are unaffected by the rearrangement process.

Thus, for T-cell receptor B chain gene analysis, *Bam*HI restriction enzyme analysis will show rearrangements involving either the B1 or B2 locus. However, the large germline band seen with *Bam*HI presents its own unique problem. Analysis with *Eco*RI restriction enzyme will demonstrate rearrangements of the B1 locus, but will not show a rearranged band if the B2 locus is involved. This is in contrast to *Hind*III restriction enzyme analysis, which will demonstrate B2 loci rearrangements, but fail to demonstrate B1 rearrangements.

Although in B-cell neoplasms, clonality can be determined using surface markers such as kappa and lambda light chains, the T-cell clonality cannot be detected by these methods. Apparent loss of T-cell antigen expression is considered to be good evidence for T-cell malignancy; however, some cases may have this feature without presence of malignant process. Hence, the molecular biologic studies for the determination of T-cell clonality are essential for lesions suspicious for T-cell malignancy.

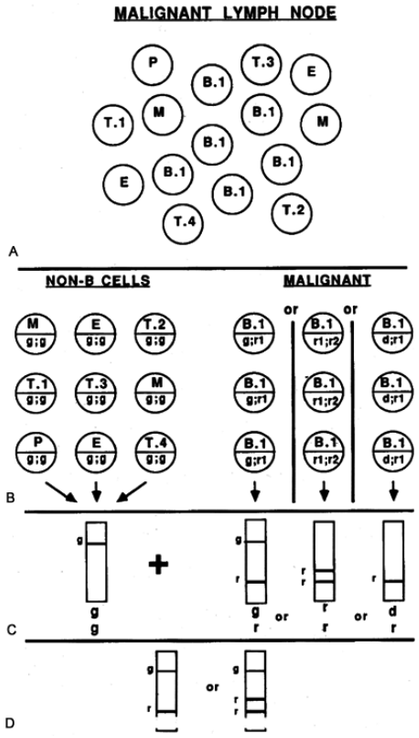


FIGURE 11.9. Schematic drawing of a malignant lymph node and resulting immunoglobulin gene rearrangement studies. (**A**) A malignant lymph node consists of a mixture of monoclonal B-cells (B.1), polyclonal T-cells (T.1–T.4), macrophage/histiocytes (M), polymorphonuclear granulocytes (P), and endothelial cells (E). (**B**) Schematic depiction of gene rearrangement analysis applied to cells from a neoplastic lymph node. The nonmalignant cells are shown on the left-hand side; these cells retain their immunoglobulin genes in germline configuration. The monoclonal B-cells depicted on the right-hand side consist of monoclonal B-cells that may be in one of three possible configurations: (i) the cells may have one rearranged immunoglobulin allele and the other allele in germline configuration; (ii) the cells may have both immunoglobulin alleles rearranged, with only one productive rearrangement; or, (iii) one allele may be deleted and the other immunoglobulin locus rearranged. (**C**) This schematic of a Southern blot audioradiograph shows that the few reactive cells contribute to the presence of a germline band. The monoclonal B-cells will have one of three possible configurations depending on the status of both immunoglobulin alleles. (**D**) Final Southern blot resulting from combination of the reactive cells and monoclonal B-cells lead to a faint germline band, contributed from the few background normal cells and possibly a germline allele in the monoclonal B-cells. The dominant band, however, appears rearranged.

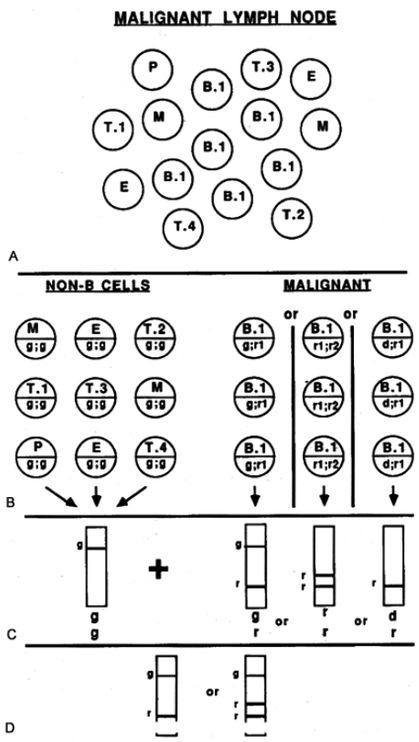


FIGURE 11.9. Schematic drawing of a malignant lymph node and resulting immunoglobulin gene rearrangement studies. (**A**) A malignant lymph node consists of a mixture of monoclonal B-cells (B.1), polyclonal T-cells (T.1–T.4), macrophage/histiocytes (M), polymorphonuclear granulocytes (P), and endothelial cells (E). (**B**) Schematic depiction of gene rearrangement analysis applied to cells from a neoplastic lymph node. The nonmalignant cells are shown on the left-hand side; these cells retain their immunoglobulin genes in germline configuration. The monoclonal B-cells depicted on the right-hand side consist of monoclonal B-cells that may be in one of three possible configurations: (i) the cells may have one rearranged immunoglobulin allele and the other allele in germline configuration; (ii) the cells may have both immunoglobulin alleles rearranged, with only one productive rearrangement; or, (iii) one allele may be deleted and the other immunoglobulin locus rearranged. (**C**) This schematic of a Southern blot audioradiograph shows that the few reactive cells contribute to the presence of a germline band. The monoclonal B-cells will have one of three possible configurations depending on the status of both immunoglobulin alleles. (**D**) Final Southern blot resulting from combination of the reactive cells and monoclonal B-cells lead to a faint germline band, contributed from the few background normal cells and possibly a germline allele in the monoclonal B-cells. The dominant band, however, appears rearranged.

Most recent amplification studies using PCR primers with GC-clamps and subsequent analysis by denaturing density gradient electrophoresis (DDGE) or temporal temperature gradient electrophoresis (TTGE) has significantly improved the TCR analysis of the TCR genes (9, 10). Assay based on the detection of PCR products of TCR genes by DDGE requires preparation of a gradient gel. Detection of a predominant monoclonal DNA fragment by TTGE relies upon the same principle in which denaturing gradient electrophoresis is based. Both TTGE and DDGE allow the separation of DNA fragments based not only on their length, but also the sequence of the fragments (Fig. 11.7). Using TTGE, the PCR fragments are electrophoresed through a denaturing gel containing urea as the denaturant. During TTGE, the temperature of the gel is increased gradually

in contrast to the DDGE in which the gel is prepared by a gradient of denaturant. Change in temperature results in a linear temperature gradient over the length of the electrophoresis run.

Both in DDGE and TTGE, the detection of similar size but different DNA fragments is based on the melting temperature of the target fragments being analyzed. In a denaturing condition of either DDGE or TTGE, double-stranded DNA fragments amplified by PCR is subjected to conditions that allow melting the fragments in discreet segments referred to as melting domains. The melting temperature (Tm) of these domains is sequence specific. Therefore, the DNA fragment partially separates (melts) when the Tm of the lowest melting domain of the fragment is reached. This partial melting causes branching in the DNA fragments, hence reduces the mobility of the fragments in a polyacrylamide gel.

Figure 11.8 and Figure 11.9 demonstrate a schematic for understanding the rearrangement process in a normal and malignant lymph node, respectively. Figure 11.10 demonstrates example of Southern blot detection of T-cell receptor gene rearrangement in a patient with peripheral T-cell lymphoma.

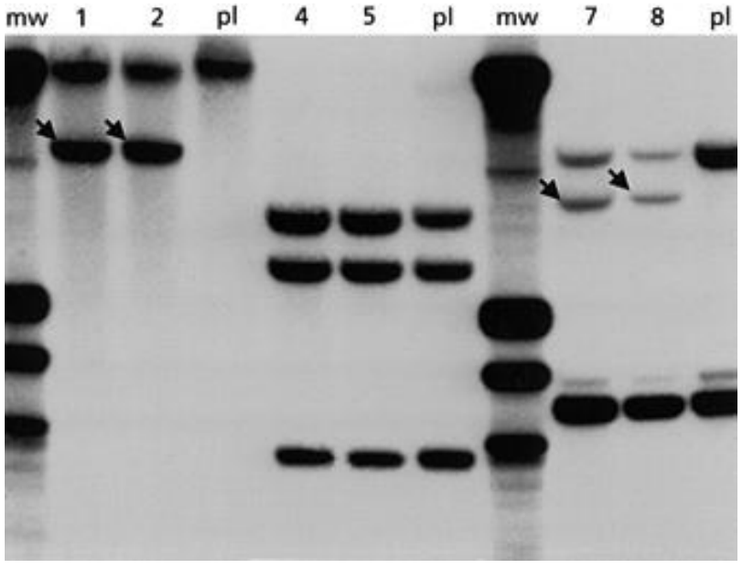


FIGURE 11.10. Southern blot detection of T-cell receptor gene rearrangement in a patient with peripheral T-cell lymphoma. Two separate lesions were investigated for evidence of a monoclonal T-cell population. Restriction digestion in the *Bam*HI (1, 2) and EcoRI (7, 8) show extra bands (arrow), which are not detected in the lane with DNA isolated from placenta (pl). Although *Hind*III digest (4, 5) shows no extra band, two rearranged bands detected by *Bam*HI and *Eco*RI are sufficient to establish clonality. Because both lesions show similar banding pattern, these lesions are considered to be clonally identical.

Non-Hodgkin's Lymphoma

The non-Hodgkin's lymphomas (NHLs) are a diverse collection of lymphoid malignancies with varied etiology, pathogenesis, and clinical course. The diversity of the NHLs is illustrated by recent classifications based primarily on morphology, immunophenotype, and chromosomal abnormalities. Significance of chromosomal abnormalities has been shown by these classifications that have taken into consideration molecular aberrations associated with specific morphologic entities (11, 12). With the currently available molecular assays, genetic abnormalities of various subtypes of lymphomas can be identified in a significant number of lesions (Table 11.1).

TABLE 11.1. COMMON RECURRING CHROMOSOMAL TRANSLOCATIONS IN NON-HODGKIN LYMPHOMA

Disease	Translocation	Oncogene Involved
Small lymphocytic lymphoma	t(14;19)(q32;q13.1)	BLC-3
Mantle cell	t(11;14)(q13;q32)	CCNDI (Cyclin D1)
Lymphoplasmacytic	t(9;14)(p13;q32)	PAX5
Marginal zone	t(11;18)(q21;q21)	Unknown
Follicle center cell	t(14;18)q32;q31)	BCL-2
Burkitt	t(8;14)(q24;q32)	C-MYC
	t(2;8)(q12;q24)	C-MYC
	t(8;22)(q24;q11)	C-MYC
Diffuse large cell	t(3;-)(q27;-)	BCL-6
Anaplastic large cell	t(2;5)(p23;q35)	NPM-ALK

The realization of association between specific chromosomal alterations, particularly chromosomal translocations, had lead to identification of dysregulated genes because of specific genetic abnormalities during the last decade. In this section, these common chromosomal and molecular abnormalities associated with NHLs are described by grouping under specific histopathologic subtypes.

Many chromosomal translocations in B-cell NHL result in juxtaposition of cellular oncogenes to the regulatory portion of immunoglobulin genes or equivalent loci. In general, coding regions of the oncogenes are not interrupted by the translocation. Therefore, the overall consequences of the translocation appear to be deregulated expression of the gene product; the gene may be expressed at inappropriate levels compared with normal lymphocytes of the equivalent stage of B-cell differentiation and/or at inappropriate times when unaffected genes usually are switched off. However, in some lymphomas, gene products involved in translocations also have been found normally in lymphoid tissues at considerable levels. Attempts have failed to demonstrate clear differences at the level of expression of a particular gene between lymphoma tissues carrying a translocation involving the relevant gene and those lacking the gene rearrangement (13). For instance, Bcl-6 protein is expressed in germinal center B cells as well as in a variety of germinal-center-derived lymphoma tissues independently of BCL6 gene rearrangement (14). This suggests that the loss of regulation also is an important factor in lymphomagenesis.

Chronic Lymphocytic Leukemia

The most frequent cytogenetic aberrations in chronic lymphocytic leukemia (CLL) are trisomy 12 and structural abnormalities involving the long arm of chromosome 13 (13q12-q32) on which retinoblastoma gene is located. Their incidence ranges 12% to 32% for trisomy 12 and 10% to 28% for 13q abnormalities. Trisomy 12 is more commonly associated with atypical morphology, higher proliferative activity, advanced disease, and short survival; however, this is not confirmed in a recent study using both karyotyping and fluorescent *in situ* hybridization (FISH) (15, 16 and 17). On the other hand, 13q deletion has been associated with typical morphology, early stage, and longer survival (18). These abnormalities have been considered as presence of single independent clones for specific subtypes of CLL. Molecular studies have demonstrated that monoallelic loss of RB1 at 13q14 band is found in about 30% of cases, but a higher incidence of deletions affecting the interval between D13S319 and the D13S25 loci also has been described (18). FISH analysis has become a more reliable assay for evaluation of these chromosomal aberrations. In fact, a recent study has demonstrated trisomy 12 in hematopoietic stem cells suggesting that the precursor cells in some of CLL patients is much earlier in development than previously considered (19, 20). It is likely that molecular-based techniques will find a higher incidence of and perhaps novel genetic abnormalities in CLL due to its low proliferation rate.

Lymphoplasmacytoid Lymphoma

Close association of t(9;14)(p13;q32) with small lymphocytic lymphoma with plasmacytoid differentiation was demonstrated in a large series of karyotypically analyzed NHL (21). Initial molecular analysis of t(9;14) suggested the PAX5 gene as the responsible gene by two independent groups (22, 23). The t(9;14) translocation involves the PAX (paired homeobox) transcription factor family which has been shown to be important for the embryonal development and organogenesis (24). PAX 5 gene shows restrictive expression in tissues mainly showing expression in the fetal brain and liver during embryogenesis, but only in the lymphocytes and testes following birth. A number of other reciprocal chromosomal fusion partners with the 9p13 breakpoint also have been described including 1q25, 3q27, 7q11, 12q13, 12q21, 10p13, and 9q13. The PAX5 gene encodes the BSAP (B-cell-specific activator protein), which is a key regulator of B-cell development and differentiation (25). Functional loss of PAX 5 causes maturation arrest in the lymphoid cells at the pro B-cell stage. In contrast, overexpression of the gene results in the proliferation of the splenic B-lymphocytes.

A high prevalence of chronic hepatitis C virus (HCV) infection recently has been shown in NHLs, most of which belong to the lymphoplasmacytoid lymphoma/immunocytoma subtype (26). The HCV also is considered to be a major etiologic factor of type II mixed cryoglobulinemia (MC), which is characterized by an underlying proliferation of monoclonal B cells (27). Type II MC frequently evolves into overt B-cell NHL, suggesting that chronic HCV infection can lead to both a benign and a malignant lymphoproliferative disorder. The pathogenesis of HCV and lymphomagenesis currently is unknown.

Mantle Cell Lymphoma

Mantle-cell lymphomas (MCL) are associated with a characteristic chromosomal translocation, t(11;14)(q13;q32). This translocation involves rearrangement of the bcl-1 proto-oncogene from chromosome 11 to the immunoglobulin heavy chain gene on chromosome 14, resulting in an overexpression of cyclin D1 mRNA (also known as bcl-1 and PRAD1) (28, 29). Cyclin D1 overexpression is considered to be directly involved with the lymphomagenesis. D-type cyclins physically associate with cyclin-dependent kinases (CDK4 and CDK6) and facilitate lymphoid proliferation. Upon cyclin-D1 activation, the retinoblastoma (RB) tumor suppressor protein (regulator of E2F transcription factor) loses its interaction with transcription factors leading to upregulation of E2F and promotion of S-phase progression and shortened G1 phase.

The bcl-1 locus is rearrangement detected in 50% to 80% of MCLs by using Southern blot or 75% to 80% by cytogenetics studies. This can be further increased using FISH. Detection of overexpression of cyclin D1 by immunohistochemical studies and molecular detection of bcl-1 translocation are very useful for the diagnosis of mantle cell lymphomas or confirmation of histologic suspicion. The breakpoints are widely scattered on chromosome 11q13; however, the majority of the breakpoints are localized to a 1-kb DNA segment known as the major translocation cluster (MTC) (30). Within the MTC, the breakpoints occur in a relatively small region on chromosome 11 and the 5′ area of one of the IgH joining (JH) regions on chromosome 14. PCR can be utilized to amplify *bcl-1* translocation; however, some of the breakpoints falling outside the MTC cannot be identified with PCR (31). PCR technique has been demonstrated to be sensitive when the MTC is present, however, because of the lack of involvement of MTC, the technique has a low sensitivity for documentation of bcl-1 translocation. It is estimated that only in 33% to 50% of patients with MCL, the breakpoints could be detected by PCR using primers in the region of the MTC. Although differential detection of cyclin D1 by quantitative RT-PCR has been reported as a diagnostic tool, this has not become a practical assay because it requires sampling of whole tissue, which also contains nonneoplastic both reactive and stromal cells. FISH-based assays have found a higher rate of *bcl-1* translocation than PCR assays, likely because of the broader-based probes that can be used (32, 33).

Follicular Lymphoma

The *bcl*-2 translocation is present in 80% to 85% of follicular lymphomas (FL) and up to 30% of diffuse B-cell lymphomas. Approximately 65% of *bcl-2*rearrangements in FL occur at the major breakpoint region (*MBR*) located in the untranslated 3′ end of the last exon and in approximately 10% to 20% of cases, the rearrangement occurs in the minor cluster region (*mcr*) located approximately 30 kb downstream of the *bcl-2*gene (Fig. 11.11) (34, 35 and 36). In a few cases, rearrangements occur at other sites such as the variant cluster region at the 5′ end of the *bcl-2* gene. Tight clustering of the breakpoints at *MBR* and *mcr*, as well as the availability of consensus sequences of the *JH* segments of the *IgH* make this a practical target for PCR amplification. In fact, it is possible to use two universal primers (one for each breakpoint) along with a primer derived from JH region to amplify the majority of the translocations at *MBR* and *mcr*, with minority of translocations failing to be amplified. The biological and clinical significance of differences in breakpoint regions in FL is unclear. However, a recent study suggested that FL patients with *MBR* or *mcr* type of translocation were less likely to show complete remission compared to patients with germline pattern (37). Expression of *Bcl-2* protein can be seen without translocation, as evidenced by its expression in a number of normal tissues, as well as a number of other lymphoproliferative disorders without a t(14;18) (13). A high level of *Bcl-2* protein confers a survival advantage on B cells by inhibiting apoptosis and more generally may block a common cell death pathway induced by chemotherapy, conferring clinical drug resistance on cells over-expressing *Bcl-2* protein. One of the hypotheses for the gene rearrangement is the presence of sequences near the *bcl-2* gene, which resembles the signal sequences recognized by the recombinase enzyme system involving excision of DNA during antigen receptor gene rearrangement. These sequences are 3′ of the *bcl-2* gene such that the gene itself is translocated as a whole functional unit in close proximity to the immunoglobulin enhancer gene regulatory region. This is believed to cause overexpression of *bcl-2* gene. Overexpression of *bcl-2* does not allow differentiation of FCL from other indolent lymphomas because they also

commonly demonstrate expression of this protein. Hence, molecular detection of *bcl-2* translocation may be required for accurate classification of some lymphomas. Fresh or frozen unfixed tissue usually is required for molecular assessment of *bcl-2* translocation with Southern blot analysis in lymphoid tissues, but detection of *bcl-2* translocation with PCR can be performed with formalin-fixed tissue embedded in paraffin (Fig. 11.11c).

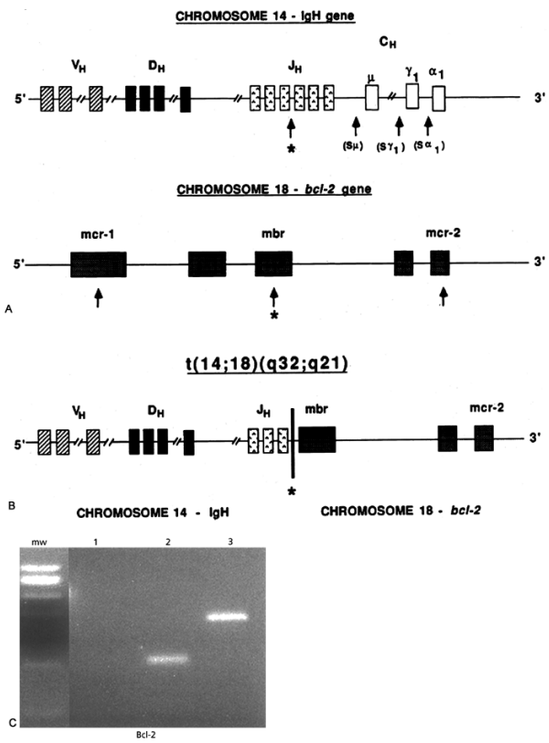


FIGURE 11.11. Molecular gene structure of chromosomes 14 and 18. See Figure 11.1, Figure 11.3, and Figure 11.17 for explanation of figure characters. (**A**) Normal gene structure of chromosomes 14 and 18. Arrows indicate locations of the most common breakpoint sites. *mcr* stands for minor cluster regions and *mbr* represents the major breakpoint region of the bcl-2 gene. *(continued)*

Initial molecular characterization of *bcl-2* significantly contributed to our understanding of apoptosis and helped in identification of other molecules involved in apoptosis regulation. *Bcl-2*, similar to *bcl-XL*, functions as inhibitor of apoptosis by protein-protein interactions with a number of homologues apoptosis regulatory proteins including death promoting protein bax. Bax heterodimerizes with *bcl-2* in homeostatic state while homodimerization (bax-bax) causes apoptosis; in contrast, overexpression of *bcl-2* favors *bcl-2-bcl-2* homodimerization, which prevents apoptosis. Hence, the ratio of *bcl-2*, as well as *bcl-XL* and bax, is very critical in determining the fate of cell. There are also additional factors such as bad and bag-1, which also influence the function of *bcl-2*. Although *bcl-2* overexpression alone is insufficient to transform B-lymphoid cells, the survival advantage provides an opportunity for genetic mutations, which will lead into lymphoma development. This was shown in the case of c-myc in which *bcl-2* was shown to complement the c-myc protein, in part by inhibiting c-myc induced apoptosis.

In addition to diagnostic usefulness of *bcl-2* translocation detection, the assessment of *bcl-2* translocation has been used for the evaluation of minimal residual disease with a sensitivity of 10^{-5} or greater. There is a good correlation between bone marrow positivity for the *bcl-2* translocation by PCR and relapse following autologous bone marrow transplantation (38, 39). However, it should be remembered that *bcl-2*-rearranged cells were documented in the normal healthy donors (40, 41).

Burkitt's Lymphoma

Burkitt's lymphoma (BL) shows two epidemiologic distribution patterns, endemic and sporadic forms. The endemic form of BL typically affects children between 5 and 10 years of age and usually presents with involvement of jaw, abdominal, and extranodal sites. In the endemic type of BLs, clonal integration of Epstein-Barr virus (EBV) is shown in 95% of the tumors, whereas EBV is seen in 15% to 20% of nonendemic sporadic form of BLs (42). The sporadic form occurs in adolescent and young adult age patients, mostly presenting with abdominal mass. Presence of EBV genome in BLs has been implicated for the neoplastic transformation.

On the basis of sequence divergence in the EBNA-2 gene, two different families of EBV, type A and type B, have been defined. Type A EBV immortalizes B cells more efficiently *in vitro* and infects immunocompetent individuals more commonly than type B EBV. However, type B EBV is seen at increasing rates in immunocompromised hosts and in lymphoid neoplasms in this setting. There is a higher level of expression of EBV lytic cycle transcripts (BZLF1, BHRF1, BLLF1, and Fp) in endemic BL tumors similar to the strain association of AIDS-related EBV-positive lymphomas. There appears to be a nearly equal mix of type-A and type-B virus in endemic BL in contrast to Hodgkin's disease (HD) and posttransplant associated lymphoproliferative disorders (PTLDs) where type-A virus predominates (43). Whether the virus type (A or B) detected in BL simply reflects the frequency

of infection in the affected population, or there is a specific virus-type tumor association in some instances remains to be determined. LMP1 is a transforming protein that interacts with tumor necrosis factor receptor (TNFR)-associated factors (TRAFs), and is involved in NFκB activation (44). LMP1 is one of the most significant gene implicated in lymphomagenesis. It resides at the carboxyl terminus (amino acids 346–355), which is a short 30-base pair segment considered to be important to the transforming ability of LMP1 protein. This segment is deleted in a variety of EBV-associated tumors and may be related to tumor pathogenesis or disease progress (45, 46). The deletion also is noted in some cases of endemic BL.

The translocation t(8;14)(q24;q32) is the characteristic chromosomal aberration of BLs. The t(8;14) juxtaposes the c-myc gene in 8q24 next to the IgH locus in 14q32, resulting in overexpression of c-Myc (Fig. 11.12). Although often used as a “characteristic” marker for BL, this translocation is not specific for BL and can be found in other large-cell lymphoma.

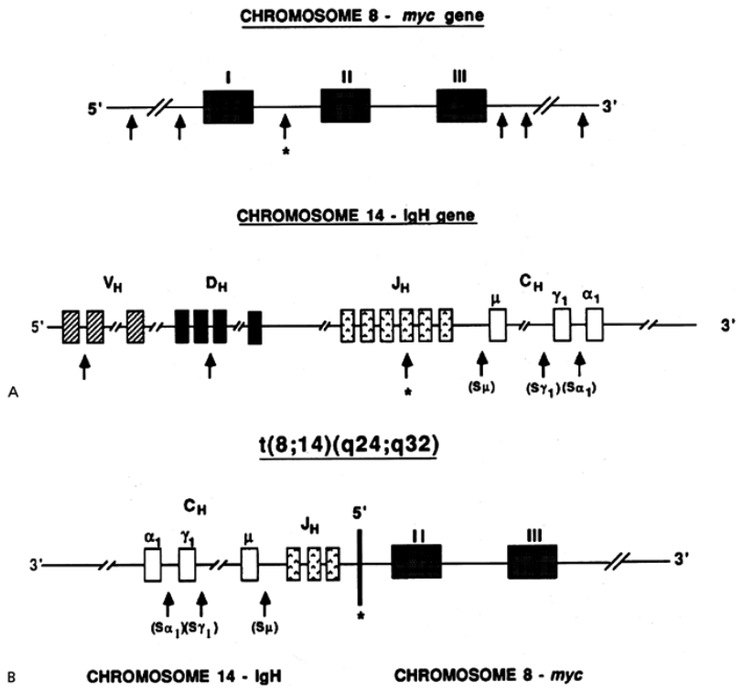


FIGURE 11.12. Molecular gene structure of chromosomes 8 and 14. (**A**) See Figure 11.1 and Figure 11.3 for explanation of figure characters. The arrows indicate preferential locations of breakpoint sites in the t(8;14) translocations. The Sμ, Sγ, and Sα represent the respective switch regions for particular constant region exons. (**B**) This is one example of a reciprocal translocation between the *myc* oncogene and the immunoglobulin heavy chain gene. The breakpoint on chromosome 8q24 has occurred between c-myc exons I and II (*). The breakpoint on chromosome 14q32 has occurred within the middle J_H region and has joined in a 5′ to 5′ fashion with the *c-myc* oncogene.

The great majority of the endemic BLs shows breakpoints far upstream of c-myc gene. In contrast, sporadic BLs almost always show breakpoints within or close to the c-myc locus. C-myc translocation causes transcriptional dysregulation as a result of juxtaposition to the immunoglobulin gene because of the enhancer elements. Mutations within the c-myc locus also may contribute to the transcription of upregulation as well as functional dysregulation of the c-myc gene, which in turn promotes cell cycle progression and inhibits differentiation. *In vitro* studies in transgenic animals have shown cellular transformation as a consequence of c-myc upregulation supporting an important role for c-myc in lymphoma development.

Mechanism of c-myc cellular transformation is not fully understood; however, it is thought to be secondary to dysregulation of transcription of genes essential for the gene regulation. One of the key regulatory functions is c-myc interaction with Helix-loop-Helix (HLH) leucine zipper protein referred to as MAX. Transcription or regulation of myc is related closely to dimer formation with MAX, MAD, and MXI, two other HLH proteins. While myc/MAX heterodimer formation causes activation of

gene expression, MAD/MAX heterodimer formation causes transrepression of target genes. In a normal homeostatic state, cells form c-myc/MAX heterodimers while MAX/MAX homodimers compete for the similar target sequences to inhibit the transcription activation. In the case of BL, overproduction of c-myc causes excessive amount of c-myc/MAX complexes, which transactivates the target genes.

Primary Effusion Lymphoma

Primary effusion lymphoma (PEL) is a recently described distinct subtype of B-cell NHL associated with infection by the human herpesvirus-8, which is also associated with Kaposi's sarcoma, multicentric Castleman's disease, angioimmunoblastic lymphadenopathy, and multiple myeloma (47, 48, 49). PEL occurs predominantly but not exclusively among HIV-infected individuals. These lymphoid neoplasms show propensity to body-cavity involvement without mass lesions. There is little or no dissemination to other solid organs. Tumor cells show usually immunoblastic morphology, late B-cell phenotype. Genotypically of PEL cells reveal clonal Ig gene rearrangement in all cases and EBV infection in most cases. There are so far no chromosomal abnormalities noted in this entity. PEL carries a clinically aggressive course and have a poor prognosis.

Hodgkin's Disease

HD shows a heterogenous morphology and classified into four histologic subgroups, including lymphocyte predominance, nodular sclerosing, mixed cellularity, and lymphocyte depleted types. In general, lymphocyte predominance is biologically considered to differ from other three categories. Reed-Sternberg cells are a characteristic feature of HD and represent only a minority of (0.1% to 1%) cells in involved tissue. The RS cells are surrounded by T lymphocytes, histiocytes, fibroblasts, eosinophils, and other nonmalignant hematopoietic cells.

The etiology of HD remains unknown. Unlike NHLs, there are no consistent structural karyotypical abnormalities associated with HD. The most frequent cytogenetic abnormalities include 1p36, 6q21-q26, 14q11, and 14q32. There have been studies looking at the bcl-2 gene rearrangement using assays similar to those used for the Ig gene analysis. These studies showed variable results; however, in general, *bcl-2* positivity is believed to be rare in HD and positivity may be related to nonneoplastic cells having rare translocation of *bcl-2* gene (50). EBV has been extensively investigated in HD in different population. In the Northwestern American and European population, approximately 50% of HD cases are EBV-positive but reaching up to 90% South American population in some endemic areas. The prevalence of EBV is higher in mixed cellularity than nodular sclerosing type and extremely rare in nodular lymphocyte predominance type (51, 52). EBV is present in clonal episomal form suggesting that the cell had been infected before the neoplastic transformation and supports a possible role in pathogenesis of HD. The EBV infection is latent and shows expression of EBNA1 and LMP-1 but no EBNA2, consistent with type II pattern (53). Despite serologic and epidemiologic studies, the precise functional role for EBV in the pathogenesis of HD remains unknown.

Because of the scarcity of RS cells in tumor tissue, it has been exceedingly difficult to use standard techniques of molecular biology on HD tumor tissues to characterize molecular events involved in neoplastic transformation. Immunoglobulin and T-cell receptor gene studies demonstrate germline pattern in the great majority of HD cases most likely because of the presence of low number of RS cells. Isolation of single cells by microdissection followed by PCR analysis has allowed molecular analysis of RS cells. Using these methods, detection of Ig gene rearrangements in single RS cells of classical HD favored for B-cell origin. In majority of the cases, Ig gene rearrangements amplified from multiple RS cells showed a clonal relationship among the RS cells supporting a clonal origin. It is believed that RS cells are derived from germinal center (GC) B cells because there are somatic mutations within the rearranged Ig genes. Based on the finding of crippling mutations in some of the cases, it was suggested that RS cells in classical HD as a rule are derived from crippled GC B cells that lost the capacity to express antigen receptor because of somatic mutations (54). The discovery that the RS cell shows an immunoglobulin gene rearrangement suggests that HD has a physiologic similarity to B cells (55, 56 and 57).

Recent studies show constitutive expression of NFκB that plays an essential role in prevention of apoptosis in various tumors (58). Studies related to activation of NFκB including cytokines, EBV, and loss of inhibitory genes appear to provide newer insight into the biology of HD (59).

Cutaneous T-Cell Lymphoma and Sezary Syndrome

Sezary syndrome is a leukemic variant of cutaneous T-cell lymphoma (CTCL) in which patients typically present with generalized erythroderma, lymphadenopathy, and abnormal T-cells (Sezary cells) in the peripheral blood (60). The Sezary cells have a characteristic morphology with cerebriform nuclei and typically show a CD4 positive, CD7 negative immunophenotype. The clonal nature of the disease also has been shown by demonstration of rearrangement of the TCR-β and TCR-γ genes. Recent modifications, especially PCR amplification of TCR-γ genes now allow monoclonality determination in small samples (61, 62). Association between mycosis fungoides (MF), its leukemic variant Sezary syndrome (SS) and the human T-cell lymphotropic virus type-I (HTLV-I) has been controversial. Previous studies demonstrated HTLV-I provirus in the skin lesions of patients with MF by PCR analysis using oligonucleotide primers directed against common genes of HTLV-I such as gag, pol, env, or tax oligonucleotide primers (63, 64). However, a recent cooperative international study investigated 127 patients from five different countries with different geographical backgrounds using PCR with primers against all regions of the HTLV-I genome and failed to demonstrate an association between MF and HTLVI (65).

Adult T-Cell Leukemia/Lymphoma

Adult T-cell leukemia/lymphoma (ATLL) is a T-cell neoplasm characterized by cutaneous disease, lymphadenopathy, hepatosplenomegaly, and high leucocyte counts with peripheral atypical lymphocyte and hypercalcemia (66). The epidemiologic studies have shown endemic areas involving Japan, Caribbean basin, tropical islands of Pacific Ocean, and the Seychelles in the Indian Ocean. The black population in the Southeastern United States show higher incidence of ATLL. ATLL is believed to be

caused by human retrovirus HTLV-I, which binds CD4-positive cells and cause uncontrolled proliferation (67, 68). The diagnosis of ATLL is made by the characteristic clinical and histologic finding with the detection of serum antibodies to HTLV-I. Histologically, some cases may be difficult to differentiate from Sezary syndrome because both lesions are proliferation of CD4 T-lymphocytes. Monoclonal integration of HTLV-I proviral DNA in cellular DNA of ATLL cells may be needed as confirmation of clinical diagnosis ATLL.

Posttransplantation Lymphoproliferative Disorders (PTLDs). PTLDs represent a heterogeneous group of lymphoid proliferations that arise in immunosuppressed transplant recipients. Clinical course is variable. Some of these lesions regress after a reduction in immunosuppressive therapy, whereas others progress despite aggressive therapy. There are morphological, immunophenotypic, and immunogenotypic criteria developed to predict clinical outcome (69).

Molecular genetic characterization of PTLDs, including an assessment of the presence of structural alterations in some oncogenes and tumor suppressor genes, is highly informative in predicting clinical outcome. Great majority of PTLDs are associated with EBV (70, 71). PTLDs with monomorphic morphology and a strong clonal immunoglobulin gene rearrangement band on Southern blotting and a c-myc gene rearrangement are more likely to have progressive disease. There are three basic clinical presentations to predict clinical course:(i) plasmacytic hyperplasia (PH): immunoglobulin gene rearrangement is polyclonal, and EBV terminal repeat analysis is polyclonal or faint monoclonal by Southern blot analysis; (ii) polymorphic PTLD: monoclonal or oligoclonal by immunoglobulin gene rearrangement or EBV terminal repeat analysis and no other identifiable genetic alterations; and (iii) lymphoma/multiple myeloma: monoclonal with structural alterations of the N-ras, p53, and/or the c-myc genes. A single clonal band representing the joined EBV genomic termini can be detected in most monomorphic PTLDs, demonstrating the presence of a monoclonal expansion of B lymphocytes carrying EBV DNA. The polymorphic lesions lack structural alterations in most of the genes involved in other lymphoid neoplasms, including BCL-1 and BCL-2, c-myc, p53, and ras. According to this clinical classification, the PTLDs classified as PH and polymorphic PTLD are more likely to regress with a reduction in immunosuppression or to undergo resolution with more aggressive therapy than those classified as lymphoma or myeloma. Recently, an additional study looking at bcl-6 gene structure was shown to be useful to predict treatment response (72). Although BCL-6 gene shows a germline pattern, sequence analysis of BCL-6 illustrates a difference among patients. Patients with polymorphic PTLD and lymphoma/myeloma usually have a mutation in the first intron segments of the bcl-6 gene, which commonly is associated with progressive disease. In summary, complete core analysis of PTLD requires molecular assays for immunoglobulin gene rearrangement, EBV terminal repeat analysis, c-myc translocation, and bcl-6 mutation.

Anaplastic Large Cell Lymphoma

Approximately 2% to 8% of all NHLs contain a t(2; 5)(p23; q35) chromosomal rearrangement, which is found in mostly CD30-positive anaplastic large-cell lymphoma (ALCL) of T-cell or null cell phenotype. This translocation juxtaposes the nucleophosmin (*NPM*) gene at 5q35 with the anaplastic lymphoma kinase (*ALK*) gene at 2p23 (73, 74). NPM is a 38-kd protein involved in shuttling ribosomal components between the nucleus and cytoplasm (75). The fusion gene created by the (2;5) translocation encodes an 80-kd chimeric NPM-ALK kinase, in which 40% of the N-terminal of NPM is linked to the entire intracytoplasmic region of ALK. The NPM-ALK hybrid protein is thought to play a key role in lymphomagenesis by aberrant phosphorylation of intracellular substrates. It is believed that *NPM* gene promotes the expression of the ALK catalytic domain present in the chiameric NPM-ALK protein. Few cases with other partner genes with ALK have been described. It was shown that constitutive ALK activity contributes to the malignant transformation of lymphoid cells.

Although the vast majority of cases with a t(2; 5) rearrangement possess anaplastic large-cell histology, this alteration can be detected in other histologic forms of lymphomas, hence difficult to recognize. The expression of the CD30/Ki-1 antigen by NHL does not correlate strictly with the presence of NPM-ALK with approximately one half of CD30/Ki-1-positive ALCL also being *NPM-ALK*-positive and occasional CD30/Ki-1-negative ALCL being positive for the fusion gene. Despite controversy, recent studies described B-cell type ALCL among the morphological variants of diffuse large B-cell lymphoma. Reported cases with t(2;5) chromosomal translocation as well as expression of ALK protein suggests existence of B-ALCL. However, there is insufficient number of B-ALCL to analyze clinical features, including prognosis.

In the clinical setting, detection of NPM-ALK fusion transcript can be demonstrated easily by RT-PCR, FISH, and Southern blotting. Most groups failed to demonstrate t(2;5) in Hodgkin's disease; hence, NPM/ALK positivity is believed to be very characteristic of ALCL. Recent studies have shown a better prognosis for the ALCL; nevertheless, 20% to 30% of these patients eventually succumb to their disease despite aggressive therapeutic intervention. Extensive immunohistochemical studies now have been performed using antibody developed against ALK protein, and new information has emerged based on pathologic and clinical features of ALK-positive lymphomas. Based on these studies, ALK-positive ALCL was shown to have a better prognosis (76, 77). It has been proposed to distinguish these tumors from the heterogeneous group of lymphomas covered by the term ALCL.

Majority of cutaneous ALCL, some possibly including lymphomatoid papulosis, are also reported to have a good prognosis. Several studies have shown the absence of t(2;5) in majority CD30+ cutaneous lymphoproliferations, as opposed to its presence in CD30+ systemic ALCLs. Great majority of cutaneous ALCLs also are negative for ALK expression.

Leukemias

Improvements in cytogenetic techniques have provided significant insight as to the importance of molecular abnormalities in the pathophysiology and prognosis of hematologic malignancies. Of all neoplasms, leukemias have been by far the most intensively investigated and account for the great majority of chromosomal

aberrations documented in human neoplasms. The majority of cases of acute lymphocytic leukemias demonstrate an abnormal karyotype, either in chromosome number or as structural changes such as translocations, inversions, or deletions (Table 11.2). Although leukemias currently are classified according to morphology and cell lineage as determined by immunological methods, it has become apparent that leukemias with a certain lineage comprise a heterogenous disease mostly because of the characteristic of molecular abnormalities associated with certain cytogenetic changes. Despite the fundamental importance of the underlying molecular abnormalities both in acute myeloid and lymphoid leukemias, there has not been an acute leukemia classification that systematically incorporates the results of such alterations. In this section, the most common translocations (see Table 11.2) found in leukemias with an emphasis on the pathophysiological and clinical implications will be discussed.

TABLE 11.2. COMMON RECURRING CHROMOSOMAL TRANSLOCATIONS ENCOUNTERED IN ACUTE LEUKEMIAS

Disease	Translocation	Genes Involved
AML	t(6;9)(p23;q34)	DEK-CAN
AML-M2	t(8;21)(q22;q22)	AML1-ETO
AML-M3	t(15;17)(q22;q21)	PML-RARα
AML-M4Eo	inv(16)(p13;q22)	CBFß-MYH
AML-M4/M5, biphenotypic	t(11;19)(q23;p13)	MLL-ENL
AML-M5, biphenotypic	t(9;11)(p21;q23)	MLL-AF9
CML	t(9;22)(q34;q11)	BCR-ABL
Precursor B-ALL	t(12;21)(p13;q22)	ETV6-AML1
	t(9;22)(q34;q11)	BCR-ABL
	t(4;11)(q21;q23)	MLL-AF4
Pre-B-ALL	t(1;19)(q23;p13)	E2A-PBX1
B-ALL	t(8;14)(q24;q32)	MYC-IgH

AMI, Acute myelogenous leukemia
ALI, acute lymphoblastic leukemia

Acute Myelogenous Leukemias AML1/ETO [t(8;21)(q22;q22)

The (8;21)(q22;q22) translocation is one of the most frequent karyotypic abnormalities detected in acute myelogenous leukemia (AML), accounting for approximately 40% of de novo AML cases that have M2 morphology in the French-American-British (FAB) classification. This karyotypic abnormality is associated with a good clinical response to chemotherapy and a high remission rate with a long-term disease free survival. As a result of this translocation, the gene encoding the *AML1* (*CBFA2*) transcription factor from chromosome 21 is fused to the "eight-twenty-one gene" (ETO, also known as MTG8 and *CBFA2T1*) on chromosome 8 (78). This fusion gene encodes a chimeric AML1-ETO protein consisting of the N-terminus of AML1 fused in frame to the C-terminus of ETO. The chimeric protein is believed to induce leukemogenesis by its abnormal activation of both myeloid and lymphoid target genes through the AML1/CBFB transcription factor complex (79). One of the targets of this translocation, AML1, is the DNA-binding subunit of the AML1/CBF transcription factor complex and binds the enhancer core motif, TGT/cGGT. DNA-binding affinity of AML1 is increased through heterodimerization with CBF, and both its DNA-binding and interaction with CBF are mediated through a central domain with high homology to the *Drosophila* segmentation gene, *runt*. The AML1-ETO chimeric product retains this domain and therefore also binds the core enhancer sequence and interacts with CBF.

Transcriptional regulation by AML1 through the enhancer core motif has been shown to be important for the tissue-specific expression of a number of hematopoietic-specific genes including interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor (GM-CSF), the receptor for CSF-1 (CSF-1R), myeloperoxidase, and subunits of the T-cell antigen receptor (TCR).

The t(8;21) can be detected by molecular studies using Southern blotting or RT-PCR. One of the important caveats for the molecular detection by PCR is that despite clinical remission with long duration patients remain PCR positive for the AML1/ETO transcript without clinical progression (80). In fact, some patients reported to be positive as long as 10 years after complete morphologic and cytogenetic remission. This data suggests that RT-PCR analysis for the minimal residual disease for this patient population should be interpreted with a caution. In addition, RT-PCR studies also have shown that AML1/ETO transcripts can be detected in patients where there is no evidence of the t(8;21) at the karyotypic level indicating very low number of cells without evidence of leukemia.

Acute Promyelocytic Leukemia, t(15; 17)(q22; q21)

Acute promyelocytic leukemia (APL) is associated characteristically with the reciprocal translocation of t(15; 17)(q22; q21), leading to the formation of *PML-RARA* as well as *RARA-PML* fusion genes (81). PML-RARA, transcribed by the fusion gene, retains all the functional domains considered to be of importance in APL. PML-RARA is believed to play a key role in leukemogenesis, as demonstrated in transgenic animal model. However, the role for del(17q)-derived RARA-PML in the development of APL remains unclear, particularly because reciprocal transcripts are not detected in all of the APL patients (82). The pathogenesis of APL in cases associated with a PML-RAR *A* rearrangement is believed to be secondary to a block in myeloid differentiation that leads into abnormal proliferation of promyelocytes. This block in differentiation may be overcome by retinoids such as all-*trans* retinoic acid (ATRA) (83). RARA is a member of the steroid hormone nuclear receptor family, acting as a transcription factor mediating the effect of retinoic acid at specific response elements. There are other translocations of the RARA gene associated with APL that have been characterized at the molecular level. These include: t(11;17)(q23;q21), t(11;17) (q13;q21) and t(5;17)(q35;q21), causing fusion of RARA to nucleophosmin (NPM), promyelocytic leukemia zinc finger (PLZF), and nuclear mitotic apparatus (NuMA) genes, respectively. One of the important aspects of APL with t(11;17)(q23;q21) is the lack of clinical response to ATRA therapy (84).

There are three major breakpoints in PML gene (intron 6, bcr1; exon 6, bcr2; intron 3, bcr3) that generate two transcripts of different length by RT-PCR. These are referred to as long (L)

and short (S) forms. L includes bcr1 and bcr2 or variable (V) transcripts, which can only be distinguished by DNA sequence or specific oligonucleotide hybridization. Great majority of PML patients show bcr1 and bcr3 transcripts. Correlations between type of PML/RAR transcript and response to therapy and outcome have been investigated; however, the impact of this separation currently is unknown.

The diagnosis of APL is mainly supported by the molecular detection of PML-RARA transcripts using RT-PCR or FISH analysis to detect reciprocal translocation because cytogenetic analysis sometimes fails because of only erythroblasts entering mitosis or no microscopically detectable rearrangement is present. The molecular analysis is particularly useful in cases classified as microgranular variant of APL. In cases with morphologic features of APL that lack PML-RARA transcript, they should be further investigated for the presence of other cytogenetic abnormalities mentioned above. In addition, most of the studies indicate that a positive PML/RARA by PCR after consolidation is a strong predictor of subsequent hematologic relapse, whereas repeatedly negative results are associated with long-term survival in the majority of patients.

Inversion 16

The pericentric inv(16)(p13q22) and related t(16;16)(p13;q22) cytogenetic abnormalities are observed in approximately 10% of *de novo* acute myeloid leukemia (AML) cases and typically associated with M4 with eosinophilia (M4EO) FAB subtype. However, some cases may lack AML-M4EO histology (85). This subtype of AML has been considered to have relatively favorable therapeutic outcome. The translocation results in the molecular fusion between the CBFB gene on 16q22 (encoding core binding factor-subunit [CBF]) and the MYH11 gene on 16p13 (encoding a type II smooth muscle myosin heavy chain [SMMHC]) (86). The fusion protein CBF-SMMHC disrupts normal myelopoiesis by inhibiting the function of endogenous CBF and its heterodimerizing AML1 (CBF) partners. This, in turn, alters the expression patterns of critical target genes in myelopoiesis. The large majority (>85%) of cases of inv(16)/t(16;16) AML are associated with a type A fusion transcript, corresponding to an in-frame CBFB nucleotide 495-MYH11 nucleotide 1921 junction. RT-PCR technique is advocated for detection CBFB-MYH11 fusion gene transcripts that encode a CBF-SMMHC fusion protein in cases histologically consistent with AML-M4 (86).

Philadelphia Chromosome, Chronic Myelogeneous Leukemia

Philadelphia chromosome primarily is seen in chronic myelogeneous leukemia (CML) and acute lymphocytic leukemia. It results from a reciprocal translocation between the long arm of chromosomes 9 and 22, transposing the c-abl gene from chromosome 9 to the bcr gene on chromosome 22, causing a hybrid BCR-ABL gene (87) (Fig. 11.13). The breakpoints in the ABL gene occupy a large segment of an intron sizing 300 kb and referred as major breakpoint cluster region (M-bcr) and minor cluster region (m-bcr). The great majority of the CML cases show translocation into M-bcr that is spanned between exons 12 and 16. This region historically has been named exons b1 through b5. Majority of bcr-abl translocation causes a fusion transcript involving the exon b2 or b3 of the M-bcr region and a2 of the ABL gene generating transcripts b2-a2 and b3-a2 (Fig. 11.14) (88).

The incidence of the BCR-ABL fusion gene in CML patients is underestimated by traditional karyotypic studies. Other molecular assays, like Southern blot analysis, FISH, PCR, and pulsed-field gel electrophoresis to complement cytogenetic analysis now are used commonly in clinical evaluation of CML patients. Although Southern blot analysis had played an important role in early description of molecular structure of BCR-ABL, this assay currently is considered to be inferior as a single diagnostic tool.

Cytogenetic analysis is essential to diagnose and follow the course of CML and for identifying karyotypic changes, but it has a limited utility for minimal residual disease evaluation of CML patients because of its lower sensitivity (1% to 5%). Therefore, more sensitive clinical assays for the molecular diagnosis and follow-up of residual disease now are being used. FISH and its variants, interphase and hypermetaphase FISH, allow analysis of a larger number of cells (>500) in a timely and efficient manner, with the added benefit that interphase FISH can be applied to nondividing, low proliferative fraction of cells.

In many centers, molecular-based assays, such as PCR and FISH, are becoming the techniques of choice for detecting residual disease on a molecular level. Particularly, the real-time quantitative PCR method appears to be a very reliable assay for detection of residual disease (89). Depending on the method used, i.e., nonnested or nested, PCR allows identification of a single Ph-positive cell among 10^4 to 10^8 normal cells, a sensitivity unparalleled by any other available method. This level of sensitivity provided opportunities to look at minimal residual disease (MRD) following eradication of leukemia based on morphologic evaluation. Parameters considered to be critical for the interpretation of MRD studies include (i) treatment protocol; (ii) time of sampling in respect to treatment; (iii) controls for housekeeping (e.g., actin, hemoglobin) and target genes; (iv) sensitivity of the PCR assay; (v) inter-laboratory standardization; (vi) duration of study; (vii) retrospective or prospective nature of the study; and (viii) number of tests per patient.

In patients who are in cytogenetic remission after treatment with interferon, PCR techniques still can commonly identify BCR-ABL transcripts within the first year (90). The PCR analysis of CML patients who underwent bone marrow transplantation has been variable. Miyamura et al. (91), using nested PCR, detected no association between PCR positivity and subsequent relapse in their series of 64 patients with CML in remission after allogeneic stem cell transplantation. Interestingly, the persistence of detectable BCR-ABL transcript for up to 2 years postremission did not result in disease recurrence in some patients. In contrast, Hughes et al. (92), on the basis of nested PCR study, reported that PCR positivity within 6 months after allogeneic transplantation did not predict a worse outcome, whereas PCR positivity later than 6 months after transplantation associated with progressive disease. In another multivariate analysis, PCR-positivity at 6 to 12 months posttransplantation was identified as an independent variable for influencing subsequent relapse (93). In a quantitative PCR analysis of BCR-ABL transcripts in contrast

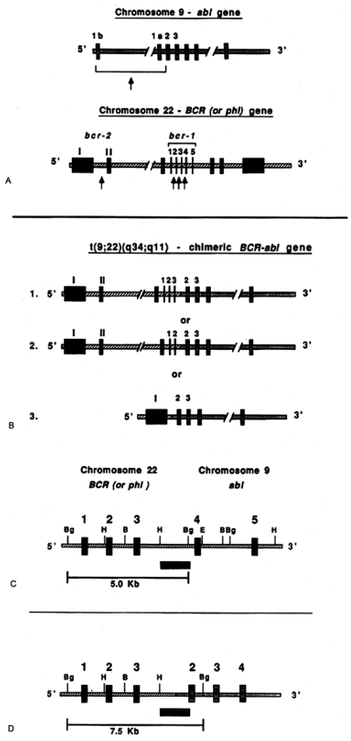


FIGURE 11.13. Molecular gene structure and schematic of the Philadelphia chromosome. The drawings do not accurately reflect distances and are drawn for schematic purposes only. (**A**) Structure of the normal chromosome 9, *c-abl* proto-oncogene, and chromosome 22, the *BCR* gene. Boxes represent exons. Arrows indicate breakpoint regions. The large bracket with the arrow at the *c-abl* gene represents a 100 kb region where virtually all translocations occur. *bcr-1* and *bcr-2* represent the breakpoint cluster regions in the *BCR* gene. Arabic numbers and roman numerals indicate the given nomenclature for the *abl* and *BCR* exons. (**B**) As a result of the reciprocal translocation between chromosome 9 and chromosome 22, a chimeric *BCR-abl* gene is formed. (1) and (2) represent over 95% of the translocations occurring in the bcr-1 region with breakpoints occurring between exons 3 and 4 or exons 2 and 3, respectively; 3) represents the reciprocal translocation of chromosome 9 into the first 5′ intron (*bcr*-2) of the *BCR* gene. (**C**) The germline configuration of the bcr-1 region (Bg = *Bgl*/II restriction enzyme) shows a 5.0-kb, *Bgl* /II fragment with the *bcr* probe. (**D**) Rearrangement of c-abl into the *bcr* region replaces the 3′ Bgl/II restriction enzyme site, leading to the formation of a larger, 7.5-kb band.

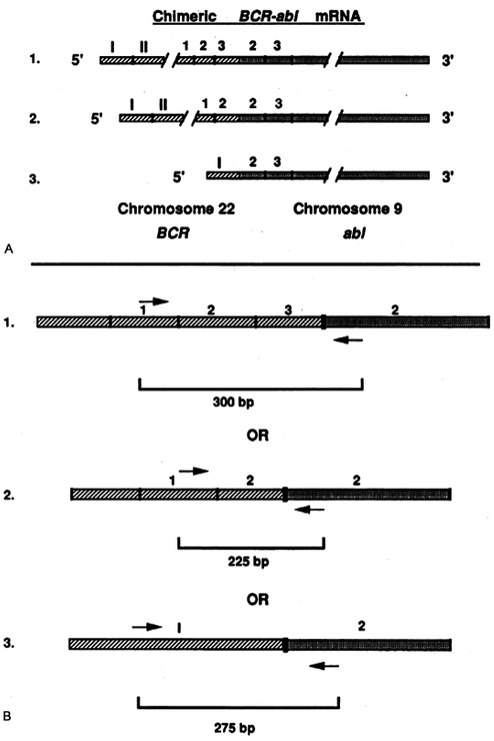


FIGURE 11.14. Gene structure of the chimeric *BCR-abl* mRNA and polymerase chain reaction (PCR) analysis. See Figure 11 for explanation of characters. (**A**) (1) and (2) represent the *bcr-abl* mRNA that is formed after translocation into *bcr*-1. (3) represents translocation of c-abl into the 5′ intron (*bcr* -2) that is associated with Philadelphia chromosome-positive ALL. (**B**) Schematic of PCR analysis for *BCR-abl* mRNA. The horizontal arrows represent the locations for oligonucleotide primers, 1) and 2). A 300 base pair fragment is amplified if an exon 3 (*bcr-1*) rearrangement has occurred or a 225 base pair is amplified if an exon 2 (*bcr* -1) rearrangement has occurred. (3) PCR amplification of a 5′ intron (*bcr*-2)-*abl* translocation.

to qualitative PCR data, the probability of relapse was shown to be significantly increased in patients with higher versus lower levels of BCR-ABL. Some studies suggest that serial PCR measurements identify patients at high risk (i.e., BCR-ABL positivity in serial samples) and low risk of relapse (i.e., transient positivity). So far, because of different levels of PCR, sensitivity and short clinical follow-up durations make it difficult to decide the clinical utility of PCR. However, it is safe to conclude that repeated PCR-negative patients are less likely to relapse. The PCR technique, although considered a valid clinical testing procedure, should be used cautiously in this setting until sufficient data are available.

Acute Lymphoblastic Leukemia (ALL)

The cytogenetic abnormalities confer important prognostic information in ALL. Complete remission rates, remission durations, as well as disease-free-survivals were significantly affected by the karyotypic abnormalities. Translocations constitute the most common changes and there are many different nonrandomly occurring rearrangements described in acute leukemias, the most common abnormalities being listed in Table 11.2.

Philadelphia Chromosome, ALL

The t(9;22) translocation is among the most frequent abnormalities encountered in ALL. In children, Ph+ ALL has a dismal prognosis even with intensive chemotherapy programs that have improved survival in other cytogenetic subgroups. Similarly, in adult ALLs, despite high remission rates comparable to those of Ph-negative ALL, remission duration and survival times are short. Therefore, patients with Ph+ ALL are suitable candidates for innovative and intensified strategies.

As discussed in the section of chronic myelogeneous leukemia, the breakpoint in the ABL region occurs anywhere in a 300-kb intron, but usually is located 5′ of ABL exon a2 in patients

with Ph-+ ALL. Deletions of ABL exon a2 and in-frame joining at the mRNA level of 5′ BCR sequences to the ABL exon a3 also have been described (94). In 50% of adult Ph ALL and in 80% of childhood Ph+ ALL, and very rarely cases of CML or AML, the breakpoint on chromosome 22 falls 5′ of the M-bcr, within a long intron segment separating the alternative exon e2′ from e2 called the minor breakpoint cluster region (m-bcr) (95). Splicing out exons e1′ and e2′ creates an e1a2 junction in the BCR-ABL transcript, translating into a smaller BCR-ABL fusion protein referred to as p190BCR-ABL. Both p210BCR-ABL (CML type) and p190BCR-ABL (ALL type) have significantly increased tyrosine kinase activity compared to the normal human c-ABL protein, p145. *In vitro* studies showed that p190BCR-ABL has a more tyrosine kinase activity than p210BCR-ABL.

Because breakpoints on the BCR gene occur outside the M-bcr region in an area too large to be reliably recognized, Southern blotting used only detection of p210BCR-ABL but not p190BCR-ABL, which is found in 50% to 80% of Ph+ ALL. Studies investigating the clinical significance of fusion protein (p190 vs p210) showed no difference between those with an M-bcr breakpoint and those with a different breakpoint location for age, immunophenotype, or outcome (96). Common association of p190BCR-ABL with ALL and p210 BCR-ABL with CML has been argued to favor a distinction between true *de novo* ALL (m-bcr, p190BCR-ABL) and CML in lymphoid blastic phase (M-bcr, p210BCR-ABL). It has been noted that there is a difference in the frequency of expression of p190BCR-ABL between pediatric and adult ALL (80% vs. 50% of Ph+ ALL cases)(97).

Interestingly, p190BCR-ABL transcripts are not confined to primary acute leukemias. Using quantitative reverse transcription-PCR (RT-PCR) for p210BCR-ABL and p190BCR-ABL mRNA, Van Rhee et al. detected p190BCR-ABL mRNA in 88% (14 of 16) of patients with CML in chronic phase, 100% (10 of 10) of patients with CML in lymphoid blastic phase, and 100% (10 of 10) of cases with p210BCR-ABL-positive ALL (98). However, the p190BCR-ABL/p210BCR-ABL ratio was 10 times greater in ALL than CML, whereas the ratio was similar in CML at diagnosis, in chronic phase, and in lymphoid blastic phase.

RT-PCR has been used to detect and monitor the bcr-abl fusion genes associated with the t(9;22) that occurs in approximately 30% of adults with ALL. Molecular diagnosis detecting bcr-abl fusion protein by RT-PCR has been shown to be a very sensitive technique compared to standard cytogenetic analysis for identification of bcr-abl positive cells. In adult ALL patients expressing B-cell phenotype, greater than 30% of the ALL associates with bcr-abl. The great majority of these patients show p190 molecular-size protein in contrast to the p210 bcr-abl fusion protein that is present in the great majority of the patients with CML. Molecular studies have revealed that even though the complete remission following induction therapy is similar in both bcr-abl positive and negative cases, ALL with bcr-abl transcript has significantly shorter duration of remission, which is associated with shorter survival. It appears that allogeneic transplantation may provide improved long-term disease-free survival for some of the bcr-abl positive ALL patients. There is no significant clinical difference in terms of response between patients with p190 bcr-abl compared to p210 bcr-abl. RT-PCR also allows detection of minimal residual disease following bone marrow transplantation. Patients who are consistently bcr-abl negative following transplantation are not likely to relapse and may become long-term survivors. However, patients who are positive for the bcr-abl transcript following transplantation carry a high risk for relapse.

The MLL gene, the t(4;11)(q21;q23)

The MLL gene, located on 11q23 has been shown to be involved with more than 30 different translocations in both acute lymphocytic and myeloid leukemias, as well as biphenotypic leukemias. Of all the various chromosomal loci participating in 11q23 translocations, the t(4;11)(q21;q23) is by far the best characterized structural lesion, with specific biological and clinical features (99). The presence of t(4;11) has been shown to be associated with a very poor prognosis in response to conventional chemotherapy in both pediatric and adult patients. Therefore, similar to patients with Philadelphia chromosome, patients with t(4;11) are recommended for more intensive regimens.

The t(4;11)(q21;q23) is seen in 2% of the pediatric patients with ALL and in approximately 3% to 5% of adult patients with ALL (100). It is especially very high in infant acute leukemias, reaching up to 60% to 70% in incidence. The frequency of the 11q23 abnormality also is higher in patients who had been previously treated with topoisomerase II inhibitors with a frequency of up to 80%. The t(4;11), regardless of the age of the patients, shows a B-cell phenotype with expression of CD10, CD19+, CD24, and commonly myeloid-associated antigens. These patients typically show hyperleukocytosis, organomegaly, a high incidence of CNS leukemia, and carry a bad prognosis, especially in infant ALL. Leukemic transformation is considered to involve a stem cell or an early committed progenitor cell with the potential for both lymphoid and myeloid differentiation (101, 102). The translocation usually involves a gene located on chromosome 11q23 referred to as mixed lineage leukemia or myeloid lymphoid leukemia (MLL) gene, which has at least 21 exons and encodes for 431 kd protein (103). The C-terminal domain shows significant homology to *Drosophila trithorax* protein. The N-terminal region has methyltransferase domain and three AT-hook motifs. The AT-hook motifs have been implicated in binding to the minor groove of the DNA double helix containing AT-rich regions. The MLL protein is considered to be an important regulator of differentiation by its effects on other DNA-binding proteins or direct interaction with DNA. The breakpoint cluster lies within the 8.3 kb region between exons 5 and 11, allowing easy identification of the translocation by Southern blot analysis (104). The translocation can be detected by RT-PCR amplification of the fusion transcript containing MLL and AF4, the gene located on 4q21. However, there are many reciprocal translocation sites for 11q23 translocations including 4q21, 9p22, 19p13, and 1p32 that can be missed by PCR. The translocation t(11;19)(q23;p13) shows similar clinical and prognostic features as the t(4;11) translocation.

The t(1;19)(q23;p13)

The t(1;19)(q23;p13) is one of the most common recurring translocations in pediatric ALL, with a frequency of 5% to 6%

overall, and of 25% in pre-B-cell ALL. It is present in less than 5% of adult ALL cases. Clinical characteristics of pre-B-ALL with t(1;19) include presentation with high WBC counts, increased LDH levels, and a DNA index of less than 1.6, which underscores that this type of leukemia is usually associated with a pseudodiploid karyotype. It frequently occurs in black people. The majority of cases are positive for CD10, CD19, and cIg (105).

Although previously considered to be a poor prognostic factor, with the use of contemporary therapies, the implication of overall clinical significance is lost (106). This translocation involves E2A, which encodes the two transcription factors E12 and E47 via alternative splicing of precursor mRNA. E12 and E47 are ubiquitous helix-loop-helix (HLH)-containing proteins that bind to the E-box element in the light-chain DNA-enhancer region and plays an important role in lymphopoiesis and regulation of B-cell development. Breakpoints in the E2A gene involves 3.5-kb intron segment between exons 13 and 14. The breakpoint region on chromosome 1q23 is dispersed and lies within an intron of at least 50 kb in size. It disrupts the homeobox-containing “pre-B-cell leukemia” gene (PRL, also referred to as PBX1), which is transcriptionally silent in lymphoid cells. The translocation fuses the 5′ sequences of the E2A gene with 3′ sequences of PBX1 resulting in fusion transcript with the activation domain of E2A and PBX1. E2A-PBX1 appears to function as a potent transcriptional activator that can be detected by RT-PCR. However, detection of E2A-PBX1 by PCR during therapy is not useful predictor of relapse (107).

t(12;21)((p13;q22).(ALL)

One of the most significant developments in pediatric hematologic neoplasms during last couple of years has been the definition of molecular abnormality seen in ALL with t(12;21) (p13;q22) (108, 109). Based on cytogenetic studies, the t(12;21) has been considered to be rare in ALL. However, t(12;21) has been difficult to detect by routine cytogenetic studies that grossly underestimate its frequency. In several studies, the application of molecular tools such as FISH, Southern blot analysis, and RT-PCR have shown TEL-AML1 fusion in up to 27% of children with B-lineage ALL, making it the most common karyotypic-molecular abnormality in pediatric ALL (110). Pediatric B-precursor ALL with TEL-AML1 rearrangement cases carry a good clinical outcome independent of age, and WBC count at presentation (111, 112). Overall, the frequency of TEL-AML1-positive ALL is low in adult B-lineage ALL (3% to 4%), and little is known about its prognostic significance.

The exact mechanism for transformation currently is unknown. Translocation generates a fusion of 5′ end of TEL gene, which contains a helix-loop-helix (HLH)-motif enabling interaction with other proteins, to functional domain of the AML1 gene. AML1 is a heterodimer protein with α and ß units. The α unit contains DNA binding and transactivation motif referred to as “runt,” because of its homology to drosophila runt gene. Hence, the fusion protein contains both HLH domain of TEL and both DNA binding and runt domain of AML1. It is interesting to note that this translocation is the first demonstration of involvement of the AML1 gene in lymphoid neoplasm. TEL-AML1 usually is associated with loss of other TEL allele indicating that loss of TEL function also may be important in leukemic transformation. Other important factors in pathogenesis are transcriptional repression of TCR enhancer and dominant-negative inhibition of AML1.

Tumor Suppressor Genes

Tumor suppressor genes are groups of genes whose expression can block the development of a tumorigenic phenotype and, upon inactivation, lead to a malignant process. The presence of these genes has been identified in certain chromosomal locations by analyzing loss of material from specific chromosomes. Most tumor suppressor gene abnormalities are detected by molecular techniques such as Southern blot analysis, PCR, single strand conformation polymorphism (SSCP), nucleotide sequencing, pulsed-field gel electrophoresis or chemical mismatch cleavage. Although tumor suppressor genes frequently are deleted in solid tumors, they appear to be present in only a fraction of cases of human leukemias. p16INK4A, with or without codeletion of p15INK4B, is the most common gene deleted in acute leukemias (mainly T-cell-derived leukemias). The incidence of homozygous deletions of p16INK4A has been as high as 80% in some series.

The retinoblastoma (Rb) and the p53 genes are the tumor suppressor genes that have been most thoroughly investigated. The Rb gene, on chromosome 13q14, was isolated and cloned by virtue of its deletion in retinoblastoma. It binds and inhibits several transcription factors (such as E2F) and causes inhibition of cell-cycle progression at the G1-S phase. It is deleted in 12% of T-ALL patients and about 20% of patients with CD10+ early pre-B-ALL. The incidence of p53 alterations in ALL is low in both pediatric and adult ALL cases. There appears to be a correlation between p53 mutation and progressive or relapsing disease, especially in T-cell ALL (113). In B-lineage disease, p53 mutations are observed frequently with translocation t(1;19), t(8;14), and the FAB L3 morphology.

NON-NEOPLASTIC DISORDERS

Part of "11 - Clinical Applications of Molecular Biology Hematopoietic Disorders"

Fanconi Anemia

Fanconi anemia (FA) is a rare autosomal recessive disease characterized by multiple congenital abnormalities and bone marrow (BM) failure. Patients with FA shows growth retardation and skin abnormalities, upper extremity defects, kidneys and gastrointestinal system abnormalities, and a higher susceptibility to develop cancer. The diagnosis of FA is usually established by demonstration of increased spontaneous chromosomal breakage that is amplified by the addition of the cross-linking agents such as diepoxybutane or mitomycin C (114). Although this clinical assay is highly sensitive and specific test for FA, it does not distinguish FA carriers from the general population. Furthermore, this assay is limited due to false-negative results, particularly in patients with cellular mosaicism (115).

Molecular studies revealed that FA is a genetically heterogenous disorder, unlike the congenital syndromes such as ataxia-telangiectasia (AT) and Bloom's syndrome, which arise from mutations in single genes. To establish molecular defect, somatic cell fusion studies, and complementation analysis were used which so far lead to identification of at least eight molecular abnormalities

(*FANC A-H*) in FA patients and four different genes (FANCA, FANCC, FANCD and *FANCG*). Localization of the major FA gene, *FANCA*, was facilitated by linkage analysis in families assigned to the FA-A complementation group. Mutations in the FANCA is found in 60-70% of patients (116).

There are a relatively small number of characteristic mutations, associated with specific ancestral backgrounds. A splice site mutation in intron 4 (IVS4+4 A > T mutation) is found in patients of Ashkenazi-Jewish ancestry and accounts for greater than 80% of FA in this population (117). Patients homozygous for this mutation show more severe FA, with multiple congenital abnormalities and early onset of hematological disease. Less common mutations include 322delG and O13X in exon 1, R185X in exon 6, R548X and L554P in exon 14. Patients with the 322delG mutation have a comparatively mild FA, with fewer congenital abnormalities and later onset of hematological disease (118). Additional pathogenic mutations identified that includes mutations in exon 1, exon 6, and exon14. Most of these mutations result in either truncation or internal deletion of the FAC protein. Only one pathogenic missense mutation (L554P) has been described.

Paroxysmal Nocturnal Hemoglobinuria

Paroxysmal nocturnal hemoglobinuria (PNH) is an acquired disorder of the hematopoietic stem cell characterized by clonal hematopoietic cells. Clinical manifestations are mainly due to intravascular hemolysis, venous thrombosis, and deficient hema-topoiesis. PNH often arises in patients with aplastic anemia (AA), accompanied by a marked improvement in their hematological indices, and it has been shown that some patients with PNH develop bone marrow suppression ultimately leading into aplastic anemia. There is a defective biosynthesis of glycosylphosphatidyl-inositol (GPI). The GPI deficiency causes defective surface expressions of at least twenty proteins that are normally anchored to the cell membrane via GPI. Deficiencies of GPI-anchored complement regulatory proteins, such as decay accelerating factor (DAF) and CD59, cause red blood cells to be very sensitive to complement resulting in complement-mediated hemolysis and hemoglobinuria. Somatic mutation in the X-linked gene PIG-A that is essential for the first step in GPI biosynthesis accounts for GPI deficiency in all patients with PNH characterized to date. Approximately, 10 or more genes are involved in biosynthesis of the GPI anchor. Although mutation in any one of these genes could result in GPI-anchor deficiency, the *PIG-A* gene is the cause of mutation in almost all PNH patients. The majority of mutations in *PIG-A* gene is a frameshift mutation of the coding region; however, there are also missense and nonsense mutations, abnormal splicing, and rarely cases with amino acid deletions have been reported. The mutations are distributed widely and heterogeneously in the coding regions and splice sites. It appears that there is no apparent clustering or a hot spot within the *PIG-A* gene.

Hereditary Hemochromatosis

Hereditary hemochromatosis (HH) is a common autosomal recessive genetic disorder of iron metabolism. It is very common among whites, with mutation affecting 1 in 300 individuals. (119, 120). It is characterized by increased iron absorption and storage that results in systemic organ damage. The clinical complications due to iron overload can be prevented by early diagnosis and treatment. Recent identification of the genetic abnormalities allowed development of a reliable genetic-based test that is expected to provide more effective preventive treatment. By positional cloning in 1996, the *HFE* gene was identified as the gene responsible for hemochromatosis. The *HFE* gene encodes an HLA class I-like protein that, in association with 2-microglobulin, has an expression pattern correlated with the localization of iron absorption. HFE seems to play a role in iron uptake by interacting with the transferrin receptor, leading to its decreased affinity for transferrin.

Two common missense mutations have been described: C282Y, accounting for 80% to 90% of HH chromosomes, and H63D, which is associated with a milder form of the disease representing 40% to 70% of non-C282Y HH chromosomes. Minority of HH patients show 193AT substitution with S65C missense substitution which accounts for 8% of HH not related to C282Y or H63D (121).

The HFE gene encodes an HLA class I-like protein that has an expression pattern correlating with the localization of iron absorption. Clinical studies have demonstrated that genetic studies requires further clinical assays since there are a number of cases within hemochromatosis families with C282Y mutation showing no iron overload. This may represent incomplete penetrance, which has not been systematically analyzed in the general population. Currently, the molecular testing is considered to be useful in a patient who is clinically suspected of having hemochromatosis or shows elevation of transferrin saturation or serum ferritin. Other clinical use is the investigation of siblings and other family members of a homozygote.

GENETIC FACTORS ASSOCIATED WITH COAGULATION DISORDERS

Part of "11 - Clinical Applications of Molecular Biology Hematopoietic Disorders"

Factor V Leiden and Prothrombin II Mutation

During the last couple of years, specific etiology of many hereditary causes of venous thromboembolism has been established. One of the most common known inherited causes of venous thrombosis include Activated Protein C Resistance (APCR) and deficiencies of the natural anticoagulants including Protein C, Protein S, antithrombin, and prothrombin. It currently is believed that the great majority of young patients with idiopathic venous thrombosis carry mutation in factor V (Factor V Leiden) or factor II genes. Molecular studies looking for genetic risk factors usually are included in the diagnostic work-up of a young patient with venous thrombosis to search for factor V and prothrombin gene mutation.

Factor V Leiden was found to be most common cause of APC resistance(122, 123). This is because of a common mutation within the Factor V gene involving a single point mutation that causes a change in aminoacid arginine (Arg) at position 506 to glutamine (Gln) in the Factor V protein (FVA-Arg 506-Gln). Only a minority of patients has been shown to have abnormality involving other portions of the Factor V gene. The FVA-Arg 506-Gln is the site at which APC cleaves factor Va; hence mutation causes resistance to proteolytic cleavage and functional activation

(Fig. 11.15). Most patients with APC resistance are identified as heterozygous for the factor V Leiden mutation; however, some individuals are homozygous who have markedly increased risk for thrombotic events compared to the heterozygotes (124). In a retrospective case-controlled study of healthy individuals over age 40, heterozygosity for the factor V Leiden was found in 12% of patients with an episode of thromboembolic event and 6% of normal individuals (125). In patients older than 60 and with an episode of venous thrombosis, heterozygosity was demonstrated in 26% of individuals. The risk for the thromboembolic disease in heterozygous individuals was reported to be considerably higher during pregnancy and oral contraceptive use. However, there appears to be no association between factor V Leiden and myocardial infarction in young women (126). A minority of individuals with APC show mutations at the arginine-306 of factor V, Factor V Cambridge, which is also reported to be associated with APC resistance (127).

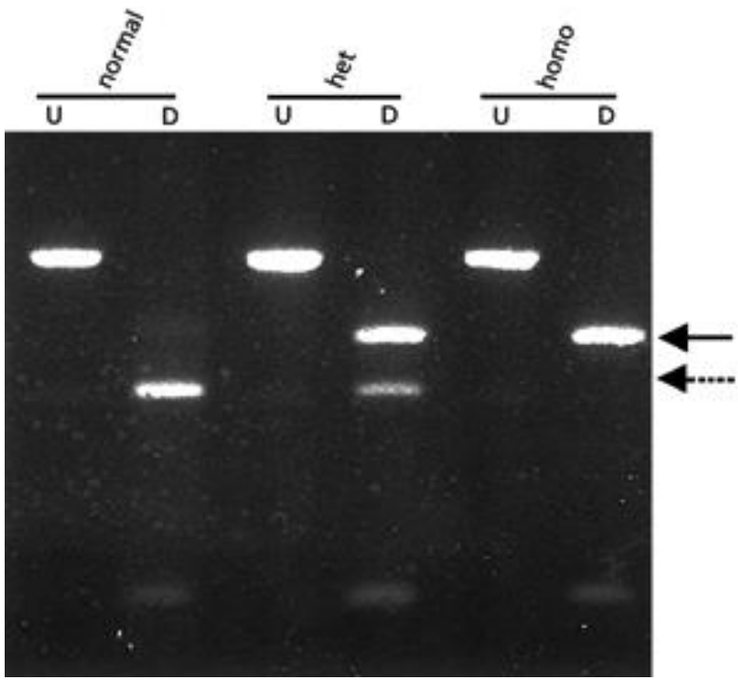


FIGURE 11.15. Detection of factor V Leiden. Exon 10 of Factor V gene is PCR amplified and cut with (d) resection enzyme *Mnl*I. DNA from normal (N) individuals contains MnlI, recognition site (GAGG), which is lost in individuals carrying mutation. Undigested DNA in normal (N) individuals shows 267 bp (upper bands) DNA fragment. After resection enzyme treatment, normal individuals show 163 bp fragment (dashed arrow). In homozygous (Hom) patients, digestion shows 200 bp fragment (solid arrow). In heterozygous (Het) patients, both 163 bp and 200 bp fragments are seen. Internal control of 67 bp in digested samples are seen in the bottom (Abbreviations: U = undigested; D = digested).

Prothrombin Gene Mutation

Mutations in the prothrombin gene cause substitution of glutamine for arginine at position 20210 in the 3′ untranslated region of the prothrombin gene (128). This mutation is associated with a 2.8-fold increase in risk for development of venous thrombosis (129, 130). In arterial thrombosis, the presence of prothrombin 20210A is considered to carry a markedly increased risk factor for the development of myocardial infarction in women, particularly with patients carrying high risk for myocardial infarction. Mutation also has been shown to carry high risk for early stroke development. Individuals with factor II 20210A show an increased level of plasma prothrombin levels; however, clinical utility is limited because of wide variability in the normal range of prothrombin assay. Thus, the DNA-based studies currently are utilized for determination of mutation. Testing for these mutations usually is done by PCR amplification of DNA isolated from peripheral blood mononuclear cells and using restriction enzyme targeted to the DNA sequences containing the mutation (Fig. 11.16).

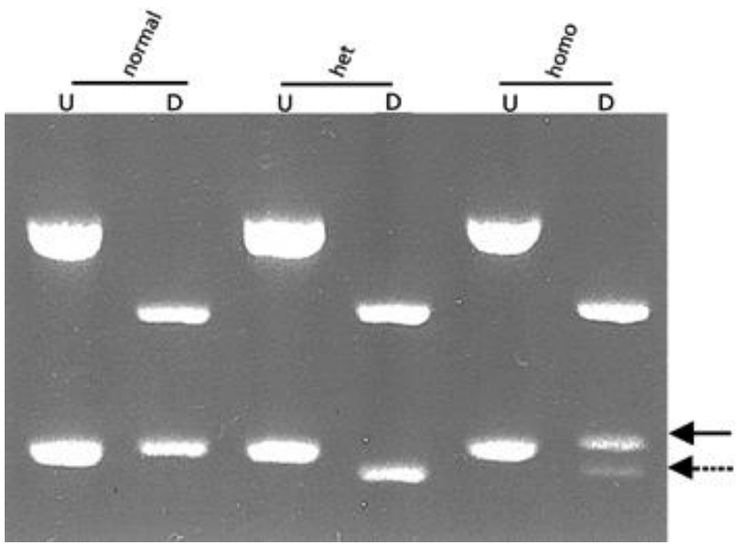


FIGURE 11.16. Detection of prothrombin gene G to A transition (20210). 3′ untranslated region of prothrombin gene, containing a G to A transition at nucleotide position 20210, was amplified and digested with HindIII restriction enzyme. Band of 345 bp (dashed arrow) after *Hind*III digestion shows a same size as undigested fragment while in homozygous (Hom) patient, digestion creates a faster moving band (322bp) (solid arrow). Heterozygous (Het) individual shows both 345 and 322 bp fragments. Most upper band is internal control for *Hind*III digestion to assure that restriction enzyme is functional. (Abbreviations: U = undigested; D = digested).

Von Willebrand Disease

Von Willebrand disease (vWD) is the most common inherited bleeding disorder. Several distinct clinical subtypes have been described with a high degree of heterogeneity and divided into two major categories depending on the presence of quantitative or qualitative defects. The quantitative deficiencies in plasma Von Willebrand factor (vWF) include types 1 and 3 vWD and qualitative abnormalities affecting vWF functions include type 2 vWD. The diagnosis usually is based on clinical findings and measurements of plasma and platelet vWF, the ability of vWF to interact with its platelet receptor, and the analysis of the multimeric composition of vWF. Because of the heterogeneity of vWF defects, a correct diagnosis of types and subtypes sometimes may be difficult but is very important for an appropriate treatment of patients with vWD.

Von Willebrand factor (vWF) is a high-molecular-weight (HMW) multimeric glycoprotein with two main functions in hemostasis: (i) stabilization and protection of factor VIII, and (ii) binding of platelets to the subendothelium at sites of vascular injury. It is synthesized in endothelial cells and megakaryocytes by a mRNA resulting from a 180-kb gene of 52 exons. There also is a partial pseudogene (exons 23–34) present which interferes with the analysis of the vWF gene. vWF is synthesized from a 9-kb mRNA exclusively in vascular endothelial cells and megakaryocytes. The prepropeptide (763 aa) and the mature

subunit (2050 aa) of vWF are composed of four repeating regions designated A, B, C, and D that are arranged as D1-D2-D′-D3-A1-A2-A3-D4-B1-B2-B3-C1-C2. Before its release, the VWF protein undergoes posttranslational modifications. It circulates, in plasma, as HMW multimers, which range in size from 500 to more than 100,000 kd. Molecular studies demonstrated that vWD shows an extensive genetic heterogeneity. Therefore, the role of molecular assays is limited in the diagnosis of this disorder.

Whereas various abnormalities (total, partial, or point deletions, point insertions, nonsense mutations) already have been identified in type 3, the molecular basis of type 1, most common type, is still unresolved in most cases.

Type 2A is the most common qualitative form of the vWD. Several causative missense mutations associated with type 2A have been identified, most of them clustered in the A2 domain (exon 28)(131). These patients show heterozygous missense mutations involving the A2 domain and they can be diagnosed by PCR (132). A novel mutation Gly 1671 Arg in type 2A and a homozygous mutation in type 2B von Willebrand disease also was described.

REFERENCES

1. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998;4:844–847.

2. Wilgenbus KK, Lichter P. DNA chip technology ante portas. *J Mol Med* 1999;77:761–768.

3. Grawunder U, West RB, Lieber MR. Antigen receptor gene rearrangement. *Curr Opin Immunol* 1998;10:172–180.

4. van Dongen JJ, Wolvers-Tettero IL. Analysis of immunoglobulin and T cell receptor genes. Part I: Basic and technical aspects. *Clin Chim Acta* 1991;198:1–91.

5. Blom B, Verschuren MCM, Heemskerk MHM, et al. TCR gene rearrangements and expression of the pre-T cell receptor complex during human T-cell differentiation. *Blood* 1999;93:3033–3043.

6. Davy MP, Bongiovanni KS, Kalsersch W, et al. Immunoglobulin and T-cell receptors gene rearrangement and expression in human lymphoid leukemia cells at different stages of maturation. *Proc Natl Acad Sci USA* 1986;83:8759–8763.

7. Greaves MS, Hurley AJW, al. e. Inappropriate rearrangement of immunoglobulin and T-cell receptor genes. *Immunol Today* 1987;8:115–116.

8. Biisaqizen A. Multiple rearrangement immunoglobulin genes in childhood acute lymphoblastic leukemia of B-cell origin. *Leukemia* 1991;5:657–667.

9. Scheller U, Muche JM, Sterry W, et al. Detection of clonal T cells in cutaneous T cell lymphoma by polymerase chain reaction: comparison of mutation detection enhancement-polyacrylamide gel electrophoresis, temperature gradient gel electrophoresis and fragment analysis of sequencing gels. *Electrophoresis* 1998;19:653–658.

10. Theodorou I, Bigorgne C, Delfau MH, et al. VJ rearrangements of the TCR gamma locus in peripheral T-cell lymphomas: analysis by polymerase chain reaction and denaturing gradient gel electrophoresis. *J Pathol* 1996;178:303–310.

11. Harris NL, Jaffe ES, Diebold J, et al. The World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues. Report of the Clinical Advisory Committee meeting, Airlie House, Virginia, November, 1997. *Ann Oncol* 1999;10:1419–1432.

12. Harris NL, Jaffe ES, Stein H, et al. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood* 1994;84:1361.

13. Pezzella F, Tse AG, Cordell JL, et al. Expression of the bcl-2 oncogene protein is not specific for the 14;18 chromosomal translocation. *Am J Pathol* 1990;137:225–232.

14. Onizuka T, Moriyama M, Yamochi T, et al. BCL-6 gene product, a 92- to 98-kD nuclear phosphoprotein, is highly expressed in germinal center B cells and their neoplastic counterparts. *Blood* 1995;86:28–37.

15. Auer RL, Bienz N, Neilson J, et al. The sequential analysis of trisomy 12 in B-cell chronic lymphocytic leukemia. *Br J Haematol* 1999;104:742–744.

16. Criel A, Michaux L, De Wolf-Peeters C. The concept of typical and atypical chronic lymphocytic leukaemia. *Leuk Lymphoma* 1999;33:33–45.

17. Geisler CH, Philip P, Christensen BE, et al. In B-cell chronic lymphocytic leukaemia chromosome 17 abnormalities and not trisomy 12 are the single most important cytogenetic abnormalities for the prognosis: a cytogenetic and immunophenotypic study of 480 unselected newly diagnosed patients. *Leuk Res* 1997;21:1011–1023.

18. Liu Y, Hermanson M, Grander D, et al. 13q deletions in lymphoid malignancies. *Blood* 1995;86:1911–1915.

19. Gahn B, Schafer C, Neef J, et al. Detection of trisomy 12 and rb-deletion in CD34+ cells of patients with B-cell chronic lymphocytic leukemia. *Blood* 1997;89:4275–4281.

20. Gahn B, Wendenburg B, Troff C, et al. Analysis of progenitor cell involvement in B-CLL by simultaneous immunophenotypic and genotypic analysis at the single cell level. *Br J Haematol* 1999;105:955–959.

21. Offit K, Parsa NZ, Filippa D, et al. t(9;14)(p13;q32) denotes a subset of low-grade non-Hodgkin's lymphoma with plasmacytoid differentiation. *Blood* 1992;80:2594–2599.

22. Busslinger M, Klix N, Pfeffer P, et al. Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc Natl Acad Sci USA* 1996;93:6129–6134.

23. Iida S, Rao PH, Nallasivam P, et al. The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood* 1996;88:4110–4117.

24. Adams B, Dorfler P, Aguzzi A, et al. Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev* 1992;6:1589–1607.

25. Neurath MF, Stuber ER, Strober W. BSAP: a key regulator of B-cell development and differentiation. *Immunol Today* 1995;16:564–569.

26. Pozzato G, Mazzaro C, Crovatto M, et al. Low-grade malignant lymphoma, hepatitis C virus infection, and mixed cryoglobulinemia. *Blood* 1994;84:3047.

27. Gorevic PD, Frangione B. Mixed cryoglobulinemia crossreactive idiotypes. Implications for the relationship of MC to rheumatic and lymphoproliferative disease. *Semin Hematol* 1991;28:79.

28. Bosch F, Jares P, Campo E, et al. PRAD-1/Cyclin D1 overexpression in chronic lymphoproliferative disorders: A highly specific marker of mantle cell lymphoma. *Blood* 1994;84:2726.

29. Ott MM, Helbing A, Ott G, et al. Bcl-1 gene rearrangement and cyclin D1 protein expression in mantle cell lymphoma. *J Pathol* 1996;179:238.

30. Rosenberg CL, Wong E, Petty EM, et al. PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma. *Proc Natl Acad Sci USA* 1991;88:9636.

31. Rimokh R, Berger F, Delsol G, et al. Rearrangement and overexpression of the Bcl-1/PRAD-1 gene in intermediate lymphocytic lymphomas and t(11q13)-bearing leukemias. *Blood* 1993;81:3063.

32. Siebert R, Matthiesen P, Harder S, et al. Application of interphase cytogenetics for the detection of t(11;14)(q13;q32) in mantle cell lymphomas. *Ann Oncol* 1998;9:519–526.

33. Vaandrager JW, Schuuring E, Zwikstra E, et al. Direct visualization of dispersed 11q13 chromosomal translocations in mantle cell lymphoma by multicolor DNA fiber fluorescence in situ hybridization. *Blood* 1996;88:1177–1182.

34. Cleary ML, Galili N, Sklar J. Detection of a second t(14;18) breakpoint cluster region in human follicular lymphomas. *J Exp Med* 1986;164:315.

35. Cleary ML, Sklar J. Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome. *Proc Natl Acad Sci USA* 1985;82:7439.

36. Lee MS. Molecular aspects of chromosomal translocation t(14;18). *Semin Hematol* 1993;30:297.

37. Lopez-Giullermo A, Cabanillas F, McDonnell TI, et al. Correlation of bcl-2 rearrangement with clinical characteristics and outcome in indolent follicular lymphoma. *Blood* 1999;93:3081–3087.

38. Gribben JG, Neuberg D, Freedman AS, et al. Detection by polymerase chain reaction of residual cells with the bcl-2 translocation is associated with increased risk of relapse after autologous bone marrow transplantation for B-cell lymphoma. *Blood* 1993;81:3449.

39. López-Guillermo A, Cabanillas F, McLaughlin P, et al. The clinical significance of molecular response in indolent follicular lymphomas. *Blood* 1998;91:2955–2960.

40. Limpens J, Stad R, Vos C, et al. Lymphoma-associated translocation t(14;18) in blood B-cells of normal individuals. *Blood* 1995;85:2528.

41. Liu Y, Hernandez AM, Shibata D, et al. Bcl-2 translocation frequency rises with age in humans. *Proc Natl Acad Sci USA* 1994;91:8910.

42. Shiramizu B, Barriga F, Neequaye J, et al. Patterns of chromosomal breakpoint locations in Burkitt's lymphoma: relevance to geography and Epstein-Barr virus association. *Blood* 1991;77:1516.

43. Goldschmidts WL, Bhatia K, Johnson JF, et al. Epstein-Barr virus genotypes in AIDS-associated lymphomas are similar to those in endemic Burkitt's lymphomas. *Leukemia* 1992;6:875.

44. Mosialos G, Birkenbach M, Yalamanchili R, et al. The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for the tumor necrosis factor receptor family. *Cell* 1995;80:389.

45. Chen M-L, Tsai C-N, Liang C-L, et al. Cloning and characterization of the latent membrane protein (LMP) of a specific Epstein-Barr virus variant derived from the nasopharyngeal carcinoma in the Taiwanese population. *Oncogene* 1992;7:2131.

46. Knecht H, Bachmann E, Brousset P, et al. Deletions within the LMP1 oncogene of Epstein-Barr virus are clustered in Hodgkin's disease and identical to those observed in nasopharyngeal carcinoma. *Blood* 1993;82:2937.

47. Luppi M, Barozzi P, Maiorana A, et al. Human herpesvirus-8 DNA sequences in human immunodeficiency virus-negative angioimmunoblastic lymphadenopathy and benign lymphadenopathy with giant germinal center hyperplasia and increased vascularity. *Blood* 1996;87:3903–3909.

48. Moore PS, Chang Y. Detection of herpesvirus-like sequences in Kaposi's sarcoma in patients with and those without HIV infection. *N Engl J Med* 1995;332:1182.

49. Dupin N, Diss TL, Kellam P, et al. HHV-8 is associated with a plasmablastic variant of Castleman disease that is linked to HHV-8-positive plasmablastic lymphoma. *Blood* 2000;95:1406–1412.

50. Poppema S, Kaleta J, Hepperle B. Chromosomal abnormalities in patients with Hodgkin's disease: evidence for frequent involvement of the 14q chromosomal region but infrequent bcl-2 gene rearrangement in Reed-Sternberg cells. *J Natl Cancer Inst* 1992;84:1789–1793.

51. Alkan S, Ross CW, Hanson CA, et al. Epstein-Barr virus and bcl-2 protein overexpression are not detected in the neoplastic cells of nodular lymphocyte predominance Hodgkin's disease. *Mod Pathol* 1995;8:544–547.

52. von Wasielewski R, Werner M, Fischer R, et al. Lymphocyte-predominant Hodgkin's disease. An immunohistochemical analysis of 208 reviewed Hodgkin's disease cases from the German Hodgkin Study Group. *Am J Pathol* 1997;150:793–803.

53. Kingma DW, Weiss WB, Jaffe ES, et al. Epstein-Barr virus latent membrane protein-1 oncogene delations: correlations with malignancy in Epstein-Barr virus-associated lymphoproliferative disorders and malignant lymphomas. *Blood* 1996;88:242–251.

54. Küppers R, Rajewsky K. The origin of Hodgkin and Reed Sternberg cells in Hodgkin's disease. *Annu Rev Immunol* 1998;16:471.

55. Marafioti T, Hummel M, Anagnostopoulos I, et al. Origin of nodular lymphocyte-predominant Hodgkin's disease from a clonal expansion of highly mutated germinal-center B cells. *N Engl J Med* 1997;337:453–458.

56. Ohno T, Stribley JA, Wu G, et al. Clonality in nodular lymphocyte-predominant Hodgkin's disease. *N Engl J Med* 1997;337:459–465.

57. Vockerodt M, Soares M, Kanzler H, et al. Detection of clonal Hodgkin and Reed-Sternberg cells with identical somatically mutated and rearranged VH genes in different biopsies in relapsed Hodgkin's disease. *Blood* 1998;92:2899–2907.

58. Krappmann D, Emmerich F, Kordes U, et al. Molecular mechanisms of constitutive NF-kappaB/Rel activation in Hodgkin/Reed-Sternberg cells. *Oncogene* 1999;18:943–953.

59. Emmerich F, Meiser M, Hummel M, et al. Overexpression of I kappa B alpha gene in Reed-Sternberg cells. *Blood* 1999;94:129–134.

60. Kuzel TM, Roenigk HH, Rosen ST. Mycosis fungoides and the Sezary syndrome: a review of pathogenesis, diagnosis, and therapy. *J Clin Oncol* 1991;9:1298–1313.

61. Delfau-Larue MH, Dalac S, Lepage E, et al. Prognostic significance of a polymerase chain reaction-detectable dominant T-lymphocyte clone in cutaneous lesions of patients with mycosis fungoides. *Blood* 1998;92:3376–3380.

62. Scheller U, Muche JM, Sterry W, et al. Detection of clonal T cells in cutaneous T cell lymphoma by polymerase chain reaction: comparison of mutation detection enhancement-polyacrylamide gel electrophoresis, temperature gradient gel electrophoresis and fragment analysis of sequencing gels. *Electrophoresis* 1998;19:653–658.

63. Hall WW, Liu CR, Schneewind O, et al. Deleted HTLV-1 provirus in blood and cutaneous lesions of patients with mycosis fungoides. *Science* 1991;253:317–320.

64. Zucker-Franklin D, Coutavas EE, Rush MG, et al. Detection of human T-lymphotropic virus like particles in cultures of peripheral blood lymphocytes from patients with mycosis fungoides. *Proc Natl Acad Sci USA* 1991;88:7630–7634.

65. Bazarbachi A, Soriano V, Pawson R, et al. Mycosis fungoides and Sezary syndrome are not associated with HTLV-1 infection: an international study. *Br J Haematol* 1997;98:927–933.

66. Shimoyama M. Diagnostic criteria and classification of clinical subtypes of adult T-cell leukemia-lymphoma: a report from the Lymphoma Study Group (1984-87). *Br J Haematol* 1991;79:428.

67. Seiki M, Eddy R, Shows TB, et al. Nonspecific integration of the HTLV provirus genome into adult T-cell leukaemia cells. *Nature* 1984;309:640.

68. Smith MR, Greene WC. Molecular biology of type 1 human T-cell leukemia virus (HTLV-1) and adult T-cell leukemia. *J Clin Invest* 1991;87:761.

69. Knowles DM, Cesarman E, Chadburn A, et al. Correlative morphologic and molecular genetic analysis demonstrates three distinct categories of posttransplantation lymphoproliferative disorders. *Blood* 1995;85:552.

70. Cleary ML, Nalesnik MA, Shearer WT, et al. Clonal analysis of transplant-associated lymphoproliferations based on the structure of the genomic termini of the Epstein-Barr virus. *Blood* 1988;72:349.

71. Locker J, Nalesnik M. Molecular genetic analysis of lymphoid tumors arising after organ transplantation. *Am J Pathol* 1989;135:977.

72. Cesarman E, Chadburn A, Liu Y-F, et al. BCL-6 gene mutations in posttransplantation lymphoproliferative disorders predict response to therapy and clinical outcome. *Blood* 1998;92:2294–2302.

73. Morris SW, Kirstein MN, Valentine MB, et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* 1994;263:1281–1284.

74. Rimokh R, Magaud JP, Berger F, et al. A translocation involving a specific breakpoint (q35) on chromosome 5 is characteristic of anaplastic large cell lymphoma ('Ki-1 lymphoma'). *Br J Haematol* 1989;71:31.

75. Borer RA, Lehner CF, Eppenberger HM, et al. Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* 1989;56:379–390.

76. Falini B, Bigerna B, Fizzotti M, et al. ALK expression defines a distinct group of T/null lymphomas ("ALK lymphomas") with a wide morphological spectrum. *Am J Pathol* 1998;153:875.

77. Gascoyne RD, Aoun P, Wu D, et al. Prognostic significance of anaplastic lymphoma kinase (ALK) protein expression in adults with anaplastic large cell lymphoma. *Blood* 1999;93:3913–3921.

78. Erickson P, Gao J, Chang K-S, et al. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* 1992;80:1825–1831.

79. Nucifora G, Dickenstein JI, Torbenson V, et al. Correlation between cell morphology and expression of the AML1/ETO chimeric transcript in patients with acute myeloid leukemia without the t(8;21). *Leukemia* 1994;8:1533–1538.

80. Kusec R, Laczika K, Knobl P, et al. AML1/ETO fusion mRNA can be detected in remission blood samples of all patients with t(8;21) acute myeloid leukemia after chemotherapy or autologous bone marrow transplantation. *Leukemia* 1994;8:735–739.

81. Grimwade D, Solomon E. Characterization of the PML/RARA rearrangement associated with t(15; 17) acute promyelocytic leukaemia. *Curr Topics Microbiol Immunol* 1997;220:81.

82. Lafage-Pochitaloff M, Alcalay M, Brunel V, et al. Acute promyelocytic leukemia cases with nonreciprocal PML/RAR or RAR/PML fusion genes. *Blood* 1995;85:1169.

83. Brown D, Kogan S, Lagasse E, et al. A PMLRAR transgene initiates murine acute promyelocytic leukemia. *Proc Natl Acad Sci USA* 1997;94:2551.

84. Licht JD, Chomienne C, Goy A, et al. Clinical and molecular characterization of a rare syndrome of acute promyelocytic leukemia associated with translocation. *Blood* 1995;85:1083.

85. Viswanatha D, Chen I-M, Liu PP, et al. Characterization and utility of an antibody detecting the CBFB-SMMHC fusion protein in inv(16)/t(16;16)-associated acute myeloid leukemias. *Blood* 1998;91:1882–1890.

86. Liu PP, Hajra A, Wijmenga C, et al. Molecular pathogenesis of the chromosome 16 inversion in the M4eo subtype of acute myeloid leukemia. *Blood* 1995;85:2289–2302.

87. Ben-Neriah Y, Daley QG, Mes-Mason AM, et al. The chronic myelogenous leukemia-specific p210 protein is the product of the bcr-abl hybrid gene. *Science* 1986;133:212.

88. Faderl S, Talpaz M, Estrov Z, et al. The biology of chronic myeloid leukemia. *N Engl J Med* 1999;341:164–172.

89. Branford S, Hughes TP, Rudzki Z. Monitoring chronic myeloid leukaemia therapy by real-time quantitative PCR in blood is a reliable alternative to bone marrow cytogenetics. *Br J Haematol* 1999;107:587–599.

90. Hochaus A, Lin F, Reiter A, et al. Variable numbers of BCR-ABL transcripts persist in CML patients who achieve complete cytogenetic remission with interferon. *Br J Haematol* 1995;91:126.

91. Miyamura K, Tahara T, Tanimoto M, et al. Long persistent bcr-abl positive transcript detected by polymerase chain reaction after marrow transplant for chronic myelogenous leukemia without clinical relapse: A study of 64 patients. *Blood* 1993;81:1089.

92. Hughes TP, Morgan GJ, Martiat P, et al. Detection of residual leukemia after bone marrow transplant for chronic myeloid leukemia: Role of polymerase chain reaction in predicting relapse. *Blood* 1991;77:874.

93. Radich JP, Gehly G, Gooley T, et al. Polymerase chain reaction detection of the *BCR-ABL* fusion transcript after allogeneic marrow transplantation for chronic myeloid leukemia: Results and implications in 346 patients. *Blood* 1995;85:2632.

94. Soekerman D, van Denderen J, Hoefsloot L, et al. A novel variant of the bcr-abl fusion product in Philadelphia chromosome-positive acute lymphoblastic leukemia. *Leukemia* 1990;4:397.

95. Kurzrock R, Shtalrid M, Romero P, et al. A novel c-abl protein product in Philadelphia-positive acute lymphoblastic leukemia. *Nature* 1987;325:631.

96. Secker-Walker LM, Craig JM. Prognostic implications of breakpoint and lineage heterogeneity in Philadelphia-positive acute lymphoblastic leukemia: a review. *Leukemia* 1993;7:147.

97. Kantarjian HM, Talpaz M, Dhingra K, et al. Significance of the p210 versus p190 molecular abnormalities in adults with Philadelphia chromosome-positive acute leukemia. *Blood* 1991;78:2411.

98. Van Rhee F, Hochhaus A, Lin F, et al. p190 bcr-abl mRNA is expressed at low levels in p210-positive chronic myeloid and acute lymphoblastic leukemias. *Blood* 1996;87:5213.

99. Rubnitz JE, Behm FG, Downing JR. 11q23 rearrangements in acute leukemia. *Leukemia* 1996;10:74.

100. Biondi A, Rambaldi A, Rossi V, et al. Detection of ALL-1/AF4 fusion transcript by reverse transcription-polymerase chain reaction for diagnosis and monitoring of acute leukemias with the t(4;11) translocation. *Blood* 1993;82:2943–2947.

101. Léglise MC, Rivière D, Brière J. Acute leukemia with a translocation t(4;11)(q21;q23): a distinct clinicopathological entity. *Leuk Lymphoma* 1990;2:353.

102. Pui CH. Acute leukemias with the t(4;11)(q21:q23). *Leuk Lymphoma* 1992;7:173.

103. Bernard OA, Berger R. Molecular basis of 11q23 rearrangements in hematopoietic malignant proliferations. *Genes Chromosomes Cancer* 1995;13:75.

104. Thirman MJ, Gill HJ, Burnett RC, et al. Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. *N Engl J Med* 1993;329:909.

105. Pui C-H, Raimondi SC, Hancock ML, et al. Immunologic, cytogenetic, and clinical characterization of childhood acute lymphoblastic leukemia with the t(1;19)(q23;p13) or its derivative. *J Clin Oncol* 1994;12:2601.

106. Uckun FM, Sensel MG, Sather HN, et al. Clinical significance of translocation t(1;19) in childhood acute lymphoblastic leukemia in the context of contemporary therapies: a report from the Children's Cancer Group. *J Clin Oncol* 1998;16:527–535.

107. Hungler SP, Fall MZ, Camitta BM, et al. E2A-PBX1 chimeric transcript status at end of consolidation is not predictive of treatment outcome in childhood acute lymphoblastic leukemias with a t(1;19)(q23;p13): a Pediatric Oncology Group study. *Blood* 1998;91:1021–1028.

108. Raynaud S, Cave H, Baens M, et al. The 12;21 translocation involving TEL and deletion of the other TEL allele: Two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* 1996;87:2891.

109. Romana SP, Poirel H, Le Coniat M, et al. High frequenty of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. *Blood* 1995;86:4263.

110. Shurtleff SA, Bujis A, Behm FG, et al. TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup with an excellent prognosis. *Leukemia* 1995;9:1985.

111. McLean T, Ringold S, Neuberg D, et al. TEL/AML-1 dimerizes and is associated with a favorable outcome in childhood acute lymphoblastic leukemia. *Blood* 1996;88:4252.

112. Rubnitz JE, Downing JR, Pui C-H, et al. TEL gene rearrangement in acute lymphoblastic leukemia: A new genetic marker with prognostic significance. *J Clin Oncol* 1997;15:1150.

113. Wada M, Bartram CR, Nakamura H, et al. Analysis of p53 mutations in a large series of lymphoid hematologic malignancies of childhood. *Blood* 1993;82:3163.

114. Auerbach AD. Fanconi anemia diagnosis and the diepoxybutane (DEB) test. *Exp Hematol* 1993;21:731.

115. Dokal I, Chase A, Morgan NV, et al. Positive diepoxybutane test in only one of two brothers found to be compound heterozygotes for Fanconi's anemia complementation group C mutations. *Br J Haematol* 1996;93:813.

116. Joenje H, Oostra AB, Wijker M, et al. Evidence for at least eight Fanconi anemia genes. *Am J Hum Genet* 1997;61:940–944.

117. Whitney MA, Saito H, Jakobs PM, et al. A common mutation in the FACC gene causes Fanconi anemia in Ashkenazi Jews. *Nat Genet* 1993;4:202.

118. Verlander PC, Lin JD, Udono MU, et al. Mutation analysis of the Fanconi anemia gene FACC. *Am J Hum Genet* 1994;54:595.

119. Edwards CQ, Griffen LM, Goldgar D, et al. Prevalence of hemochromatosis among 11065 presumably healthy blood donors. *N Engl J Med* 1988;318:1355.

120. Leggett BA, Halliday JW, Brown NN, et al. Prevalence of haemochromatosis among asymptomatic Australians. *Br J Haematol* 1990;74:525.

121. Mura C, Raguenes O. HFE mutations analysis in 711 hemochromatosis probands: evidence for S65C implication in mild form of hemochromatosis. *Blood* 1999;93:2502–2505.

122. Bertina RM, Koeleman RPC, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994;369:64.

123. Greengard JS, Sun X, Xu X, et al. Activated protein C resistance caused by Arg506Gln mutation in factor Va. *Lancet* 1994;343:1362.

124. Rosendaal FR, Koster T, Vandenbroucke JP, et al. High risk of thrombosis in patients homozygous for factor V Leiden (activated protein C resistance). *Blood* 1995;85:1504–1508.

125. Ridker PM, Hennekens CH, Lindpaintner K, et al. Mutation in the gene coding for coagulation factor V and the risk of myocardial infarction, stroke, and venous thrombosis in apparently healthy men. *N Engl J Med* 1995;332:912–917.

126. Amowitz LL, Komaroff AL, Miletich JP, et al. Factor V Leiden is not a risk factor for myocardial infarction among young women. *Blood* 1999;93:1432-1433.

127. Williamson D, Brown K, Luddington R, et al. Factor V Cambridge: A new mutation (Arg306Thr) associated with resistance to activated Protein C. *Blood* 1998;91:1140–1144.

128. Poort SR, Rosendaal FR, Reitsma PH, et al. A common genetic variation in the 3′-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 1996;88:3698–3703.

129. Hillarp A, Zoller B, Svensson PJ, et al. The 20210 A allele of the prothrombin gene is a common risk factor among Swedish outpatients with verified deep venous thrombosis. *Thromb Haemost* 1997;78:990–992.

130. Makris M, Preston FE, Beauchamp NJ, et al. Co-inheritance of the 20210A allele of the prothrombin gene increases the risk of thrombosis in subjects with familial thrombophilia. *Throm Haemost* 1997;1426–1429.

131. Hagiwara t, Inaba H, Yoshida S, et al. A novel mutation Gly 1671 Arg in type 2A and a homozygous mutation in type 2B von Willebrand disease. *Thromb Haemost* 1996;76:253–257.

132. Lyons SE, Bruck ME, Bowie EJW, et al. Impaired intracellular transport produced by a subset of type IIA von Willebrand disease mutations. *J Biol Chem* 1992;267:4424–4430.

Section 3 Clinical Chemistry

Clinical Chemistry - Introduction

Kent Lewandrowski

Section Chief

Elizabeth Lee-Lewandrowski

Section Chief

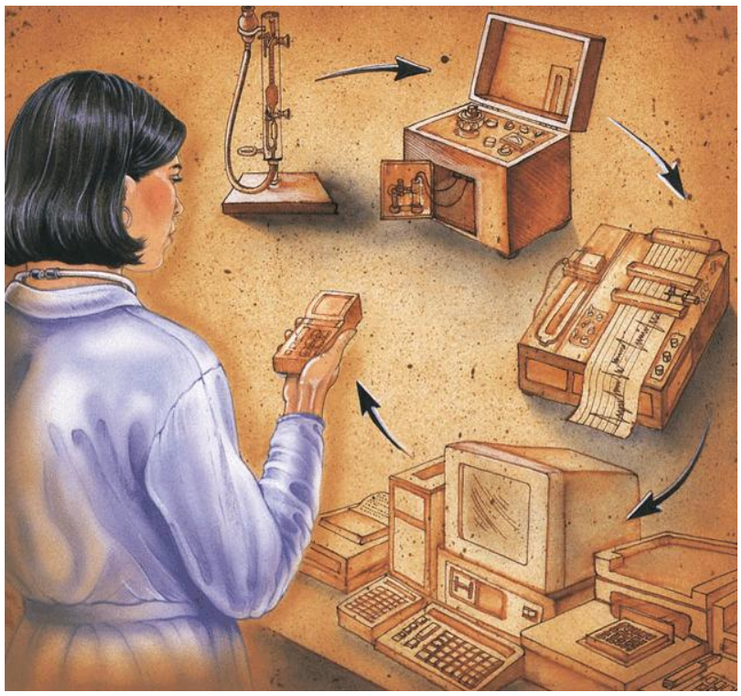


Figure.

Most illnesses either have a primary biochemical origin or result in a secondary disturbance to the intricate biochemical equilibria that characterize the healthy human organism. It follows that careful measurement of selected analytes in blood or other body fluids can provide information useful in diagnosis, and such measurements also are useful in monitoring disease progression and the response to therapy.

Although clinical chemistry sometimes is viewed as a narrow subspecialty, it is in fact a very broad field encompassing a large variety of quantitative and qualitative analytical techniques. Analytes range from simple ionic species and small organic molecules to protein, hormones, and drugs. The range of concentration of the various species of interest spans more than nine orders of magnitude. Methodologies range from difficult and labor-intensive manual methods to highly automated instrument methods.

This section is intended to cover the most important analytes of interest in clinical chemistry. When applicable, chapters include normal biochemistry and physiology, an outline of laboratory techniques for measuring the analyte, and the clinical relevance of the laboratory results. It is hoped that this section will serve as a reference for the most commonly used information about clinical chemistry, as well as an introduction to the more detailed discussion found in specialized textbooks and journal articles.

12
Immunochemical Methods

Robert E. Moore

Immunoassay, as a diagnostic tool, originates with the much older clinical discipline of serology. Approximately 100 years ago, it was demonstrated that certain types of bacteria could be agglutinated by antiserum. It was reasoned that this could be an aid in the diagnosis of typhoid fever, and the procedure was named the Grüber-Widal test. From this modest beginning, immunoassay developed into the fastest-growing methodology, at that time, in the clinical laboratory. The transition, however, was not easy. Numerous advances have allowed immunoassay to achieve its present popularity. The first of these came in the late 1950s when Drs. Berson and Yalow published a series of studies on ^{131}I-labeled insulin to detect antibodies to insulin and later to measure the insulin itself. These studies demonstrated beyond any doubt that a new analytical technique with exquisite sensitivity and specificity was possible. The principle was proven but there were practical problems that kept the technique from being adopted for general use. First, the technique was a radioactive procedure and the additional training, equipment, and regulatory issues were deterrents. Second, antibodies primarily were responsible for the specificity and to a lesser extent the sensitivity of the assay. These reagent antibodies were not commonly available, and investigators had to produce and characterize their own antibodies for each analyte that was measured. Antibody production was not a technique common to the clinical laboratory, and laboratories that were capable of undertaking antibody production found it to be a labor-intensive effort. Third, these early assays were based on competitive binding, a technique that depends on a labeled antigen competing equivalently with the natural antigen for a fixed number of sites, which meant most antigens had to be labeled in the laboratory. The short half-life of ^{131}I, and the difficulty in controlling the organic reactions required to attach this radioactive atom to the analyte, made this a very inefficient and expensive procedure. The instability of the labeled compounds required this labeling procedure to be done just before performing the analysis. Generally, these radioimmunoassay (RIA) methods were long and tedious and if the results were unacceptable, the process had to be repeated, including the labeling procedure.

It was apparent that the analytical potential of the immunoassay with a radioactive label was virtually unlimited and had allowed the quantitation of hormones and peptides for the first time. It also was evident that this potential would be sufficient to stimulate commercial vendors to begin supplying assay materials in a more convenient form, thereby giving smaller laboratories access to this technology. By the late 1960s and early 1970s, commercial preparation of antibodies was common. Investigators could pick from an ever-increasing menu of prepared antibodies and prelabeled antigens. It was a small step from antibody and antigen availability to the appearance of kit packaging. These kits contained all the reagents that the researcher would need to perform a specific assay, and this made immunoassays available to significantly more laboratories. With this proliferation of immunoassay procedures, the users demanded higher quality and standardized reagents.

A parallel development occurred with small organic molecules such as steroids, thyroid hormones, and drugs. These molecules had been synthesized for some time with tritium or ^{14}C atoms at specific structural sites. Antibodies for these materials were not available. One problem with this type of molecule is that it is not immunogenic by itself. That is, these compounds will bind with antibodies but are not capable of evoking an immunologic response in an animal to produce an antibody. This was an additional problem that was not encountered with the larger peptide hormones or proteins. The solution to this problem was to **attach, through covalent linkages**, these small molecules to much larger immunogenic molecules. As investigators began to couple these haptens to proteins, it became apparent that the stereochemistry of the coupling reactions was extremely important. Because there is a great deal of homology among steroids, drugs, and thyroid hormones, judicious choice of coupling site was required to improve the specificity of antibodies. As with the peptide hormones, it soon became apparent that the commercial manufacturers would take a leading role in supplying antibodies and prepackaged reagent systems for these assays.

As more investigators became interested in the RIA technique, it was inevitable that some effort would be invested in finding alternatives to radioactive methods. Because the specificity of these reactions is a result of the antibody, and the sensitivity is mainly because of the amplification system (i.e., multiple radioactive atoms per molecule of analyte), any chemical method that had amplification potential became a candidate to replace the radioactive label. The only limitation seems to be the imagination as multiple types of labels are combined with detection systems to produce the wide spectrum of immunoassays currently available.

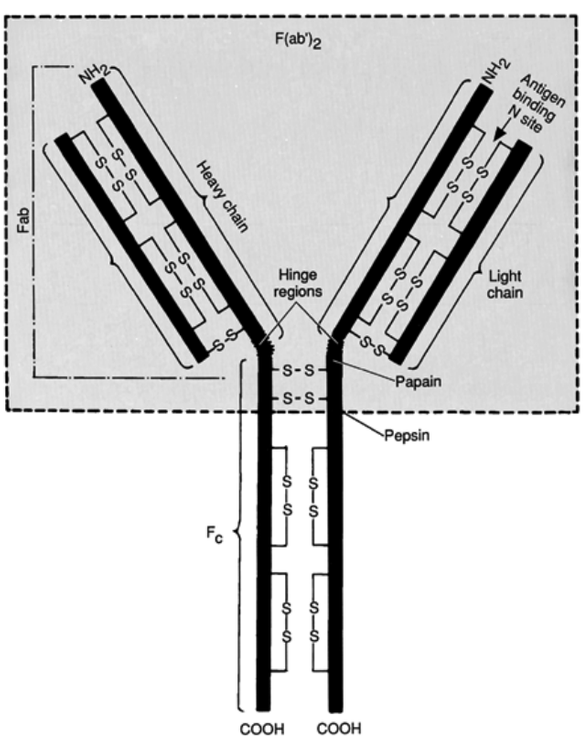


FIGURE 12.1. Schematic representation of IgG antibody. Papain digestion releases two Fab fragments and one Fc fragment. Pepsin digestion releases one $F(ab')_2$ fragment. Both enzyme digestions conserve antigen-binding sites.

All immunoassays share one common feature. They use one or more antibodies to effect the analytic measurement. They owe their specificity and to some extent their sensitivity to the quality of the antibody used in the assay.

- ANTIGENS
- IMMUNOGENS
- ANTIBODIES
- PRECIPITATION TECHNIQUES
- GEL SYSTEMS
- ELECTROPHORESIS
- INSTRUMENTAL METHODS
- HETEROGENEOUS AND HOMOGENEOUS ASSAY SYSTEMS
- SPECIFIC ASSAY TYPES

ANTIGENS

Part of "12 - Immunochemical Methods"

To try to define antigen in a chapter on immunoassay is problematic. Immunologists and practitioners of the immunologic sciences have struggled, with limited success, to provide a precise and all-inclusive meaning for antigen. For purposes of this discussion, the following functional definition will be used: antigens are substances capable of binding to an antibody. This is in contrast to immunogens, which are substances capable of eliciting a humoral antibody response when presented to a host. The definitions in this chapter are not refined, but they are attempts to minimize confusion.

IMMUNOGENS

Part of "12 - Immunochemical Methods"

Immunogens have several qualities that have been determined empirically. Good immunogens usually are large, complex molecules that differ significantly from normal molecular species found in the host. Although the term large is difficult to describe, molecules with molecular weights greater than 10,000 are usually good immunogens, while, conversely, molecules with molecular weights below 10,000 are poorer.

Molecular size is not the only requirement. Structure is a significant characteristic in determining the efficacy of the immunogen. If the overall structure has incorporated in it lipids, glycoproteins, polysaccharides, or nucleic acids, these will generate a better antibody response in the host than will synthetic peptides composed of one or two simple amino acids. In addition to primary structure, there also is a relationship between immunogenicity and tertiary structure. The more complex spatial relationships found in biomolecules tend to be more immunogenic than less-complex synthetic linear peptides and proteins.

The quality of the immune response of the host improves dramatically as the host is able to identify the immunogen as a foreign substance. When the host animal is challenged, there is an attempt to categorize the material as self or nonself. For obvious reasons, materials that are identified as nonself elicit a much more vigorous response. If the origin of the nonself material is from an evolutionary divergent species, the immune response is expected to be far greater than if the nonself is from a closely related or same species.

Finally, not all immunogens are equal in their ability to elicit antibody formation. The immunogen can be administered as a simple solution or it can be suspended with an adjuvant that will help in amplifying the host response. Administration of the immunogen usually is by injection and the site is one of personal preference. In any case, the host should not become incapacitated (footpad, or splenic injections) because of the administration. Convenient and effective routes commonly are employed.

For biomolecules of clinical interest that are not immunogenic in and of themselves (e.g., steroids, thyroid hormones, drugs, and small peptides), they can be made effective immunogens by covalently coupling them to a much larger protein backbone. Multiple moles of a small molecule can be covalently linked per mole of large protein to produce a very effective immunogen. Molecules that require coupling to other molecules to produce immunogenicity are referred to as haptens. Because many of these small molecules share similar structure, as in the case of steroids, it is possible to select the organic coupling reaction so that the common portion of the structure is used for attachment to the protein backbone, while the unique portion of the structure is left exposed to the environment. This process increases the probability that antibodies will be generated against this unique structure.

ANTIBODIES

Part of "12 - Immunochemical Methods"

Once the immunogen has been presented to an appropriate host, a type of lymphocyte called a “B cell” will begin to produce antibodies against the immunogen. These antibodies are immunoglobulin class proteins predominantly of the IgG and IgM type, with the IgG making up most of the analytic antibodies in use today. As shown in Figure 12.1, IgG is composed of two identical heavy and two identical light chains held together by a series of interdisulfide and intradisulfide bonds. The heavy chains are designated (H) and the light chains (L). The specific heavy chain associated with IgG is designated as a γ chain, while the light chains are either κ or λ, but both do not occur on the same molecule. There are two binding sites on each IgG molecule, and these are located at the ends of the “arms” at the juncture of the heavy and light chains.

Enzymatic digestion of the IgG molecule yields some interesting structural characteristics. Pepsin digestion cleaves the two arms of the “Y” from the body of the molecule, and these arms remain attached to each other and retain the divalent antibody-binding characteristics. The body of the IgG is digested into small fragments. If papain is used to digest the IgG molecule, three fragments are obtained. The two arms are cleaved as independent fragments, each with its antibody-binding characteristic, and the third fragment is the body of the IgG molecule, which remains intact. When the two antibody fragments remain together, as in the case of pepsin digestion, it is referred to as a Fab'_2 fragment, indicating that the two antibody fragments remain together. In the case of papain, the antibody fragments are designated as Fab, because they have been cleaved into single binding sites by the enzyme. The third fragment, the body of the IgG, is the Fc fragment. The c here stands for crystallizable, referring to a characteristic of that fragment after papain digestion. The actual binding site on the IgG molecule is at the amino end, with one binding site in each of the arms.

It is possible to generate analytic antibodies in one of two ways. The historical method is to challenge an animal species with an immunogen to elicit an immune response. Each B cell or lymphocyte that is stimulated to produce antibody to the immunogen ultimately will produce one antibody with specific binding characteristics. This “clonal theory” was introduced in 1957 by Burnet (1). The large number of lymphocytes stimulated each with a single antibody response, gives rise to a broad

spectrum of antibodies that recognize different antigenic determinants over the entire structure of the immunogen. These antibodies differ in quantity and quality as to their ability to identify the specific immunogen presented, but with some luck and considerable effort in purifying and characterizing these antibodies, the result can yield very high-quality reagents for *in vitro* use. These polyclonal antibodies have been used for years in analytical methods.

The second and more recent approach is to generate monoclonal antibodies. Kühler and Milstein exploited the fact that each lymphocyte clone would produce only one type antibody. They developed a technique for continuous production of a selected antibody from cultured lymphocytes (2). This method produces very specific antibodies that recognize a single antigenic determinant on a given molecule, and the clone can be cultured virtually forever. The procedure requires inoculating mice, taking the spleen cells from responsive animals, and fusing them with a myeloma cell line. The fusion products are cultured and screened for the clone that generates the desired antibody. In theory, this approach should have produced the ideal analytic antibody. Practice has demonstrated that monoclonal antibodies (MAbs) have some deficiencies. The affinity of MAbs often is less than that required for analytic procedures resulting in low limits of detection, and the specificity for small molecules is difficult to achieve (3). Some of these drawbacks can be overcome by combining multiple clones from a given fusion, in essence producing a polyclonal serum from several MAb sources, but this is not effective for every application. For large analytes, the combination of MAbs with differing affinities for different epitopes has improved the overall avidity. This effect has not been observed with haptens (4).

Antigen-Antibody Interactions

Immunoassays can be designed to measure either the formation of an antigen-antibody complex or the inhibition of that complex

formation. Individual investigators and commercial suppliers design immunoassays to take advantage of a particular detection system or to maximize analytical sensitivity, if that is critical. It is necessary, therefore, to understand the basic qualities of the antigen-antibody complex that yield good analytic assays. As equations 12.1 and 12.2 demonstrate, the antigen-antibody reaction is a reversible reaction that can be rearranged in conventional form to yield an expression for the association constant, K_a, in terms of the reactant and product concentrations.

$$A_g + A_b \underset{k_2}{\overset{k_1}{\rightleftharpoons}} A_g \cdot A_b \qquad (12.1)$$

where

- A_g = antigen
- A_b = antibody
- $A_g \cdot A_b$ = antigen-antibody complex
- k_1 = forward reaction rate
- k_2 = reverse reaction rate

$$\frac{[A_g \cdot A_b]}{[A_g]\,[A_b]} = k_1 / k_2 = K_a \text{ association constant} \qquad (12.2A)$$

$$\frac{[A_g]\,[A_b]}{[A_g \cdot A_b]} = k_2 / k_1 = K_d \text{ dissociation constant} \qquad (12.2B)$$

Good antibodies, and consequently good assays, are characterized by large K_a values. The interpretation of this is that the equilibrium reaction lies far to the right and there is minimal dissociation of the antigen-antibody complex once it has formed. A small dissociation constant (K_d) is interpreted to mean the antibody is bound very tightly to the specific antigen. In other words, the fit between the antigen and the antibody is optimal and the binding forces have been maximized; therefore, the dissociation rate constant, k_2, is minimized. The practical implication of this is that assays designed around antigen-antibody complexes of this quality are called “robust” and tend to be less sensitive to minor changes in assay conditions such as pH, reaction time, temperature, and ionic strength. If the K_a were small, then the dissociation of the complex would become appreciable and the assay would be insensitive because of a low concentration of complexes at any given time.

The forces that determine the antigen-antibody complex all are intermolecular forces with the exception of covalent bonding. Ionic bonds, hydrogen bonds, van der Walls' forces, and hydrophobic-hydrophilic repulsion constitute the forces that are responsible for maintaining the integrity of the complex. These forces act over short molecular distances. As the fit between the antigen and the antibody improves, more and more of these noncovalent bonds are formed and the rate of dissociation of the complex is reduced. This property is called antibody affinity for a single site, for complexes that are formed using multiple antibodies on multiple antigenic sites; the sum of these forces is referred to as antibody avidity. It is an interesting property of these bonds that the total adhesive force in the complex exceeds the sum of the individual forces (5).

The stoichiometry of the antigen-antibody reaction plays a critical role in assays that are dependent on maximum precipitation. As Figure 12.2 illustrates, when a polyvalent antibody reacts with an antigen with multiple determinants, the precipitation increases gradually to a maximum and then begins to decrease. Initially, there is excess antibody for the antigen in the solution. In this situation, each antigen molecule has the maximum number of antibodies bound, but a different antibody binds each epitope on the antigen. With excess antibody available, there is little opportunity for the cross-linking of antigens and the subsequent formation of large molecular complexes. As the concentration of antigen increases, a point is reached where the antigen and antibody concentrations are balanced so that a macromolecular lattice can be formed. These lattice structures are so large that precipitation is at a maximum. The point at which maximum precipitation occurs is called the equivalence point. At this point, if one were to analyze the solution after precipitation of the antigen-antibody complex, there would be very little antigen or antibody remaining. If antigen continues to increase to a concentration where it is in excess, then each antibody becomes saturated with antigen molecules. Using IgG as an example, two antigen molecules are bound per antibody molecule. Similar to the case of antibody excess, antigen excess results in different antigens binding to the antibody, again limiting the possibility for cross-linking. The complexes formed under conditions of antigen excess are not large enough to affect total precipitation, and so considerable complex is left in solution. Note also that the *x*-axis is calibrated in concentration and the y-axis calibrated for a response to complex formation. The resulting curve, due to its symmetry, has a low and a high concentration that will yield the same response. Assays that depend on measuring complex formation, particularly if they are fully automated assays, must consider this effect and be prepared to test for antigen excess conditions.

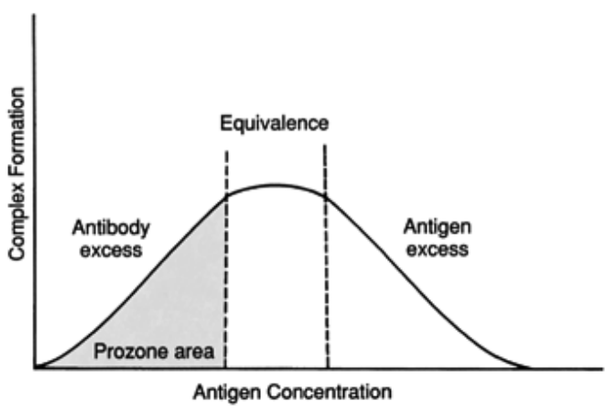


FIGURE 12.2. Titration curve for a constant quantity of antibody. In antibody excess (prozone region) there are insufficient antigen molecules to produce maximum complex formation. At equivalence, the complex formation is optimized. If the complexes are separated from the liquid phase, no free antigen or antibody is detected in the remaining solution. In antigen excess, each antibody is saturated with antigen, and complex formation is reduced.

PRECIPITATION TECHNIQUES

Immunodiffusion: Liquid Systems

The formation of precipitates between antigens and antibodies in liquid media are the simplest and oldest precipitation reactions conducted in the laboratory. This approach takes advantage of the diffusive property of molecules and the precipitation of antigen-antibody complexes. The reaction is carried out in small test tubes, small-diameter capillary tubes, or microtiter wells. A mixture of antigen and antibody is allowed to incubate for several hours. At the end of this time a precipitant will have formed and settled to the bottom of the reaction vessel. If this reaction is carried out with a constant amount of antigen and varying dilutions of antibody, a qualitative assessment of the equivalence point can be determined by evaluating the amount of precipitate that has formed in each of the tubes. The equivalence point will be equal to the concentration that formed the greatest amount of precipitant.

Another way to perform this reaction is to add the antibody solution to the tube and then very carefully layer the antigen solution on top, trying to avoid actual mixing of the solutions. As the antigen-antibody complex is formed at the interface, a ring will be visible.

In each of the aforementioned cases, the antigen and antibody were soluble and multivalent so that a macromolecular species could form that would eventually precipitate out of solution. Techniques have been developed for small antigens that do not have multiple binding sites and therefore tend to form complexes that remain soluble. One approach is to coat larger particles with the antigen before exposing it to the specific antibody. Typical supports are sheep erythrocytes or latex particles coated with the antigen of interest. Once these large support particles contain many antigens, multiple antibodies will bind and the particles will become aggregated in solution. The term aggregation generally applies to reactions in which the antigen or antibody is coated on the surface of a cell or synthetic particle used as a support medium. In this reaction, the aggregation of particles indicates the presence of the analyte. When these reactions are carried out in a series of increasing dilutions, a qualitative assessment of the concentration of analyte can be determined.

A variation on the aggregation reaction is the aggregation inhibition reaction. Like the aggregation reaction, antigens or antibodies are coated on a support such as animal erythrocytes or latex beads. The first step of the reaction is to incubate the solution to be analyzed with free antibody. This reaction proceeds to form soluble antigen-antibody complexes. At the end of this initial incubation period, the support particles with the antigens coated on the surface are incubated with the preformed soluble antigen-antibody complexes. The quantity of antigen in the test solution will determine the number of free antibodies available for agglutinating the support particles. In this case, the concentration of analyte is inversely proportional to the amount of agglutination. If a series of standard solutions is analyzed concurrently, the concentration of the antigen in the unknown can be estimated. Both the agglutination and the agglutination inhibition assays have been used in the past as slide tests for screening procedures to detect infectious diseases, pregnancy, and several red cell surface antigens.

All of the reactions described so far in liquid medium require a precalibration of the antigen and antibody solution so that precipitation or agglutination will take place. Most clinical specimens that are encountered in the laboratory fall within a relatively narrow range of concentrations. Assay optimization can thus be done for the anticipated concentrations. For the few specimens that have concentrations outside the expected range, concentration or dilution can adjust those samples so the analytes will be within the analytical range of the assay. It is important to check the analytic reagents periodically, especially as new lots of reagent are incorporated into the procedure, to confirm the analytic range.

GEL SYSTEMS

Part of "12 - Immunochemical Methods"

These reactions are carried out in a gel support, usually an agar or agarose gel, that has been heated to liquid state and then poured onto a solid support and allowed to solidify. This solid support can be glass, acetate sheets, plastic, or any appropriate transparent material. It is important that the seal between the gel and the solid support is intact to prevent the solutions from wicking along the interface. As shown in Figure 12.3, two wells are cut into the gel. An antigen solution is placed in one and an antibody solution in the other. The solutions are allowed to incubate at room temperature, during which time the solutions diffuse through the gel in a radial fashion. The rate of diffusion is dependent on several conditions: the concentration of the respective solutions, the molecular size of the antigens and antibody, and the lattice structure of the gel. As these solutions diffuse, their fronts will meet and at a point where the antigen and antibody are in equivalence, a precipitation line will be formed.

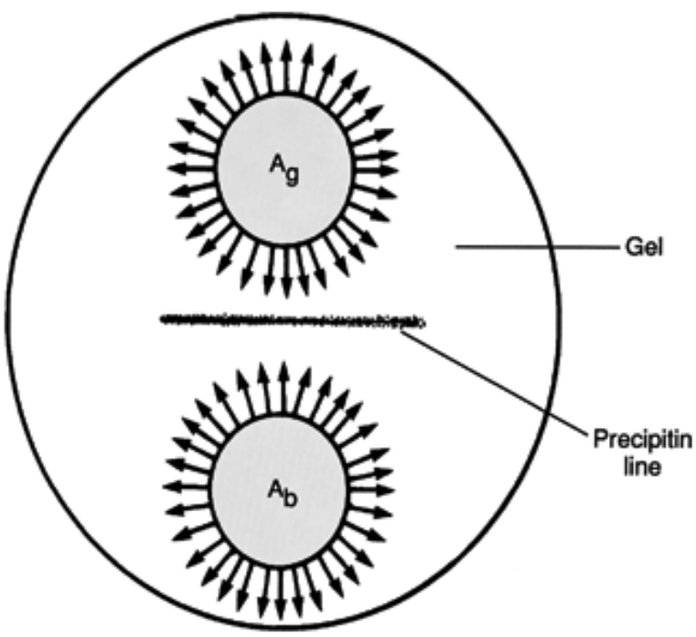


FIGURE 12.3. Immunodiffusion. Two wells are created with a gel punch in a gel slab. Antigen (*Ag*) is placed in one well and antibody (*Ab*) is placed in the other. As the solutions diffuse through the gel, antigen-antibody complexes form and a precipitin line results. If antigen and antibody are not pure, a precipitin line will form for each antigen-antibody pair of complexes generated.

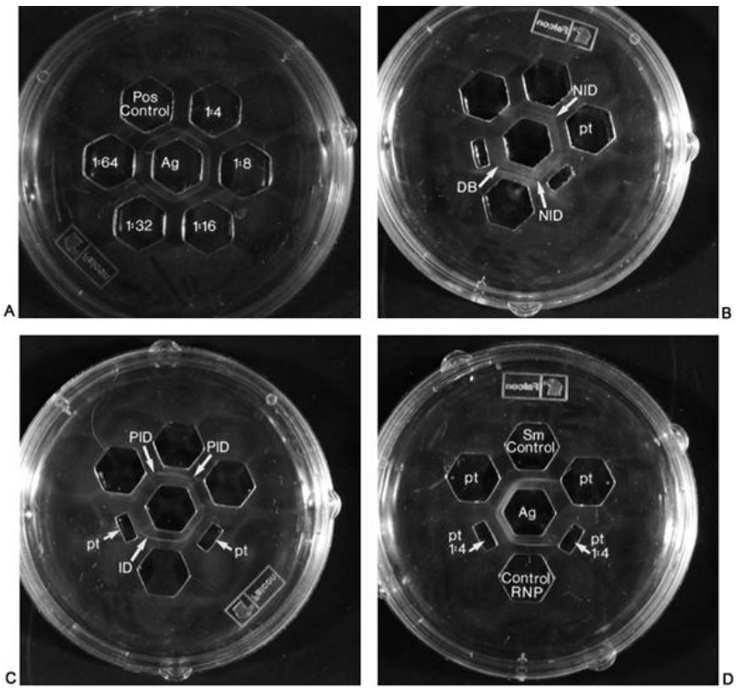


FIGURE 12.4. (**A**), Radial immunodiffusion plate. Antigen (Ag) is placed in the center well. Positive control for antigen is placed in the well marked *Pos Control*. Patient serum dilutions are made clockwise from 1:4 to 1:64. Notice that the precipitin line migrates closer to the patient wells that contain the more dilute specimen. The precipitin line forms a circle with positive control and no "spurs" or cross lines are observed. These are arcs of identity. (**B**) Immunodiffusion plate with arc of nonidentity (NID). Precipitin lines cross at the point of the *arrow*. Also visible is a double precipitin line (*DB*). This occurs when multiple antigens and antibodies are present in the sample and reagents. (**C**) Radial immunodiffusion plate with lines of partial identity (PID) and identity (*ID*). PID lines have a "spur" at the *arrows* unlike the smooth arc of identity. The precipitin line moves closer to the wells with higher dilution of specimen. (**D**) Immunodiffusion plate showing multiple precipitin lines. Patient sample (*pt*) on the left demonstrates antibodies for both Sm control and RNP control. The small rectangular well contains patient serum equivalent to 1:4 dilution.

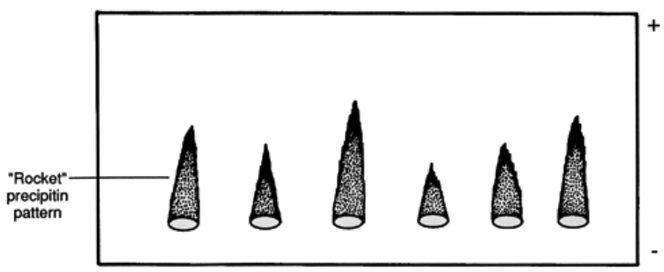
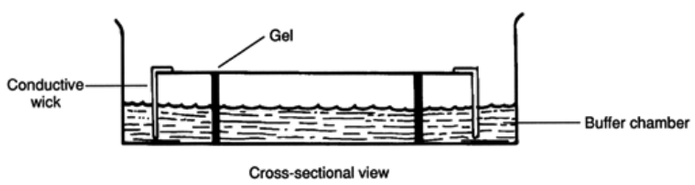


FIGURE 12.5. Laurell rocket electrophoresis. Antibody is mixed with gel and allowed to cool on a plate, forming a conductive surface. Small wells are punched at the cathode (-) side of the gel. The antigen forms, dissolves, and reforms the complexes at the electrophoresis front until the patterns are stabilized. The triangular shape of the precipitation front gives rise to the name rocket immunoelectrophoresis.



Some characteristics of the precipitin line are quite useful in the interpretation. First, the solution of higher concentration will diffuse more rapidly, and so the precipitin arc will be closer to the lower-concentration solution. Larger molecules will diffuse more slowly than smaller molecules, and if the antibody molecular weight is known, the molecular weight of the antigen can be approximated, assuming the two concentrations are similar. If there are multiple antigens or multiple antibodies, a separate line will be formed for each antigen-antibody pair. When this technique is set up in a circular fashion with multiple wells so that the antibody solution is in the center well and the circumference has several wells with different antigen solutions, the precipitin lines can be quite telling about the antigen-antibody complexes formed. As seen in Figure 12.4A, if the antigen is similar in two adjacent wells, the precipitin line is a continuous arc. If the antigens in adjacent wells are dissimilar, there are two separate arcs that cross (Fig. 12.4B), and if the antigens are similar but not identical, the precipitin line will have a single spur on one of the arcs (Fig. 12.4C). The terminology applied to these precipitin lines is "arc of identity," "arc of nonidentity," and "arc of partial identity,"

respectively. This is known as the Ouchterlony technique. Figure 12.4D demonstrates a real case with multiple precipitin lines showing positivity for the antibody controls, with additional unknown precipitin lines.

This procedure is reasonably simple and inexpensive but lacks sensitivity for quantitative applications. It is useful for applications that identify an antigen or antibody or determine their relative purities. Some experimentation has to be done to ensure that there is a reasonable relationship between the concentration of antigen and the concentration of antibody, or no precipitin line will form. If the starting protein solutions must be adjusted for concentration, adjustment of the final ionic strength and pH also may be required.

Another common diffusion technique is referred to as the Mancini technique. This procedure incorporates antibody into the agarose gel. The gel is heated to 50° to 55°C and the antibody is dispersed in this liquefied gel. If the temperature of the gel is allowed to exceed 55°C, significant antibody denaturation will take place. The gel then is poured onto a solid support and allowed to solidify, at which time small wells are cut into the gel. Each of these wells will hold a solution to be analyzed, and on a single slide a range of standard solutions as well as unknown samples can be assayed. The antigen in the well diffuses radially and precipitates as it encounters antibody in the gel. The antigen creates a concentration gradient as it diffuses from the well, and at the diffusion front when the antigen and antibody are at equivalence, a precipitate forms. As more antigen flows to the front, this precipitate redissolves because of antigen excess, and a precipitate forms farther away from the well. After several hours of incubation, this process stabilizes. Quantitative results can be obtained by plotting the diameter of the ring formed around the antigen well versus the concentration of the standards. Unknown concentrations then can be extrapolated from the curve by finding the diameter of the ring and reading the corresponding concentration.

This technique is dependent on a well-characterized monospecific antibody and reasonably simple antibody solutions. The formation of multiple rings generally means that the antibody is recognizing more than one antigen in the antigen solution, while the formation of asymmetrical rings usually is traced to a defect in the gel formation. With a little practice, this can be a moderately sensitive, simple, and cost-effective approach to analyzing some clinically important antigenic proteins.

ELECTROPHORESIS

Part of "12 - Immunochemical Methods"

The previously described diffusion methods are quite adequate for solutions that have relatively few components. In situations where the solution being analyzed is complex because of the number of components, or if it has multiple components of very similar chemical character, diffusion methods are less acceptable. An alternative to diffusion is electrophoresis. With this method, the components of a solution are separated based on size and charge in an electric field, and efficient separations can be effected on very similar molecules. Once the components of the solution have been separated, immunotechniques can be combined with the electrophoretic results to produce a quantitative analytical method.

Laurell Rocket Immunoelectrophoresis

This particular technique is named for the originator of the procedure, C. B. Laurell, and the rocket or triangular shape that the antigen-antibody complexes form at the end of the analysis (6). In this method, the antibody is dispersed in the agarose gel and the pH of the gel is adjusted to 8.6. This is an approximate isoelectric point for most IgG immunoglobulins, meaning that in the electrophoresis phase of the analysis, the antibodies will not migrate through the gel. When the gel is cool, wells are cut out at the bottom of the gel and filled with the antigen solution to be analyzed. Electrode wicks are carefully applied to the top and bottom of the gel plate as illustrated in Figure 12.5. The wicks must be attached carefully to the gel so that there is a constant current per cross-sectional areas of the gel. It is preferable to have a cooling device to keep the gel cool during the electrophoresis so there is no “drying and cooking” of the gel. As the antigens are electrophoresed from the well, they encounter the antibody that has been dispersed through the gel and the antigen-antibody complex forms at the electrophoresis front. As the antigen concentration changes, the complex dissolves and reforms, resulting in a triangular or rocket-shaped pattern at the end of the electrophoresis run. The area inside these triangular patterns is directly proportional to the concentration of antigen in the well. If the triangular shapes are symmetrical, peak height can be substituted for the area in determining the concentration of unknowns.

This technique is particularly sensitive to antigen-antibody concentrations and electric field strengths. It is important to use as low an antibody concentration as possible in the gel, and one indication of too high an antibody concentration would be the formation of excessive precipitant within the outline of the rocket.

Crossed Immunoelectrophoresis

This variation on rocket immunoelectrophoresis also is referred to as double-crossed immunoelectrophoresis or two-dimensional immunoelectrophoresis. The technique starts by electrophoresing the antigen solution on standard agarose gel under any set of electrophoretic conditions that will effect the separation of the antigens of interest. Once these antigens are separated, the gel is sliced so that the separated antigens are contained in a long thin strip of agarose. This gel strip of separated antigens is placed in or up against a second gel that has antibody dispersed through it. A second electrophoresis is conducted with the field at right angles to the original separation. From this point, the procedure is very similar to rocket immunoelectrophoresis in that, as the antigens enter and migrate through the antibody-containing gel, a precipitant forms, dissolves, and reforms, ultimately yielding triangular or triangular-like patterns of precipitation. Because several antigens can be analyzed at the same time, this technique is quite useful for metabolic studies, enzyme studies, and kinetic studies where there is interconversion of initial reactant and product.

The technique is susceptible to the same errors as Laurell rocket immunoelectrophoresis. The same conditions of low antibody concentration and low electric field strength should be observed.

Immunoelectrophoresis

One of the more common techniques used in the clinical laboratory is immunoelectrophoresis (IEP). This technique is straightforward, cost-effective, and amenable to processing large numbers of specimens as well as potentially being automated. The technique finds its primary application in the identification of serum proteins, and particularly in identifying pathological conditions such as monoclonal gammopathies, light chains, and general protein deficiencies.

The technique consists of preparing or purchasing a special gel plate for electrophoresis of the specimen containing the antigens. The special preparation of the plate requires cutting a trough in the gel and punching two small wells on either side of the trough. This is accomplished easily by using a commercially available template and gel punch or by purchasing the prepared gel from one of several manufacturers. The antigen solutions are placed in the small punched-out wells and electrophoresed long enough to separate the components of the antigen solution. On completion of the electrophoresis, antiserum is placed in the trough and allowed to diffuse into the gel. Either this antiserum can be directed against numerous proteins, as in the case of antiserum directed against human serum, or it can be specific for a single protein. In any case, precipitin arcs form where the antiserum and antigen meet and are at equivalence. These plates are read by looking at the opacity because of the precipitation of the antigen-antibody complexes, or the gel can be stained with one of the commercially available blue, red, or black protein stains. The arcs in the specimen are compared to controls and an interpretation is made as to the presence or absence of proteins and the presence of any unusual or pathological forms. Figure 12.6A depicts the gel plate with the trough and wells properly positioned and demonstrates the stained protein bands on a normal serum specimen. Figure 12.6B is an IgG myeloma.

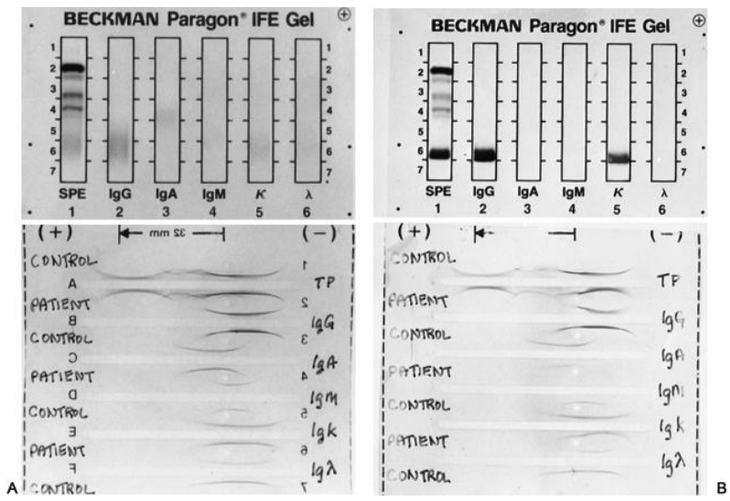


FIGURE 12.6. (**A**) Normal immunofixation electrophoresis (IFE) (top) and the same serum on immunoelectrophoresis (IEP) (bottom). Dark bands on IFE indicate the presence of a serum protein precipitated with antibody. Diffuse bands are indicative of polyclonal production of protein. IEP develops serum proteins as arcs with antibody. Control and patient sera are alternated in the gel wells so arcs can be compared directly. Antiserum is placed in the trough between the wells. (**B**), Monoclonal gammopathies of the IgG κ type. Dark bands in the serum lane and the IgG lane indicate increased concentration of IgG. The sharp band in the κ lane indicates a monoclonal increase in IgG with a κ light chain. This also is observed in the IEP (bottom) with a broad, thickened arc above the IgG trough. The arc below the κ trough demonstrates a nonsmooth arc resembling a spoon shape. All other arcs compare well with the normal control serum.

Immunofixation Electrophoresis

A modification of the immunoelectrophoresis technique first was described in the early 1960s by Alfonso and later refined by Alper and Johnson (7). The technique, known as immunofixation electrophoresis (IFE), has the advantages of rapidity and ease of interpretation over immunoelectrophoresis but suffers the disadvantages of being technique-dependent and a more costly procedure. Commercial materials are available, and the need for investigators to pour and generate their own electrophoresis plates no longer is necessary. Figure 12.6A depicts a normal serum specimen on an IFE plate with six lanes labeled

serum proteins, IgG, IgA, IgM, κ and λ. The clinical specimen, be it serum, cerebral spinal fluid, or urine, is placed at the indicated origin on each side of the lanes. Total sample volume is approximately 5 µl. The specimens are electrophoresed to separate the individual proteins in the sample. At the end of the electrophoresis time, a template is placed over the gel and pressed gently but firmly into the gel to create a seal around each of the lanes. For the lane marked serum protein, a protein fixative is added to denature and immobilize the serum proteins. For each of the other lanes, a specific antiserum is placed over the lane. The plate is covered and allowed to incubate for approximately 30 minutes. At the end of this time, the plate is washed in saline several times to remove all the proteins that have not been complexed or fixed. Therefore, the only proteins remaining in the individual lanes are those that were denatured with the fixative or those that formed complexes with the specific antiserum placed on the lane. Once the plate has been washed thoroughly, it is stained with a standard protein stain and the bands are visualized. Lanes marked G, A, and M indicate G, A, and M antisera and the presence of the respective immunoglobulin band. A diffuse area of stain is indicative of polyclonal proteins complexing with the antisera. A sharp, well-defined, darkly stained band indicates the presence of a monoclonal protein. The lanes marked κ and λ will indicate the type of light chain associated with the heavy-chain globulin or will indicate the presence of free light chains. A more detailed discussion of immunoglobulin gammopathies and protein band interpretation is available elsewhere in this edition. The main advantage of IFE is that the individual globulins and the light chains can be compared directly to the band staining of the whole sample serum proteins. If there are unusual intensities or the presence of abnormal bands in the whole serum sample, they can be compared directly to bands that are generated with specific antisera on the same plate. Figure 12.6B is an IgG monoclonal gammopathy of the κ light chain type.

INSTRUMENTAL METHODS

Part of "12 - Immunochemical Methods"

As immunoassays began to develop, there was an evolution from early qualitative precipitation techniques to procedures that separated the antigen and antibody, isolated the complex on a gel,

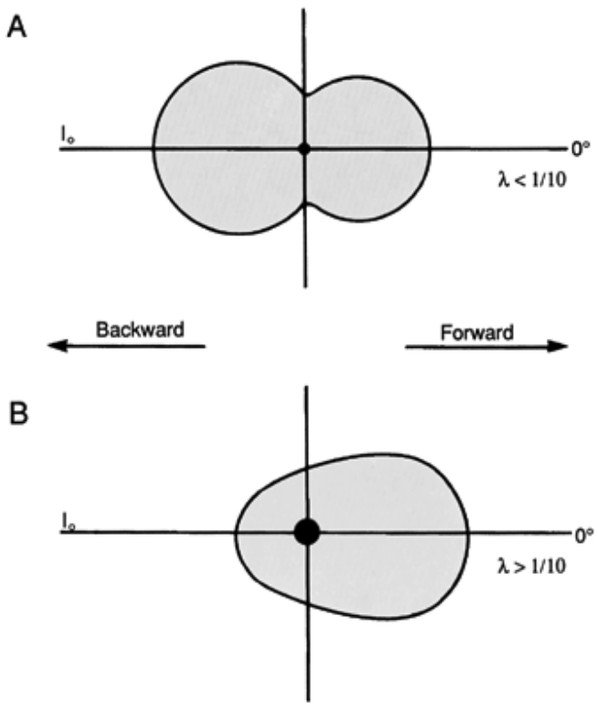


FIGURE 12.7. Light scattering as a function of particle size and angle. (**A**) Small particles (particle size less than one-tenth wavelength) scatter in a symmetrical envelope about the center of the particle. (**B**) Large particles (greater than one-tenth λ) scatter asymmetrically about the particle center, with greater intensity in the forward direction. The light source is unpolarized in both diagrams.

stained it, and quantified the stain densitometrically. The work by Drs. Berson and Yalow opened the door to a new era of instrumental techniques that had the requisite precision and sensitivity to make immunoassays a more controllable and routine analytical technique in the clinical laboratory (8). The original work of these investigators used a labeled technique whereby a molecule with a different physical characteristic was linked to the antigen-antibody reaction so a quantitative measurement could be made. From this, investigators began to pursue alternative methods. One approach took advantage of the formation of the antigen-antibody complexes without the introduction of additional molecules or atoms. In another method, investigators sought alternative labels that would have the same characteristics as the radioactive label but not share some of the disadvantages. The next sections describe these instrumental techniques. They have been arbitrarily divided into those that do not employ a label or marker for the measurement of the antigen-antibody complex and those that do.

Non-Labeled Methods

Light Scattering

Light scattering is the result of the interaction of electromagnetic radiation on particles as it passes through a sample. Light, being electromagnetic radiation, has both an electric field and a magnetic field component. As the particles in solution come under the influence of the electric field, the electrons move in one direction and the nuclei in the other, causing a relative separation. Because the electric field is sinusoidal in nature, the particle oscillates or resonates in synchrony with the electric component. This oscillation results in the reemission of energy with the same frequency as the incidence light. This reemission phenomenon is spherical in nature; that is, with reference to the specimen it can be measured in all directions. There are several characteristics of both the incident light and the particles that will affect the nature of the scattered signal. Among these are the wavelength of the incident light, whether or not the incident light is polarized, the size of the particle, the molecular weight, the concentration of particles, and the distance at which the observation or measurement is made.

Particle size generally is divided into two categories: one in which the particle is less than one tenth the wavelength of the incident light, and the other in which the particles are greater than one tenth the wavelength of light. In the first instance, the particle is small enough to be totally within the influence of the electric field. In this case, all portions of the particle see the field strength equally and become a point source of reradiation of the incident light. This phenomenon was first described by Lord Rayleigh and is described by the Rayleigh equation for small particle scattering.

$$I_s = I_i[16\pi^2 a \sin^2\xi/\lambda^4 r^2] \qquad (12.3)$$

In this expression, I_s is the intensity of the scattered light; I_i is the intensity of the incident light; a is the polarizability of the small molecule (i.e., the ability of the electric field to separate the electronic and the nuclear components of the particle); ξ is the angle of observation measured from the line of the incident light; λ is the wavelength of the incident light; and r is the distance of the observation. Although several important relationships can be taken from this equation, it is important to remember that this was derived for particles suspended in a gaseous medium and requires some corrections for solutions. The main characteristics of this expression are:

- The intensity of the scattered light is a function of the fourth power of the wavelength of the incident light; that is, short wavelengths will scatter more intensely than longer wavelengths.
- The intensity of the scattered light will be reduced in relation to the square of the distance of the obstruction. Therefore, the collectors or the measurement device should be as close to the source of scatter as is physically possible.
- As this equation is written, the right-hand side has essentially two components. The first is the intensity of the incident light, and the second is the combination of factors and constants that describe a given system. Therefore, for any given system, the scattered intensity will increase directly as the intensity of the incident light is increased. This is the rationale for the use of high-intensity sources such as lasers to yield higher scatter intensities.

The corresponding equation for liquid systems with particles in suspension and a nonpolarized source of incident radiation is given below:

$$I_s = I_i\,[4\pi^2\,(^{dn}/_{dc})^2\,Mc\,(1+\cos\xi)\,/\,N\lambda^4 r^2] \qquad (12.4)$$

The symbols remain the same as in the previous equation, with the addition of the dn/dc, a correction factor for the incremental change in refractive index of the solution with an incremental change in the concentration of particles; M is the molecular weight of the particles in question; c is the concentration; N is Avogadro's number. This expression now relates the concentration and the molecular weight of particles to the intensity of the scattered light. This is the more useful form for solution chemistry and the one that is utilized in immunoassay systems.

There are other considerations in using light scattering for immunoassay. Among these is whether the source is polarized. Other sections in this text give a detailed description of the difference between polarized and nonpolarized light. For the nonpolarized source, the scattering intensity generally is greater. Looking at the scattering intensity diagram associated with small molecules, one realizes that the scattering envelope is symmetrical about a 90° axis to the particle. The scattering envelope has a remote resemblance to the figure eight. This is in contrast to the scattering envelope of large molecules, which is asymmetrical with the greatest intensity in the forward direction and minimal intensity in the backward direction. Figure 12.7 depicts these differences. By comparing these two diagrams, it is clear that, when using scattered intensity as a measure of the presence of small molecules in a sample composed of both large and small molecules, it would be most advantageous to measure the backward scatter, because this would be least affected by large particles.

Large-particle scattering differs from small-particle scattering in that the particle generally is considered larger than one tenth the wavelength of incident light. This larger particle size means

that as the electric fields impinge on the particle, different portions of the particle will respond to differences in the electric field strength. Therefore, a large molecule can have several points of oscillation and reemission. These will constructively or destructively interfere with each other. The net result of multiple points of scattering is that the strongest intensity is in the forward direction and the weakest in the backward direction.

The molecular weight characteristic that is measured in light scattering is a weight average molecular weight. In immunoassay measurements, this means any reaction that takes place must increase the average molecular weight of the particles in solution. The greater this average molecular weight change, the greater the changes in intensity of the scattered signal. This last point is one of the reasons that light scattering techniques have not enjoyed the popularity that was initially expected. The principle of light scattering should allow the direct measurement of the complex formation. There is no requirement for additional labels or secondary indicator molecules. The problem arises when one tries to measure steroids, drugs, or other low-molecular–weight compounds in moderate to low concentrations in a matrix made up of serum proteins and lipids. The combination of an antibody to these monovalent antigens does not shift the average molecular weight sufficiently to generate a strong signal above background. Many things have been done to decrease the background signal and improve the scattered signal from the solution being investigated, but once additional steps are required, light scattering methodologies begin to have practical limitations. When an immunoassay system is set up to measure high-molecular–weight proteins in moderate to large concentrations, light-scattering is very effective. If the concentration of large molecules is too high, the macromolecular complexes that form begin to take on the appearance of aggregates. Aggregate formation does one of several things. First, the *number* of molecules present in the solution changes. Aggregates are seen as a single particle, thereby effectively reducing the number of particles in solution. Second, the average molecular weight changes by a very significant amount. Third, the physical size of these complexes exceeds the limitations of the Rayleigh-Debye scattering theory. A much more complex physical-chemical model must be employed to describe the scattering phenomenon as associated with aggregates in solution. Because of these multiple changes in solution, the results are nonlinear. Fortunately, this nonlinearity of the response can be handled by the conventional data handling techniques employed in other immunoassay procedures. These are discussed under Data Reduction.

Turbidimetry and Nephelometry

Light scattering is measured in the laboratory by one of two methods, turbidimetry or nephelometry. The difference is in the angle at which the analytic signal is measured. For turbidimetry, the detector is in line with the incident source and usually is referred to as either zero angle or 180°. For nephelometry, any other angle is used to measure the scattered signal. Turbidimetry measures a decrease in signal because of to the scattering of incident light by particles in the solution. This means that you are measuring a small difference between large numbers. Nephelometry, on the other hand, measures scattered signal against no signal because it is out of line from the incident radiation. The sophistication of modern spectrophotometers is such that turbidimetric measurements are adequate for many laboratory procedures when the antigen or antibody of interest is in moderate to high concentration. Nephelometry, however, is suited to making those measurements when the antibody or antigen is in low concentration. The much smaller signal generated by these molecules will be more easily detected against a very low background.

The sample preparation for light scattering measurements is the same whether one uses a turbidimetric or a nephelometric approach. Samples must be free of any material that is capable of scattering light such as chylomycra, dust, or other particulate introduced from the collection and preparation process. It also is important that the wavelength of light used to measure the scattering is chosen so that it is not in a transition band of a component in the solution. This means that radiating a specimen with a wavelength of light that will be absorbed by any of the components in the serum will reduce the emitted light. Turbidimetric measurements will interpret this as high scattering as a result of a high concentration of analyte, while nephelometric measurements will be made with an extremely weak signal. Interference from phosphorescence and fluorescence can be controlled with good optic design or by using polarizing filters. A potential problem for scattering and all methods that measure antigen-antibody

reactions is that of antigen excess. Instrumentation that relies on endpoint determination, that is, takes a measurement at the beginning of the reaction and one at a later time, must be capable of detecting antigen excess. This becomes less of a problem in automated instruments that use a kinetic approach, because algorithms are available for the microprocessor to determine whether the kinetics are taking place in the antigen-deficient, the equivalent, or the antigen excess zone of the curve.

Several instruments currently are on the market that are capable of doing turbidimetric or nephelometric measurements in either an endpoint or kinetic mode or both. The computer processing power within these instruments is significant, and most have complex data handling routines that allow analysis of complex curves.

Labeled Methods

It is obvious from the previous sections that there is no single immunoassay system that is suited for all the analytic measurements in the clinical laboratory. To increase the sensitivity of immunoassays, labels or markers were introduced. The first of these was the radioactive label, but that has been followed by enzymes; fluorescent, luminescent, and phosphorescent compounds, as well as metals and particle labels. The objective in using labels in immunoassay is to improve the sensitivity so that heretofore-unmeasurable analytes can be quantitated. A second objective, particularly of more recent labels, is to make the methodology simple and straightforward so increasing numbers of laboratorians would have access to the technology using standard clinical laboratory equipment. The second objective has been amended because newer labels employ technologies that are quite complex, rely on equipment that has not historically been in the clinical laboratory, and necessitate the use of powerful on-board computers for data reduction. Most of this complexity has been hidden from the technologists with the “black box” design. In addition, although the objectives are universal, the types, methods, and approaches to labeling are diverse. Labels have been linked covalently to antigens and antibodies, both the primary antibody used to detect an analyte, and secondary antibodies used for the detection of the antigen-antibody complex. Also, molecules such as enzyme inhibitors, cofactors, and fluorescent quenchers have been used as the covalently linked molecule but are there to affect the action of another molecule that is monitored for analyte quantitation. The current laboratory and research literature is replete with methods and variations on methods dealing with new labels, the chemistry of label analyte attachment, and system design. Although the details of these methods differ, they can be grouped into broad classifications based on conceptual similarities.

Labeled Antigen

There are several points to consider when selecting labels to be covalently attached to an antigen. One is the size of the antigen. For small antigens such as drugs, vitamins, and steroids, the antigenic binding site may comprise a significant portion of the molecule. If a very large label, such as an enzyme, were to be linked to this antigen, it must be done in a manner that will not sterically block the epitope on the antigen. Small labels such as isotopes, spin labels, and fluorophores must be attached in such a way as to not sterically hinder binding or change the electronic environment of the binding site. In all cases, the label must be attached so that it remains an efficient monitor of the antigen-antibody interaction but must not destabilize or significantly alter that antigen-antibody interaction.

Some successful approaches to these problems have been the use of carbon spacer arms, that is, long carbon chains that attach the antigen on one end and the label on the other. This spatially separates the label and the antigen so that steric and electronic interference is minimized. Synthesizing the antigen with ^{14}C or tritium atoms, so that it is radioactive, has the advantage of providing a true label that is structurally indistinguishable from the natural antigen. If the antigen is large, such as proteins, then the labels can be added much more easily and with less probability of interfering with the antigen-antibody interaction. Fortunately, except for research and very specialized applications, the antigens of clinical interest have been prepared with various types of labels and are commercially available. If an investigator elected to label his or her own antigen, a first approach would be to investigate using derivatives of carbodiimide, succinimide, isothiocyanate, or other organic reactions that would be appropriate for the reactive groups available on the antigen.

Antibody labeling is similar to labeling peptide or protein antigens. Most of the antibodies used in clinical chemistry are commercially available with a wide selection of labels.

Assay Design

Once the reagents have been assembled, there are several possible configurations for the ultimate assay system. These include competitive, noncompetitive, homogeneous, and heterogeneous systems. Each of these has its specific advantages and disadvantages. The choice of an essay system is generally a function of the type and quality of label and antibody available.

Competitive Assay

In this type of assay, the labeled antigen and native antigen compete directly for a limited number of binding sites in the reaction mixture. These binding sites are usually present as a fixed concentration of primary antibody. The reaction follows the law of mass action. Equations 12.1 and 12.2 illustrate the relationship between labeled and unlabeled antigen in the final antigen-antibody complex. To achieve maximum sensitivity in this assay system, it is imperative that the labeled antigen has the same binding characteristics to the antibody as the unlabeled antigen. Labels that are very large compared with the native antigen or that significantly alter the electronic environment of the antigen can alter the binding characteristics of the antigen with the antibody. This would make one of the antigens a favored binding partner, and, although the assay may work, it could have several technical difficulties, including a lack of sensitivity. The favored binding partner here is not always the native antigen.

When the antibodies are generated against an antigen with a covalently linked carrier, one of the epitopes recognized by the antibody is the organic linkage. When the antigen is labeled, if

that same organic linkage is used, the antibody could bind most strongly to that epitope. In this case, the native antigen would have difficulty in displacing the labeled antigen, and a serious reduction in sensitivity would result. These competitive designs are sometimes referred to as limited reagent systems because the concentration of antibody is held constant at a level that is insufficient to bind all antigens present. This was the predominant assay design of the original RIA methodologies.

Noncompetitive Assay

In noncompetitive methods, there is sufficient reagent to bind the antigen present in the specimen. One very popular approach to this technique is to fix an antibody to a solid support. That support can be a microtiter well, test tube, bead, or other appropriate solid support. Sample containing the antigen is added with the appropriate buffer to the solid support and allowed to incubate. During the incubation time, the antigen binds to the fixed antibody. The support material is then washed and incubated with a second antibody, which may be labeled. This second antibody then binds to a different epitope on the antigen and forms what is commonly referred to as a "sandwich." The support is washed a second time to remove all the unbound antibody, and, if the second antibody had been labeled, the measurement is made. If the second antibody is not labeled, a third antibody that will recognize the second antibody is added. This third antibody carries the label. The requisite reactants are added, and the measurement is made.

HETEROGENEOUS AND HOMOGENEOUS ASSAY SYSTEMS

Part of "12 - Immunochemical Methods"

In heterogeneous assay systems the physical characteristic of the label is unchanged by the binding process. The label that indicates the presence of the antigen-antibody complex must be separated from the extraneous label to make the quantitative measurement. An obvious example would be the radioimmunoassay system, in which the bound radioactive label emits the same signal as the unbound radioactive label. Techniques have been developed that will allow the separation of these two forms of label, allowing the measurement of either fraction. Separation techniques such as precipitation with other antibodies or chemicals, adsorption onto solid surfaces, chromatography, electrophoresis, and some innovative approaches such as coated magnetic particles, have all been successfully employed. The choice of the separation technique used in a heterogeneous assay is primarily dependent on the label used and the clinical application. For example, if the application were one that requires processing large numbers of specimens in a short period of time, then a precipitation, adsorption, or decanting technique would be most appropriate. Conversely, in some research applications, one may be able to process small numbers of specimens by using chromatographic or electrophoretic separation.

In homogeneous assays, the physical characteristic of the label changes upon binding; therefore no separation technique is required. If enzymes are used as labels, the enzymatic activity can either increase or decrease when the antigen-antibody complex formation takes place. Likewise, fluorophores can be either activated or quenched upon formation of the complex. These assay methods are generally thought to be simpler and more straightforward because the analytic measurement can be made in the presence of all the reactants.

Immunoassays designed around any of these techniques require a significant number of controls. It is important to monitor antibody characteristics, background signals, nonspecific signals, and potential interference.

SPECIFIC ASSAY TYPES

Part of "12 - Immunochemical Methods"

Radioimmunoassay

Radioimmunoassays take advantage of a radioactive atom being incorporated into one of the reactants as a label. The most common clinical atoms used are ^{125}I, ^{57}Co, ^{3}H, and ^{14}C. The γ-emitting nuclides ^{125}I and ^{57}Co can be incorporated directly into the structure of the antigen, as in the case of the thyroid hormones (^{125}I) and vitamin B_{12} (^{57}Co). The ^{125}I also can be inserted through organic reactions into the ring structure of the amino acids of peptides and proteins. These γ emitters decay at a rate that is characteristic of the atom itself. This decay is unaffected by the environment of the immunoassay reaction and can be counted with reasonable efficiency with standard γ-counting instrumentation.

The ß emitters ^{14}C and tritium generally are reserved for small molecules such as drugs and steroids and are usually substituted for "cold" atoms in the structure. As with the γ emitters, the physical decay of ß emitters is unaffected by the environment of the assay. However, the energy of the ß emission is such that it is easily absorbed by many naturally occurring compounds, making it impossible to measure ß emission directly. In order to detect ß emission, the samples are counted in a scintillation cocktail. This cocktail contains a carefully chosen solvent and organic flurophore that absorb the ß energy and reemit it as a flash of light. This flash then is counted in a scintillation counter that detects these flashes as individual events. Because energy transfer is involved, the efficiency of the detection of events is much lower than with γ emitters.

Because of the characteristics of radioactive labels, all RIAs are heterogeneous assays. The separation step in RIA generally is designed to fit the type of label employed. For labels that are small, like drugs and steroids, an adsorbent such as activated charcoal can be used to adsorb out the unbound label. This is a carefully timed process so that the charcoal does not strip away the label from the antibody, thus destroying the antigen-antibody complex. At the end of the incubation time, the samples are centrifuged and the supernatant is decanted into a second tube. The investigator has the choice of counting the supernatant that contains the bound complexes, the charcoal precipitant that contains free label, or both fractions so that a total and percent bound can be calculated on each specimen. This is older technology that has generally been replaced by simpler methods.

Another procedure is to precipitate the antigen-antibody complex with agents such as ammonium sulfate, polyethylene glycol, or additional antibodies that recognize the first antibody

used to form the antigen-antibody complex. Here the second antibody is used to make a large lattice of the antigen-antibody complex in solution. All of these methods require centrifugation to pellet the precipitant. The supernatant is decanted, and the pelleted material can be counted as the antigen-antibody complex.

Other methods use a first antibody that is fixed to a solid surface. The solutions containing the antigen and the label are added, and at the end of the incubation time, the solution is decanted. The bound antigen being fixed to the surface of the reaction vessel is then counted directly.

RIAs are capable of precise and accurate quantitation of antigen-antibody complexes if sufficient controls are included in each assay. Obviously, standard concentrations of unlabeled antigen must be used to construct a dose-response or standard curve. In addition, controls must be added to monitor the total amount of radiation added to each tube, the nonspecific binding of the radioactive label to other proteins and surfaces in the reaction medium, and, in the case of ß emitters, quench controls that monitor the energy transfer process in the specimen matrix under investigation. It is not unusual in an RIA to find a large percentage of the samples being processed to be control or calibration specimens.

The advantages of RIA are its exquisite sensitivity and the ease with which multiple specimens can be processed. With a little experience, a laboratorian can batch-process 100 or more specimens in duplicate with all the requisite controls. This makes it an attractive technology for the medium to large clinical laboratory.

The disadvantages of RIA are the short shelf-life of the labeled reagents because of the short half-life of the radioactive nuclides themselves and the damage that the nuclides do to the proteins in the reagent. In addition, there is a requirement for strict record keeping and personnel monitoring. This makes RIA somewhat more costly than other immunoassay methods, although in recent years, competition from nonisotopic methods has forced the price of prepackaged materials to a more affordable level.

Enzyme Immunoassay

Enzyme immunoassay (EIA) is a generic term that encompasses all assays that fulfill the following criteria. First, they employ immunologic elements to detect the analyte of interest, and second, enzyme activity is used to quantitate the analyte. It will become obvious that although these two criteria are relatively straightforward, some of the chemistry and assay designs are not. Many reasons can account for this. There is the general attractiveness of monitoring immunoassays with enzymes. Enzyme reactions are well-defined, familiar measurements that have been made in the clinical laboratory for decades. The enzymes themselves have been isolated, purified, and characterized and are made available by commercial suppliers at relatively low cost. The enzyme preparations are stable with long shelf lives, and they represent a minimal hazard to the laboratorian. Theoretical sensitivities range from 10^{-14} to 10^{-16} mol/liter, giving these assays the sensitivity required for many clinical assays (9, 10). Measurements either can be made on standard spectrophotometers or can be automated on high-throughput instrumentation. In addition, unlike the RIAs, EIAs require no extra licensing, record keeping, or expensive disposal procedures that ultimately contribute to the cost of laboratory analysis.

When laboratorians and researchers are left to their own devices, they tend to “tinker.” This inherent inquisitiveness, coupled with the almost unlimited number of combinations and permutations of reactions that can be put together, has led to some very ingenious and complicated assay systems. Obviously, only the most reliable and useful of these techniques will find their way to the clinical laboratory for routine use.

The function of the enzyme in the EIA is to act as an amplifier. Each enzyme linked to the antigen-antibody complex will operate on the substrate and produce products that can be measured to quantitate the formation of the antigen-antibody complex. Enzymes will continue to turn over substrate, so a few bound enzymes will result in multiple moles of product formed, resulting in an amplified signal. The assay can be designed so that there is maximum activity when no free analyte is present or there can be maximum activity when maximum free analyte is present. The following sections provide the basic requirements of the enzyme and the substrate that must be considered when designing an assay.

Enzyme Availability

Enzymes can be of plant or animal origin as long as they are abundant and isolated without difficulty. The enzymes should be robust in the sense that they can be subjected to mild organic reactions without losing appreciable activity, and they should have significant activity in the environment in which they will ultimately be used to make analytic measurements. Enzyme activity can be defined as the quantity (micromole, nanomole, picomole) of product formed (or substrate consumed) per unit time under a given set of reaction conditions. The more product formed per unit time, the greater the potential sensitivity of the assay. The enzymes should be inexpensive and available in virtually unlimited quantity. These criteria have essentially limited commercial assays to the use of horseradish peroxidase and animal source alkaline phosphatase.

Substrate

The substrate should share some of the characteristics of the preferred enzymes in that it should be available at reasonable cost in a purified form. There should be a significant difference in the measured characteristic of the substrate compared with the products formed in the enzyme reaction. This means that if the substrate is colorless, the generated product should be colored or fluorescent or should possess any other characteristic that allows confident measurement of the products in the presence of the substrate and other potential interferents. The substrate should also be stable so that it undergoes minimal nonenzymatic conversion to products. As with other reagents, biological hazards should be minimized.

One very popular type of EIA, known as the enzyme-multiplied immunoassay technique (EMIT) assay, has been developed commercially. This is a homogeneous-type assay in which the

enzyme has been conjugated to the analyte in such a way that the enzyme retains its activity in the absence of analyte-binding antibody. When free analyte is added as either a control or a specimen with a fixed amount of antianalyte antibody, the free analyte and the enzyme conjugated analyte compete for the limited number of binding sites on the antibody. When the antibody binds with an analyte that has been conjugated to an enzyme, the enzyme activity is inhibited. Consequently, the inhibition of the enzyme activity is directly proportional to the concentration of free analyte in the solution. This approach has been used very successfully in the therapeutic drug monitoring area. This is a particularly good example of a homogeneous assay because no separation step is required to measure enzyme activity. The final activity measurement can be made in the presence of all the components in the reaction.

A second type of EIA that has been popularized by researchers and commercial vendors is the enzyme-linked immunosorbent assay (ELISA). Although the term ELISA is used interchangeably with EIA, it is more appropriately applied to assays of the heterogeneous solid-phase type. Any solid material can be a candidate for the support component in an ELISA as long as it is capable of irreversibly adsorbing protein to its surface, or the surface can react chemically to covalently bond protein to the surface. Materials that have been used successfully in the past include membranes, microtiter plates, plastic test tubes, plastic beads, various polymers, and coated metal particles. These solid surfaces all perform the same function in the ELISA—to fix one of the components so that the bound enzyme can be separated from the free enzyme. In this assay, the enzyme usually retains its activity independent of its bound or free state.

In the competitive ELISA, controlled quantities of the analytic antibody are fixed to the solid phase. Then, the analyte from the specimen is mixed with the enzyme and conjugated analyte, and these are allowed to compete for the limited number of antibody-binding sites on the solid surface. After a sufficient incubation time, substrate is added, and the amount of conjugated analyte bound to the fixed antibody will generate measurable product. The concentration of the substrate product is inversely proportional to the concentration of free analyte in the test specimen.

Another approach to an ELISA system is the “sandwich” or capture design. In this procedure, an analytic antibody is fixed to the solid phase. The analyte from the test specimen standard or control is added and allowed to react with the antibody on the fixed surface. After a sufficient incubation period, the surface is washed and a second analytic antibody conjugated to an enzyme is added and allowed to incubate. The second antibody recognized a different epitope on the antigen; therefore, this system works best with analytes or antigens that are large enough to have epitopes separated on the molecule so that the two antibodies can bind freely. After the second enzyme-conjugated antibody has incubated, the surface is washed and substrate is added. The generated product is directly proportional to the concentration of analyte in the test specimens.

ELISAs also may be designed to determine antibody concentrations. Applications to screen cultures for antibody production or to measure antibody in the serum of patients suffering from various autoimmune diseases have been used in the clinical laboratory. In these assays, the antigen is fixed to the solid surface and the test specimen is added so that the antibody will bind to the fixed antigen. The surfaces are washed, as in the other assays, and a second antibody-directed against the IgG of the species generating the first antibody, e.g., human IgG in the case of autoimmune disease, which has been conjugated to an enzyme, is added. After incubation, the surfaces are washed, the substrate is added, and the concentration of the generated chromogen is directly proportional to the antibody concentration in the test specimen.

It should be obvious that there are numerous combinations of antigen-antibody, enzyme, cofactor, and substrate that can be combined in various ways to detect antigen-antibody interaction. The more complex the assay design and the more steps involved, the more difficult it will be to maintain a consistent and reproducible assay over a long period. In the simplest of these assays, one still has to be concerned about maintaining an optimum environment for both the antigen-antibody binding and the enzyme activity. As more steps are added, the potential for poisoning the enzyme, reducing the antigen-antibody interaction, or just generally increasing the background signal, becomes greater. Some of the more common enzyme reactions are listed in Table 12.1.

TABLE 12.1. ENZYMES COMMONLY USED IN EIA

Enzyme	Some Substrates
Alkaline phosphatase	4-Methylumbelliferone phosphate *p*-nitrophenyl phosphate
Glucose-6-phosphate dehydrogenase	Glucose-6-phosphate + $NADP^+$
Glucose oxidase	Glucose
Peroxidase	H_2O_2
ß-Galactosidase	ß-Galactosides

Fluorescent Immunoassay

Fluorescence is a phenomenon exhibited by molecules that have the ability to absorb light energy and then dissipate that energy, with some of its being re-emitted as light of a longer wavelength. Fluorescence lifetimes are on the order of 10^{-9} to 10^{-5} seconds, which means there is a delay between the absorption of the exciting light and the emission of the fluorescent light (11). Figure 12.8 depicts the relationship between the wavelength of the absorbed light and the wavelength of the emitted light. The difference between these two wavelengths is referred to as Stokes' shift. Fluorometers have been designed to take advantage of both of these characteristics. Excitation sources can be pulsed, i.e., turned on and off very quickly, so that measurement of the emitted light is less encumbered with the more intense exciting light. The delay between absorption and emission makes this possible. This can be combined with detectors chosen to be more sensitive to the emitted light than the exciting light for increased performance. In addition, fluorescence is emitted in all directions, so detectors can be placed anywhere around the sample and are not in direct line with the exciting light, as is the case in most adsorption spectrophotometers.

Molecules that fluoresce generally have some common characteristics.

First, they are usually ring compounds that exhibit extensive conjugation. The greater the number of alternating single and double bonds the greater the stability of the excited molecule, which leads to enhanced fluorescence. It is also observed that the greater the molecular planarity and the molecular rigidity, the greater the tendency to fluoresce. Some molecules, particularly chelaters, can be made strong fluorophores by adding the metal ion for chelation (11). It also is known that concentration, temperature, pH, and ionic strength among other things have a significant effect on the fluorescent efficiency of a molecule.

Fluorescent efficiency is termed quantum yield and is a number between 0 and 1 that expresses the amount of emitted light compared with the amount of absorbed light. Any condition that will allow an excited molecule to dissipate its excess energy through processes other than light emission will decrease the quantum yield. Conditions that are commonly encountered allowing a molecule to dissipate its energy are intermolecular collisions caused by high concentrations of fluorophore, elevated temperatures of reaction mixtures, and destabilization of resident structures by other components in the reaction mixture.

Very dilute solutions of fluorophores, in the range of micromolar to picomolar concentrations, generally are used at room temperature or cooler. Fluorescence can be measured with common clinical laboratory instrumentation and present relatively inexpensive, nonhazardous methodologies. It should be noted that fluorescence is a complex process usually dealt with in advanced chemistry and physics textbooks. The observations outlined here are presented as broad concepts so the reader has an appreciation for the requirements of a usable fluorescent immunoassay.

Assay designs in fluoroimmunoassay are very similar to the EIA designs. A fluorescent molecule can be conjugated to an antibody or an analyte and the reaction can be monitored by the presence or absence of fluorescence. The concentration of analyte in the specimen is determined by comparison to standards. The assays can be designed to be heterogeneous, homogeneous, competitive, or any other design that satisfies basic analytical principles. The more interesting uses of fluorescence are in the techniques of fluorescence polarization and time-resolved fluorescence.

Fluorescence Polarization

Fluorescence and fluorescence polarization measurements take advantage of three of the characteristics of fluorescent molecules. These are quantum yield, the time differential between adsorption and emission, and emitted light of a different wavelength than exciting light. The specific nature of the interaction of electromagnetic radiation and matter can be found elsewhere in this book and in other texts. Suffice it to say that a molecule must have a specific orientation in relation to the exciting light for the molecule to absorb energy and be elevated to an excited state. Normal light has a mixture of orientations of electric and magnetic fields. If a polarizing filter were placed between the light source and the sample, the sample would see only one orientation of the electric and magnetic component. Samples with the appropriate orientation would then absorb this light and be elevated to an excited state. If this system were frozen in time, the fluorescence emission would have the same orientation as the exciting light. However, due to the molecular rotation and the length of time required for a molecule to go from exited state to ground state, the resulting fluorescence is a mixture of electric and magnetic orientations. By imposing another polarizing filter between the sample and the detector, only the fluorescent light with the appropriate orientation will be detected. Therefore, the intensity of the polarized fluorescent signal is related to molecular orientation. Conceptually, this means if something can be done to decrease the rotational velocity of the fluorescent molecules so they retain their original orientation until they fluoresce, the intensity of the fluorescence signal will increase. In fluorescent immunoassays employing polarized signals, the rationale is that the analyte is labeled with a small fluorophore and it will rotate rapidly. In addition, relatively few molecules will emit the appropriate plain polarized light for detection. Once these analyte molecules are bound to an antibody, however, their hydrodynamic radius increases significantly. This reduces the rotational velocity, increasing the number of molecules that have the appropriate orientation. The fluorescent signal is thus increased. Under these conditions, a competitive fluorescent immunoassay (FIA) may be set up so that maximum fluorescence signal is obtained in the absence of any unlabeled analyte. When unlabeled analyte is added, competition between the labeled and unlabeled species will reduce the fluorescent signal. The concentration of analyte is inversely proportional to the fluorescent signal. Quantitation is achieved by comparison to a standard response curve.

It was mentioned earlier that the quantum yield is also taken advantage of by fluorescent and fluorescent polarization methodologies. Remembering that quantum yield is defined as the number of photons emitted divided by the number of photons absorbed, and using the Beer-Lambert equation, a series of substitutions and algebraic rearrangements can be made to show that the intensity of the fluorescent signal is directly proportional to the intensity of the exciting signal. That means for a given system, if the intensity of the exciting light is increased (i.e.,

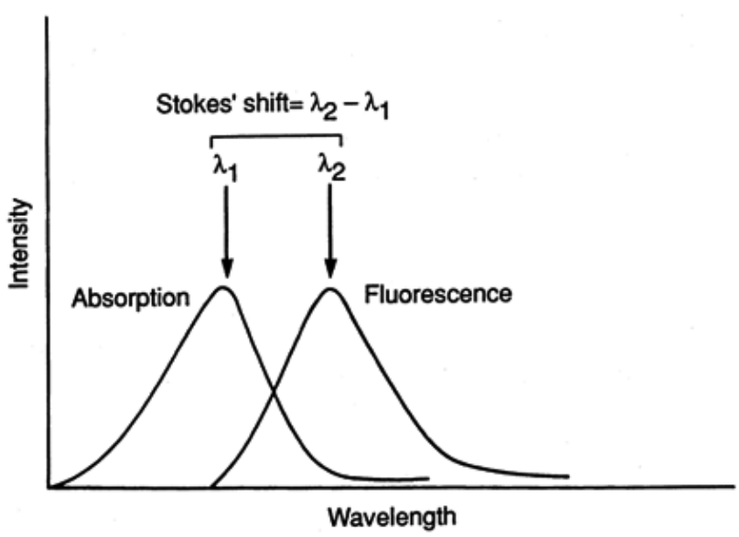


FIGURE 12.8. Relationship of absorption spectrum to emission spectrum of fluorescent molecule. Peak absorption is indicated by λ, and peak emission by λ_2. Stokes' shift is the difference between the wavelengths at the spectral peaks.

through high-intensity lamps and/or lasers), the intensity of the fluorescent signal can be increased. This is one of the reasons that fluorescence measurements are intrinsically more sensitive than absorbance measurements. Absorbance is dependent on the concentration of absorbing species and independent of incident light. Fluorescence is dependent on the intensity of the exciting signal, which can be changed in any given system. High-intensity sources coupled with polarization techniques and photon counting detectors give fluorescence immunoassay a sensitivity that is adequate for many assays. The best fluorescent polarization assays appear to be those dealing with small to medium-sized analytes.

Time-Resolved Fluorescence

Another innovative technique employing fluorescence is time-resolved fluorescence. Again, the basic design of the immunoassay is similar to the other immunoassay designs discussed. The unique feature about time-resolved fluorescence is the selection of a fluorophore that has a large Stokes' shift and an extended decay time. The large Stokes' shift means that the emitted light is separated from the exciting light, reducing the contributions from the excitation source in analyzing the signal. It is, however, from the time-resolved feature of this system that significant advantages are realized.

When a fluorophore has an extended decay time, it means that the fluorescence from a single excitation is extended over a long period. One of the characteristics of the lanthanide chelate complexes is that the fluorescence decay extends into the microsecond range, compared with nanoseconds for conventional fluorophores (11). For assays employing this type of fluorescent detection, the instrumentation excites the fluorophore with a high-intensity pulsed excitation. There is a short delay to allow the background and nonspecific fluorescence to decay before a measurement of the analytic fluorescence due to the lanthanide chelate complex is made. If this process is repeated several times over the course of a few seconds, a significant number of measurements can be made, and a statistically reliable value can be generated for the complex being measured. Because all of the background and nonspecific fluorescence is separated (resolved) from the specific analytic measurement based on time, the process is called time-resolved fluorescence. The analytic measurement is separated from the non-analytic measurements by the passage of time. Coupled with the fact that there is a large Stokes' shift, and the emission wavelength is very sharp, only a very small correction has to be made for the nonspecific signal of the lanthanide chelate complex. In principle, this methodology can potentially have sensitivity equal to or beyond that of radioimmunoassays. At least two commercial systems that take advantage of lanthanide chelate, time-resolved fluorescence.

All fluorescent techniques suffer from several limiting factors. First, fluorescence is sensitive to temperature and viscosity. Reaction conditions must accordingly be held within very narrow tolerances. In addition, most biological specimens contain materials that have natural fluorescence. In the serum or plasma specimen commonly found in the clinical laboratory, serum proteins and molecules such as bilirubin contribute to a significant background fluorescence signal. In these cases, extensive controls or corrections must be applied to separate the analytic measurement from the background signal. Time-resolved fluorescence tends to minimize these problems but does not do away with them totally. Additional interference can be a result of the internal filter effect. This internal filtering occurs when there are molecules in the measurement system that absorb light of the same wavelength as the light being emitted by the fluorescing molecules. One can try to minimize these internal filter effects by choosing fluorescent labels that emit at longer wavelengths. Biological specimens absorb the shorter visible wavelengths more often than the longer wavelengths of light.

Bioluminescence and Chemiluminescence

Chemiluminescent and bioluminescent reactions are similar to fluorescent reactions in that the measured signal is emitted light generated by the relaxation of an excited molecule. The difference between fluorescence and chemiluminescent or bioluminescent reactions is the method by which the molecule becomes excited. In fluorescence, adsorbed light energy raises the molecules to the excited state. In chemiluminescent and bioluminescent reactions, the energy is supplied by a chemical reaction. If the reaction occurs *in vivo*, it is classified as a bioluminescent reaction. Reactions that require reactants or conditions not found in biological systems are referred to as chemiluminescent reactions.

Although many organisms produce bioluminescent reactions, probably the best known is the firefly reaction. In that reaction, the substrate firefly luciferin is oxidized by the enzyme luciferase in the presence of magnesium ion and oxygen ATP, yielding a photon. This is an extremely sensitive reaction and can be used to measure ATP at the femtomole level. The quantum efficiency of this reaction approaches unity because each molecule will emit one photon of light for each photon absorbed. This and other bioluminescent reactions have found limited applicability in the immunoassay field because the reagents are expensive and the reaction works best with small molecules.

Chemiluminescent reactions have been far more successful in immunoassay applications because they are less expensive and easier to work with in assay systems. Chemiluminescent reactions never approach the efficiency of the firefly reaction, but modern photon counting equipment has been able to compensate for this deficiency. The most commonly used chemiluminescent molecules are the luminals, acridinium esters, and dioxetanes.

Luminal

Luminal, its isomer, and derivatives belong to the class of compounds known as aromatic hydrazides. When these compounds are oxidized, they emit light. This reaction usually is carried out with a catalytic agent such as peroxidase in an alkaline medium. When luminal is conjugated to one of the reactants in the immunoassay, the efficiency of the light production decreases significantly. Therefore, one of the more successful designs has been to conjugate the peroxidase to one of the reagents and then use luminal as an indicator of the peroxidase. Competitive and noncompetitive solid-phase systems are the most commonly employed designs using luminal.

Acridinium Esters

Acridinium esters are perhaps the most common of the chemiluminescent labels in use today. These compounds can be oxidized without the requirement of a catalyst. The efficiency of this reaction is greater than that of luminal. For these reasons, acridinium esters form the basis of many of today's automated immunoassay systems. Acridinium esters can be conjugated to large or small molecules without significant loss of light-emitting efficiency. They work equally well in most assay system designs; thus, instrument considerations usually dictate the type of assay system.

Dioxetanes

Dioxetanes are thermally unstable molecules that stabilize by thermal cleavage with the emission of light. These compounds can be made stable by adding a phosphate group to the phenyl ring. Under these conditions, an immunoassay can be designed that uses alkaline phosphatase conjugated to one of the reactants. Then, the phosphorylated dioxetane can be added as an indicator. Once the phosphate group is removed from the dioxetane by the alkaline phosphatase, the compound reverts to its thermally unstable condition and will emit light. Enhancers, which are essentially carefully chosen inner filters, can be used to lengthen the time that the light is emitted from the reaction.

The instrumentation required for chemiluminescence is a photometer with special adaptations. Generally, the instrument is equipped with a reagent injection system so reagents can be added while the reaction cuvette is in front of the detector. The detector is usually a high-efficiency photomultiplier tube that counts photons for a prescribed time and then employs some type of averaging algorithm to quantitate the signal. If enhancers or other types of chemistry are used to extend the light emission, less stringent requirements are placed on the photomultiplier tube.

Data Reduction

In the early days of radioimmunoassay, data reduction was a do-it-yourself project. Investigators did what they thought was appropriate to calibrate an assay and in turn to calculate results on unknown specimens. In addition, an effort was made to glean appropriate quality control indicators from these numerical procedures. The literature was replete with reviews of data reduction schemes, exploring the strengths and weaknesses of the proposed systems. The evolution of both immunoassays and data reduction systems has evolved to a point where there is more agreement on the approaches that should be used as well as significant improvement in instrumentation, desktop computers, and software. That is not to say, however, that the procedure for data reduction and quality control should be accepted from the vendor without a thorough understanding of the procedures and approaches used. All data reduction packages share a similar objective, which is to make an accurate and precise estimate of the concentration of an analyte in a given solution. Some of the more common approaches to achieving this objective are described below.

One of the first questions asked about a new assay is, “What is its limit of detection or sensitivity?” Because of the number of steps in most assays and the imprecision associated with any physical measurement process; there is a degree of uncertainty associated with each calculated result. The appropriate way to determine the limit of detection is to evaluate a zero concentration sample, calculate the statistical mean and standard deviation, and use this information to determine the analyte concentration indistinguishable from zero. A simple and practical approach to accomplishing this is to make multiple measurements on several concentrations of analyte within the range of interest, calculate the mean and standard deviations for each of these concentrations, and plot the standard deviations versus the concentration. Extrapolation of this line to zero concentration gives an appreciation for the smallest detectable analyte concentration that differs from zero (Fig. 12.9). This plot, also referred to as a precision profile, can be informative in terms of the working range of the assay. An effective approach is to plot CV percentage versus several analyte concentrations. Then decide on the maximum allowable CV the procedure can tolerate. Draw a line parallel to the x-axis at this point. The intersection with the original plot yields the working range. By evaluating the precision profile, an investigator can determine if at any concentration the standard deviation of the measurement exceeds an acceptable limit.

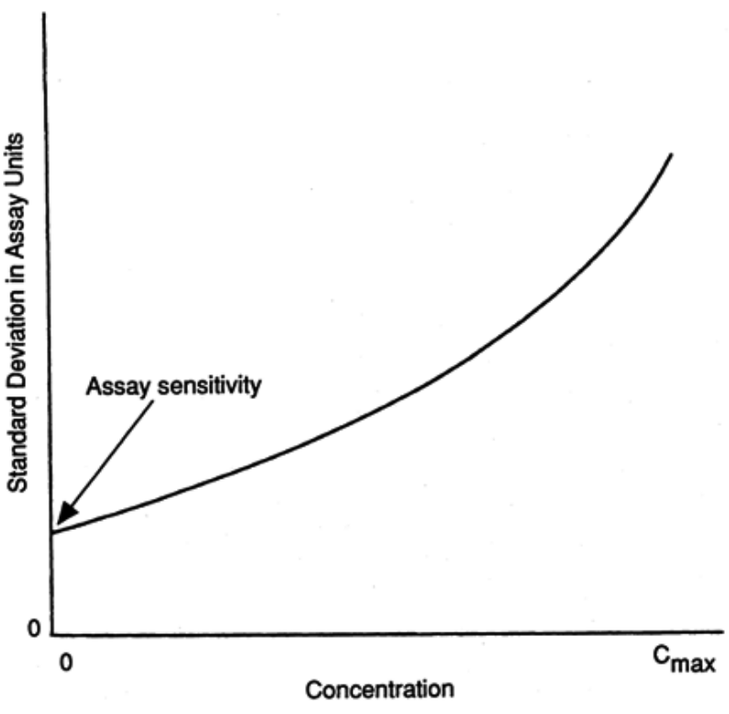


FIGURE 12.9. Precision profile diagrams. The standard deviation for several analyte concentrations has been plotted versus the analyte concentration. The units are the same as the analyte units. The line is extrapolated to zero analyte concentration and yields the level of analyte that does not differ from zero concentration in the assay. If 1 SD is used for the plot, the actual sensitivity will be two or three times the intercept value, depending on the level of confidence required.

Once a series of measurements has been made under the assay conditions, the investigator is faced with the dilemma of

selecting the best approach for using this information to calculate the concentration of analyte in other specimens. This calculation procedure generally is a part of automated methodologies, most of which present the user with a menu of data reduction choices. Manual calculations and plotting of data for the purposes of calculating the concentration of analyte in specimens is less common in the modern laboratory than it was at the time of the original radioimmunoassay. Manual plotting of data and the subsequent extrapolation of concentration for unknowns is an error-prone process. Each individual will connect the points in a slightly different manner and will read the resultant concentrations for unknowns the same way. In addition, manual data reduction is a slow and expensive process that is more important as a teaching procedure than as a daily analytic procedure.

Data reduction methods can be grossly divided into two types. The first are the point-to-point methods in which an attempt is made to draw a smooth line through the data points progressing from one point to the next. Approaches that utilize this method are point-to-point, spline, and polynomial interpolative methods. Computers can be programmed to generate these lines very quickly, but the computer does not correct the underlying weakness of these approaches. That weakness is that there are errors inherent in the values assigned to each concentration and the assumption that all concentrations throughout the working range behave identically. Once a line has been generated through these concentration points, it is assumed that the intervening concentrations, which have not been measured, will fall on that line. One cautionary note when using the polynomial or spline-fit procedures. If the data are “noisy,” the line fitting algorithms will generate a line that has so-called “hooks” because the computer has generated multiple inflection points to go through all of the data. If this situation is severe enough, a single response on an unknown can generate multiple potential concentrations from the curve. If this type of plot is used, it is essential that the curve be visualized to ensure that the standard concentrations are following an anticipated response line. Figure 12.10 illustrates a spline fit to actual assay data.

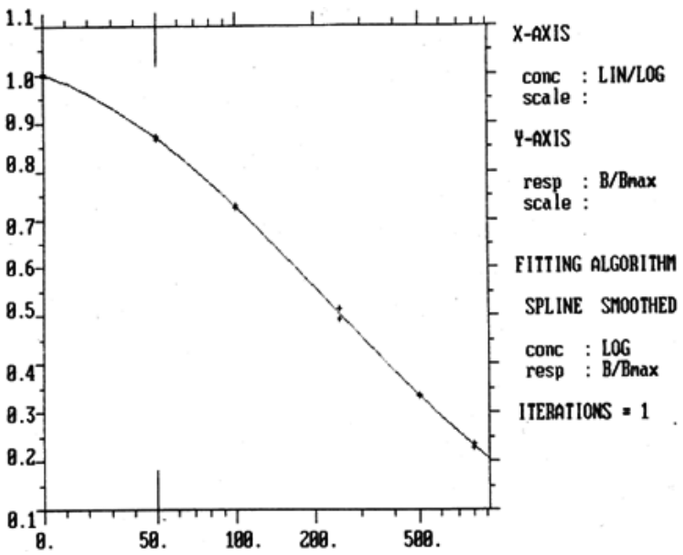


FIGURE 12.10. Spline fit to assay data. The data represent a spline fit to actual laboratory data. The x-axis is Lin/Log because of a zero concentration of analyte that does not have a log value. The response plotted on the y-axis is percent of maximum response. In this example, the spline fit required one iteration to develop an acceptable curve.

Regression methods are the other approach to data reduction. Regression methods fit a line through the data points in such a way that each data point is a minimum distance from the line. In other words, if you were to measure the distance from each data point to the line, that would be a minimum compared with any other line that could be drawn in relation to those data points. The potential problem with this approach is that it treats every data point equally. Therefore, if the method has greater imprecision at the ends of the curve or if there are outliers, the standard regression analysis fails to recognize this and treats all points equally. The method most software packages use to overcome this weakness is that of weighting. In a weighted regression, each point is adjusted by a weighting factor in accordance with its reliability. The points in a region of the curve where the error is greater and the reliability of the measurement is lower are given less influence in the calculation than those points that lie in regions of the curve where the certainty of the measurement is greater. These weighting factors can be determined by calculating the variance of multiple measurements or can be generated from mathematical models if there is sufficient information known about the behavior of the curve under study. Figure 12.11 shows a regression and a weighted regression.

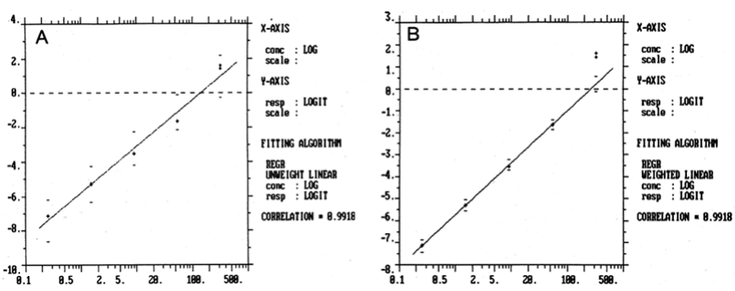


FIGURE 12.11. The figure presents an unweighted (**A**) and a weighted (**B**) linear regression. The x-axis is in log of concentration, and the y-axis is in log of the response. Notice the relative error around the line is smaller in the weighted regression, and the point at a concentration of 200 units has less effect on the line because of the unreliability of the measurement at that concentration.

A popular family of algorithms for fitting data is the logistic fit equations. These are derived from statistical functions that have been adapted to immunoassay data. Equation 12.5 shows several parameters that are related to the analyte concentration, represented by *x*, and the response measured, represented by *y*.

$$y = \frac{a - d}{1 + (x/c)^b} + d \qquad (12.5)$$

The four parameters *a*, *b*, *c*, and *d* have interesting interpretation in terms of the curve generated. Parameter *a* is the predicted response at zero analyte concentration; *d* is the predicted response at infinite analyte concentration; *c* is the inflection point on the sigmoidal curve; and *b* is the slope at that inflection point. Using this equation places no demand on any of the equation parameters. Only the concentration of the solutions and the response associated with those concentrations are required. The computer then assigns a first estimate to the value of the parameters and begins solving the equation in an iterative process. This means that the computer solves the equation, evaluates the results, makes a small change in the parameters, and continues this process until successive solutions begin to converge on the same line through the data points. This is a very effective approach to fitting a curve to experimental data, and in principle, there should be only one best-fit line that goes through a given set of data points. Because of the differences in assumptions, approaches, and software programs, it is possible to have variations from program to program. However, once it has been determined that a four-parameter logistic fit procedure is appropriate, this approach yields some additional advantages. First, there are the four parameters that can be used as quality control points for

the long-term performance of the assay. Significant changes in endpoints or slopes indicate that the assay should be reviewed. In addition, concentrations can be read from the curve at 20%, 50%, and 80% as another quality control monitor. The one precaution that should be kept in mind when employing four-parameter logistic fit equations is that you are at high risk once you exceed your calibrating solutions. That is, to extrapolate the concentration of an analyte in an unknown solution beyond the highest or lowest concentration of a calibrator leaves the analysis vulnerable to significant error. The reason for this is the four-parameter logistic fit sacrifices the ends of the curves, where the variance or error may be considerable.

Even in this day of computerized calculations, it is critical that the operator know the algorithm and assumptions used to calculate calibration curves. This information should not only be logical but should employ basic statistical concepts wherever possible. Once this information is known to the researchers, they should process their laboratory data, control specimens, and previous assays and evaluate the quality of the calculation in their hands. There are numerous software packages, both stand-alone and incorporated into instrumentation that will calculate results for any given immunoassay. The process of choosing the correct software is primarily one of evaluating the candidate products and being comfortable with the assumptions and procedures used.

Automation in Immunoassay

As might be expected, immunoassay is following a course close to classic clinical chemistry; that is, initially there were laborious manual methods designed to measure one analyte at a time in batch fashion. This was followed by assisted manual procedures in which dilutors, pumps, and spectrophotometers took over some of the steps in the assay system. Nevertheless, it remained essentially a manual procedure. Next came automation wherein specimens were prepared and presented to the system at one end and completed results were returned at the other, with relatively little hands-on intervention. In the past several years, immunoassay has reached the point of full automation. Several companies are now marketing instrumentation that takes advantage of the nonisotopic methodologies. These instruments are complete with bar-code readers, reagent monitors, and total software packages for calculating results, monitoring quality control, and generating management reports. In addition, most have the capability of being interfaced with the main laboratory computer for bidirectional communication. Companies such as Ciba Corning, Cirrus Diagnostics, Becton Dickinson, Technicon, Boehringer Mannheim, Biotrol, Abbott, and Baxter, along with several others, have marketed instruments described as automated immunoassay systems. Each of these instruments takes advantage of a slightly different technique and solves a particular problem. Several of these same companies are now integrating immunoassays on “conventional” chemistry instrumentation or have research projects to combine clinical chemistry and immunoassay into one instrument (12).

In fact, before automated immunoassays have fully matured, research and development is well established in the field of miniaturization. These techniques will define the future of laboratory methods and testing, including immunoassays (13, 14, 15). Several authors have speculated that microchips will be capable of measuring hundreds or thousands of analytes at a time on volumes in the microliter to nanoliter range (12, 13, 14). The advent of such technology will change the approach to point-of-care testing, outpatient testing, laboratory structure and size and the cost structure of laboratory medicine.

REFERENCES

1

. Burnet FM. A modification of Jerne's theory of antibody production using the concept of clonal selection. *Aust J Sci* 1957;20:67–69.

2. Kühler G, Milstein C. Continuous cultures of fused cells secreting antibody of predetermined specificity. *Nature* 1975;256:495–497.

3. Chappey ON, Sandouk P, Scherrmann JMG. Monoclonal antibodies in hapten immunoassays. *Pharm Res* 1992;9(11):1375–1379.

4. Moyle WR, Lin C, Corson RL, Erlich PH. Quantitative explanation for increased affinity shown by mixtures of monoclonal antibodies: importance of circular complex. *Mol Immunol* 1983;20:439–452.

5. Tedford MC, Stimson WH. Molecular recognition in antibodies and its application. *Experientia* 1991;47:1129–1138.

6. Laurell CB. Quantitative estimation of proteins by electrophoresis in agarose gel containing antibodies. *Anal Biochem* 1966;15:45.

7. Apler CC, Johnson AM. Immunofixation electrophoresis: a technique for the study of protein polymorphism. *Vox Sang* 1969;17:445.

8. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. *J Clin Invest* 1960;39:1157–1175.

9. Ekins R. Merits and disadvantages of different labels and methods of immunoassay. In: Voller A, Bartlett A, Bidwell D, eds. *Immunoassay for the 80's*. Baltimore: University Park Press, 1981.

10. Jackson T, Marshall N, Ekins R. Optimization of immunoradiometric (labelled antibody) assays. In: Hunter WM, Corrie JET, eds. *Immunoassays for Clinical Chemistry*. Edinburgh: Churchill Livingstone, 1983.

11. Diamandis EP. Immunoassays with time-resolved fluorescence spectroscopy: principles and applications. *Clin Biochem* 1988;21:139–150.

12. Blick KE. Current trends in Automation of Immunoassays. *Jour of Clinical Ligand Assay* 1999;22:1:6–12.

13. Kricka LJ. Miniaturization of analytical systems. *Clin Chem* 1998;44:2008–2014.

14. Ekins R, Chu FW. Multianalyte microspot immunoassay-microanalytical “compact disk” of the future. *Clin Chem* 1991;37:1955–1967.

15. Ekins R, Chu FW. Microspot®, array-based multianalyte binding assays: the ultimate microanalytical technology? In: Price CP, Newmann DJ. eds. *Principles and practice of immunoassay*, 2nd edition. New York: Stockton, 1997:625–646.

SUGGESTED READINGS

Bioluminescence and Chemiluminescence

Abelson JN, et al., eds. *Methods in Enzymology*, volume 184. New York: Academic Press, 1990.

Bonini PA, Banfi G, Murone M. Enhanced chemiluminescence in the measurement of proteins and haptens: evaluation of choriogonadotropin (hCG) and free thyroxin. *J Bioluminescence and Chemiluminescence* 1990;5:193–195.

Colowick SP, ed. Methods in enzymology, volume 133. New York: Academic Press, 1986.

De Boever J, Kohen F, Leyseele D, et al. Isolaminol as a marker in direct chemiluminescence immunoassays for steroid hormones. *J Bioluminescence and chemiluminescence* 1990;5:5–10.

Fert V, Baret A. Preparation and characterization of xanthine oxidase-antibody and -hapten conjugates for use in sensitive chemiluminescent immunoassays. *J Immunol Methods* 1990;131:237–247.

Gorus F, Schram E. Applications of bio- and chemiluminescence in the clinical laboratory. *Clin Chem* 1979;25(4):512–519.

Roche D, Susini de Luca H, Tugendhaft N. Chemiluminescence and radioimmunology compared for 10 allergens. *Clin Chem* 1991;37(3):474–475.

Woodhead JS, Weeks I. Immunochemiluminometric assays based on acridinium labels with a microtiter plate luminometer. *Clin Chem* 1991;37(3):472.

Data Analysis

Ezan E, Tiberghien C, Dray F. Practical methods for optimizing radioimmunoassay detection and precision limits. *Clin Chem* 1991;37(2):226–230.

Haven MC, Orsulak PJ, Arnold LL, et al. Data-reduction methods for immunoradiometric assays of thyrotropin compared. *Clin Chem* 1987;33(7):1207–1210.

Klee GG, Post G. Effect of counting errors on immunoassay precision. *Clin Chem* 1989;35(7):1362–1366.

Miller JJ, Valdes R Jr. Approaches to minimizing interference by cross-reacting molecules in immunoassays. *Clin Chem* 1991;37(2):144–153.

Sadler WA, Smith MH, Legge HM. A method for direct estimation of imprecision profiles, with reference to immunoassay data. *Clin Chem* 1988;34(6):1058–1061.

Seth J. Estimation of sensitivity of immunoassays. *Clin Chem* 1990;36(1):178.

Fluorescence

Calvin J, Burling K, Blow C, et al. Evaluation of fluorescence excitation transfer immunoassay for measurement of specific proteins. *J Immunol Methods* 1986;86:249–256.

Diamandis EP. Immunoassays with time-resolved fluorescence spectroscopy: principles and applications. *Clin Biochem* 1988;21:139–150.

Hemmilä I. Fluoroimmunoassays and immunofluorometric assays. *Clin Chem* 1985;31(1):359–370.

Schwan WR, Waltenbaugh C, Duncan JL. Bacteria as solid phase in a concentration fluorescence immunoassay analysis of antibodies to surface antigens. *J Immunol Methods* 1990;126:247–252.

Immunoassay

Abelson JN, Simon MI, eds. *Methods in Enzymology*, volume 182. New York: Academic Press, 1990.

Gosling JP. A decade of development in immunoassay methodology. *Clin Chem* 1990;36:1408–1427.

Latner AL, Schwartz MK, eds. *Advances in clinical chemistry*, volume 24. New York: Academic Press, 1985.

Miscellaneous

Azimzadeh A, Van Regenmortel MHV. Antibody affinity measurements. *J Mol Recognition* 1990;3:108–116.

Demers LM. Monoclonal antibodies to lutropin: are our immunoassays too specific? *Clin Chem* 1991;37(3):311–312.

Pfund WP, Bourdage JS. The confirmation-sensitive immunoassay: a membrane based ELISA system for identifying antibodies sensitive to alterations of protein conformation. *Mol Immunol* 1990;27:495–502.

Sorensen K, Brodbeck U. Assessment of coating-efficiency in ELISA plates by direct protein determination. *J Immunolog Methods* 1986;95:291–293.

Vilja P. Flexibility of noncompetitive avidin-biotin immunoassay: immunoassay of Lutropin applied to different signal detection systems. *Clin Chem* 1990;36:1897–1901.

Nephelometry

Cuilliere ML, Montagne P, Bessou TH, et al. Microparticle-enhanced nephelometric immunoassay (Nephelia®) for immunoglobulins G, A, and M. *Clin Chem* 1991;37:20–25.

APPENDIX GLOSSARY

GLOSSARY

Affinity

Describes the strength of binding of a single binding site on an antigen with its corresponding antibody. For well-defined

systems, this can be equated to the association constant of the reaction.

Antibody

An immunoglobulin class protein that is produced by a special class of lymphocyte in response to its stimulation by foreign material.

Antigen

Any material capable of reacting with an antibody to produce an antibody-antigen complex.

Avidity

Describes the stability of antigen-antibody complexes. It is the net of all the forces of all the binding sites on an antigen and generally is greater than the simple sum of these forces.

Bioluminescence

A special category of chemiluminescence that takes place in living organisms. Photons are emitted as a result of biologically mediated reactions. Characteristically, these reactions have very high quantum yields.

Chemiluminescence

A reaction in which the product molecules are in an excited state as a result of the energy released by the reaction. Relaxation of these excited molecules results in the emission of a photon. Generally, quantum yields are low.

Detection limit

The smallest amount of analyte that is measurable above zero. Usually arrived at by measuring a zero response and taking the mean of that response plus three standard deviations.

Fluorescence

The emission of a photon from a molecule that was raised to an excited state by external radiation.

Hapten

A small organic molecule that represents a single antigenic determinant. Usually is not capable of generating an immune response by itself.

Heterogeneous assay

Assay that requires the separation of bound and free species before the analytic measurement is made.

Homogeneous assay

Assay that is designed to have the analytic measurement made in the presence of both bound and free species.

Immunogen

Any substance capable of stimulating lymphocytes to produce antibodies.

Nephelometry

A technique that measures the amount of scattered light from a sample at angles other than 0°.

Sensitivity

Synonymous with detection limit.

Specificity

A characteristic of antibodies that denotes the antibody's ability to recognize a specific antigenic determinant in the presence of similar structures. The greater the specificity the less cross-reactivity.

Turbidimetry

A technique that measures the reduction in incident light after the incident light passes through a sample. Angle of measurement is 0°.

13

The Plasma Proteins

Alexander Kratz

Elizabeth Lee-Lewandrowski

Kent B. Lewandrowski

The plasma proteins are a diverse group of molecules that perform a variety of functions. More than 500 plasma proteins have been identified, and new ones continue to be discovered. Some plasma proteins such as albumin and transferrin are routinely measured in the clinical laboratory, whereas others are of interest only in specialized situations. While it is assumed that all proteins perform some function in the body, in many cases the precise role has not been defined. With the growth of more sophisticated and sensitive methods to study these molecules (1), new applications for protein analysis in clinical practice will surely arise.

Some of the plasma proteins can be conveniently grouped as related molecules that collectively carry out a specific function. Examples include the immunoglobulins, the complement proteins, the coagulation cascade, lipoproteins, and a variety of protein hormones and hormone-binding globulins. These are covered in other chapters and are mentioned only briefly here. This chapter is concerned with the physiology, pathology, and laboratory assessment of the remaining plasma proteins with an emphasis on those that are important in clinical medicine.

True plasma proteins should be distinguished from proteins that occur only incidentally in the blood. For example, a number of enzymes may be released from damaged cells and appear in plasma but perform no known function outside of the cell. Other proteins are found in plasma only during certain pathological states. In contrast, true plasma proteins are specifically synthesized for release into the blood, where they carry out their respective functions.

Despite their name, most of the plasma proteins are not confined to the vascular compartment (2). In the capillary vessels, proteins escape from the blood either by active transport mechanisms or by diffusion through the junctions between endothelial cells. In some cases, the extravascular fraction exceeds the amount in the intravascular pool. Many of these proteins carry out part of their function in the interstitial space, ultimately returning to the vascular compartment by the lymphatics. Some plasma proteins cross the blood–brain barrier and thus enter the cerebrospinal fluid (CSF). Small amounts are found in virtually all other extracellular fluids.

Several systems have been used to classify the plasma proteins. Historically, albumin was distinguished from the globulins (nonalbumin proteins) based on its greater solubility in water (2). However, this distinction is not absolute because some other proteins are also soluble in low ionic-strength solutions. Alternatively, proteins may be classified according to function (Table 13.1). For example, some proteins perform mainly transport functions for metals (transferrin), lipids (lipoproteins), or hormones (thyroid-binding globulin). Others are involved in the clotting cascade (fibrinogen) or function as enzyme inhibitors [α_1-antitrypsin (AAT)]. Some proteins perform more than one function. Albumin, for example, serves as a carrier protein for bilirubin, some hormones, and inorganic ions. It also provides a protein source for cells and contributes the major portion of the plasma colloid oncotic pressure (COP), which is essential for normal hemodynamics.

TABLE 13.1. FUNCTIONAL CLASSIFICATION OF SELECTED PLASMA PROTEINS

Proteinase inhibitors
α-Antitrypsin
Antithrombin III
α-Antichymotrypsin
α_2-Macroglobulin
Antiplasmin
Inter-α-trypsin inhibitor
Carrier/transport proteins
Retinol binding protein
Gc-globulin (vitamin D binding protein)
Lipoproteins
Hormone binding globulins
Prealbumin
Iron transport and metabolism
Transferrin
Haptoglobin
Hemopexin
Mixed function proteins
Albumin
Immune function and host defense
Immunoglobulins
Complement proteins
C-Reactive protein
Blood clotting
Clotting factors
Fibrinogen
Fetal proteins
α-Fetoprotein
Carcinoembryonic antigen
Unknown or controversial
Ceruloplasmin
α-Acid glycoprotein (orosomucoid)
β_2-Microglobulin

Plasma proteins are often classified based on their electrophoretic mobility on cellulose acetate or agarose gel. Electrophoresis separates proteins into five regions designated albumin α_1, α_2, β, and γ (Fig. 13.1). Individual proteins characteristically migrate in one of these regions (Table 13.2). The albumin band is relatively homogeneous, while the remaining regions contain many different proteins and are simply groups of proteins with similar mobilities.

With some exceptions (e.g., immunoglobulins, protein hormones, and some clotting factors), most of the plasma proteins are synthesized in the liver. The liver is also the major site of plasma protein catabolism. Most of the plasma proteins are synthesized on the rough endoplasmic reticulum of the liver hepatocytes. After assembly of the amino acid backbone, a variety of posttranslational modifications may occur, including proteolytic removal of proprotein fragments, phosphorylation, and the addition of sugars and other prosthetic groups. Ultimately, the proteins pass through the Golgi apparatus and are packed into secretory vesicles for release into the blood. Further details of protein structure and synthesis may be found in standard biochemistry textbooks.

Hepatocytes are capable of synthesizing more than one protein simultaneously. The balance between synthesis and degradation maintains the level of plasma proteins within a relatively narrow range. Most proteins are regulated independently of the others. The level of the individual protein in plasma serves as the signal for regulation, and the rates of synthesis and catabolism are the mechanism for modulating its concentration (2). Details of the regulatory steps for specific proteins are provided under separate headings later in this chapter.

Given the diversity of the plasma proteins, it is not surprising that a variety of diseases and physiological alterations may affect protein levels. Plasma protein abnormalities arise from one or more of the following:

- Congenital abnormalities affecting a specific protein (e.g., analbuminemia, AAT deficiency)
- Acquired abnormalities affecting a specific protein (e.g., transferrin in anemia, haptoglobin in hemolysis)
- Alterations affecting multiple proteins reflecting variations in the physiologic state (e.g., age, race, gender, and pregnancy)
- Alterations affecting multiple proteins secondary to disease (e.g., cirrhosis, acute-phase response, and protein-losing syndromes)

In this chapter, we begin with a discussion of the total plasma protein followed by a description of the acute-phase response. Then the individual plasma proteins are considered. In the final section of the chapter, several methods for protein separation, identification, and quantitation are described. Assays of specific proteins are described under individual headings.

TOTAL PROTEIN

Part of "13 - The Plasma Proteins"

The total serum protein concentration is one of the more common measurements in the clinical laboratory. Several factors affect the protein concentration, including alterations in fluid balance, changes in synthesis or catabolism, and protein losses.

TABLE 13.2. CLASSIFICATION OF SELECTED MAJOR[a] AND MINOR PLASMA PROTEINS BY ELECTROPHORETIC MOBILITY

Albumin zone
Albumin
Prealbumin
α Zone
α_1-Antitrypsin
High-density lipoprotein (α-lipoproteins)
α-Antichymotrypsin
Orosomucoid
α-Fetoprotein
α_2 Zone
α_2-Macroglobulin
Haptoglobin
Ceruloplasmin
Gc-Globulin
β Zone
Low-density lipoprotein (β-lipoproteins)
Transferrin
C3
β_2-Microglobulin
Hemopexin
Fibrinogen (may be in γ zone)
γ Zone
Immunoglobulins
C-Reactive protein
Fibrinogen
Lysozyme

[a]In boldface.

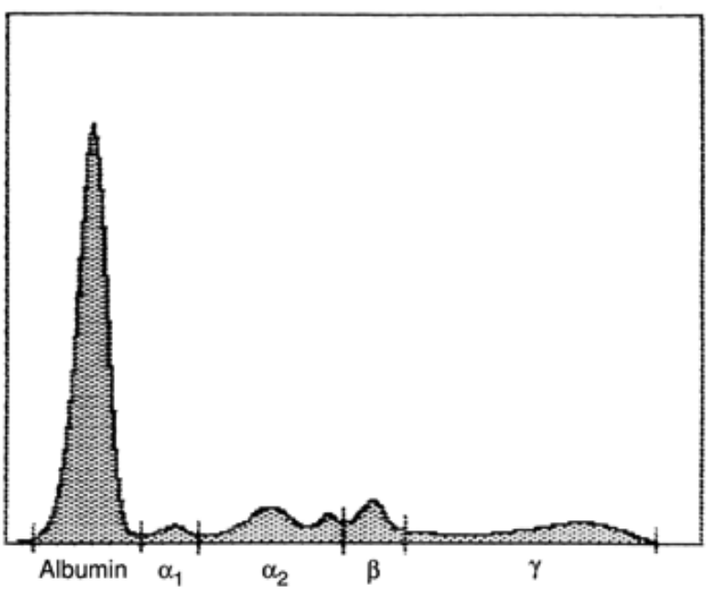


FIGURE 13.1. Densitometric scan of a normal serum protein electrophoresis pattern showing the relative position of the albumin, α_1, α_2, β, and γ regions.

Measurement of total protein is therefore a relatively sensitive but nonspecific screening test of overall health status.

The total protein may be measured on either serum or plasma. Serum samples are generally preferred because they are more convenient to use on automated equipment (3). If plasma is used, the total protein is 3% to 5% higher owing to the contribution of fibrinogen. Several preanalytic errors may alter the serum protein value (3). Prolonged tourniquet application during venipuncture results in hemoconcentration, which can increase the total protein by as much as 5 g/L. Values in recumbent subjects may be 10% lower than in ambulatory individuals. There is also a normal diurnal variation of 4 g/L and a seasonal fluctuation with peak levels in November and lowest values in June. Exercise increases the serum protein concentration by as much as 10%, but this effect is transient.

The total serum protein includes both albumin (60%) and the globulins (40%). All contribute some fraction to the COP of plasma, which is essential for normal hemodynamics. Unlike sodium salts and glucose, the movement of proteins from plasma to interstitial fluid is limited (4). The contribution of proteins to the plasma osmotic pressure tends to retain water in the vascular space. The osmotic pressure generated by the plasma proteins is called the COP, or simply oncotic pressure. Because of its abundance and small size, albumin is responsible for 80% of the total COP.

The movement of fluid from the capillaries to the interstitial space is governed by the balance between four forces. The plasma oncotic pressure acts to hold fluid in the vascular space. This force is counterbalanced by the capillary hydrostatic (hydraulic) pressure, which tends to force water out of the vascular compartment. The interstitial oncotic pressure is derived from the proteins and proteoglycans of the interstitial space and acts to draw fluid out of the capillary vessels. The small interstitial hydrostatic pressure is usually negative, and thus it also pulls fluid out of the vascular compartment. The sum of these four forces varies in different tissues (4), and water may move either into or out of the plasma compartment depending on the balance between the forces. Ordinarily the net effect is a 0.3- to 0.5-mm Hg gradient favoring movement out of the capillaries. Conversely, in the postcapillary venules, the forces normally allow return of some of the water to the circulation. The remainder is returned to the vascular space by the lymphatics. A decrease in the total protein concentration will decrease the plasma COP, which may result in edema. However, because hypoproteinemia will also decrease the amount of protein in the interstitial space, the decrease in plasma oncotic pressure will be partially offset by a decrease in the interstitial oncotic pressure. In general, hypoproteinemia must be severe to result in edema.

The total protein concentration may be altered by changes in fluid balance (plasma water). Dehydration causes a proportional increase in all the serum proteins. This occurs with decreased water intake or with increased water losses such as excessive sweating, diarrhea, vomiting, salt-losing syndromes, and osmotic or drug-induced diuresis. Conversely, volume expansion causes a proportional decrease in serum proteins. Examples include the administration of volume expanders (such as dextran), excessive administration of intravenous fluids, pregnancy, and salt-retention syndromes. These changes do not affect the actual body pools of serum proteins, and correction of the underlying fluid imbalance will correct the total protein concentration.

Changes in the amounts of plasma proteins may result from alterations in synthesis or catabolism or from protein losses. Increased synthesis of some proteins is seen in the acute-phase response to infection or tissue injury, but this is offset to some degree by a concomitant decrease in albumin synthesis. The use of steroids, including oral contraceptives, will increase protein synthesis and may increase the total protein concentration (3). Marked hyperproteinemia in the absence of dehydration usually reflects increased synthesis of γ-globulins. Hyperproteinemia owing to increased albumin synthesis is quite rare if it occurs at all.

Hypoproteinemia can be caused by decreased protein synthesis and is seen in malnutrition and in chronic liver disease. Protein calorie malnutrition deprives the liver of the necessary substrates to make plasma proteins and is a common problem in hospitalized patients (5). In chronic liver disease, the protein synthetic machinery is disturbed, which also results in hypoproteinemia. However, this is often accompanied by increased (extrahepatic) synthesis of γ-globulins, which may normalize or occasionally even increase the total protein value. Increased protein catabolism is seen in a variety of inflammatory and neoplastic conditions.

Excessive protein losses can occur through the kidney, the gastrointestinal tract, or the skin. Examples include the nephrotic syndrome, inflammatory bowel disease, extensive burns, and any severe exudative process. The most severe losses are seen in the nephrotic syndrome, in which a number of plasma proteins, especially albumin, are passed through the glomerulus into the urine. In severe cases of nephrotic syndrome, the serum albumin level may fall to less than 5 g/L (3).

Changes in the total serum protein may result from changes in albumin, globulins, or both. As mentioned earlier, a change in one protein may be offset by a change in the opposite direction of another. For this reason, it may be useful to determine the ratio of the albumin concentration to the globulin concentration (A:G ratio). The ratio may be markedly abnormal despite a normal total protein, such as in the acute-phase response, in which increased globulins are associated with a decreased albumin concentration.

Methods for measuring the total protein are discussed at the end of this chapter.

ACUTE-PHASE REACTANTS

Part of "13 - The Plasma Proteins"

The acute-phase reactants (APRs) are a group of proteins whose plasma concentration changes in response to a variety of inflammatory states including infection, surgery, trauma, myocardial infarction, malignancy, and any condition associated with tissue necrosis. Collectively, the changes in plasma proteins are referred to as the acute-phase response, even though they may accompany both acute and chronic inflammatory disorders (6). Some of these proteins are called positive APRs, meaning that their plasma levels increase, while others are designated negative APRs

to indicate that their levels decrease (Table 13.3). The acute-phase response is regulated by cytokines. Interleukin-6 (IL-6) is the major stimulator of most acute-phase proteins; other cytokines produced during inflammation, like interleukin-1ß(IL-1ß), tumor necrosis factor-α (TNF-α), interferon gamma (IFN-γ), and transforming growth factor ß (TGF-ß), upregu-late subgroups of the APRs (6). The major sources of these cytokines are macrophages and monocytes at sites of inflammation.

TABLE 13.3. MAJOR ACUTE PHASE REACTANTS

Positive	Negative
α_1-Antitrypsin	Albumin
α-Antichymotrypsin	Prealbumin
Haptoglobin	Retinol binding protein
Ceruloplasmin	Transferrin (rises in late acute phase)
Fibrinogen	
C3	
C-Reactive protein	
Hemopexin	
Serum amyloid A protein	

Measurement of APRs, especially C-reactive protein (CRP), may be useful to detect and follow patients with acute inflammatory

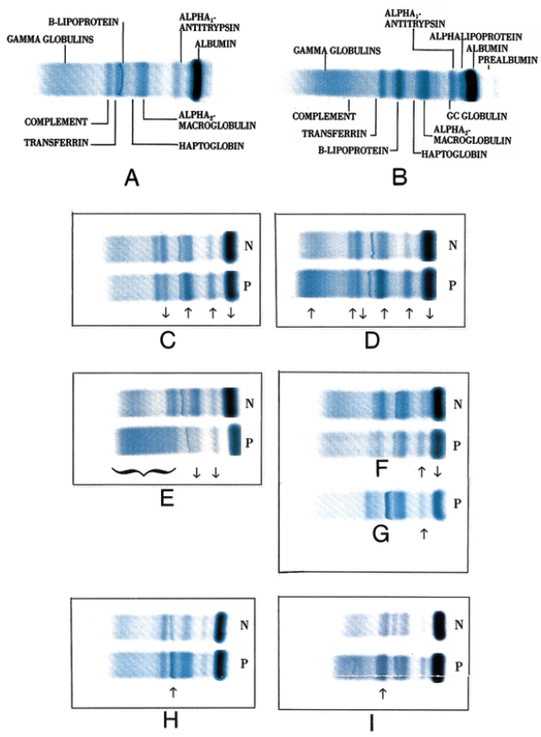


FIGURE 13.2. Representative patterns of serum proteins seen on serum protein electrophoresis. **A**, normal pattern; **B**, normal pattern on high-resolution electrophoresis; **C**, acute inflammatory pattern; **D**, chronic inflammatory pattern;**E**, cirrhotic pattern; **F**, Nonselective protein loss. **G**, nephrotic syndrome; **H**, hyperbetalipoproteinemic pattern; **I**, iron deficiency anemia. (Courtesy of Michael Burke, Beckman Instruments, Inc., Beckman Paragon Serum Protein Electrophoresis Patterns.)

disorders. However, these changes are nonspecific and provide no information on the cause or source of the inflammation. Elevations in the white blood cell count and erythrocyte sedimentation rate (ESR) provide similar information and have been considered more practical because they are easier to measure and are more readily available. However, it should be noted that the increase in the ESR seen in inflammatory states directly results from changes in the APRs, most notably an increase in fibrinogen.

The advent of more sensitive assays for CRP has aroused new interest in the potential applications of APRs to clinical practice. Some reports suggest that assay of CRP may be superior to the ESR in a variety of clinical situations (7). Serial measurements may provide an index of disease activity and are useful to assess the response to therapy. Potential applications of APRs include monitoring postoperative patients for infectious complications, detection of occult infections, monitoring disease activity in patients with rheumatic diseases and inflammatory bowel disease, and detecting renal allograft transplant rejection. Further studies are needed to define more fully the role of APRs in clinical practice.

The changes in plasma proteins seen during the acute-phase response potentially serve a variety of functions. Among the positive APRs are the protease inhibitors AAT and α_1-antichymotrypsin. These proteins protect the body from indiscriminate proteolysis by proteases released from leukocytes and macrophages at sites of tissue injury and repair. Other proteins such as haptoglobin and hemopexin protect against reactive oxygen species (6). CRP appears to function as an immunomodulator by activating the complement system as well as neutrophils and monocyte-macrophages (7). Conversely, transferrin is a late APR, its level rising in plasma only after the other APRs have reached their maximum. It is thought that the elevation in transferrin is probably a secondary event brought about by the decrease in plasma iron associated with inflammation. Wound healing is affected by the increase in fibrinogen, which causes endothelial cell adhesion, spreading, and proliferation (6). The function of some APRs such as α_1-acid glycoprotein is not known. The role of the negative APRs is more difficult to conceptualize. The decrease in albumin during the acute-phase response may serve to divert the protein synthetic machinery of the liver away from production of this nonessential protein. In theory, this would facilitate the synthesis of critical elements needed for the inflammatory response.

The individual APRs respond at different rates, and the factor by which their concentration changes in plasma varies. Within hours, CRP and α_1-antichymotrypsin rise. Soon thereafter, α_1-acid glycoprotein increases, followed by AAT, haptoglobin, C4, and fibrinogen. Later C3, ceruloplasmin (CER), and transferrin become elevated. Within 2 to 5 days, most of the APRs have reached their maximum (8). Quantitatively, the most striking elevation is seen with CRP, which may increase by a factor of 3,000. The other APRs generally show much more modest increases (factors of 2 to 4). The early rise in CRP together with the large factor by which it increases makes it the most useful APR to measure in the clinical laboratory.

Although APRs may be measured individually, the overall pattern of the acute-phase response is most easily appreciated on serum protein electrophoresis (SPE) (Fig. 13.2C and Fig. 13.2D). The total protein is usually normal, owing to a decrease in albumin with a concomitant marked increase in the α_1 and α_2 globulins. CRP may appear as a distinct band in the slow γ to mid-ß region. The γ-globulin fraction is often normal during the early acute phase. In the course of an infectious process, oligoclonal bands may be seen, followed later by a diffuse polyclonal increase in the γ-globulins.

Patients with severe liver disease, malnutrition, or protein-losing syndromes often show an absent or blunted acute-phase response. Individuals with congenital protein deficiencies may fail to increase a specific protein or may exhibit only a partial response. This can have the effect of normalizing the level of the deficient protein and masking the underlying congenital defect.

INDIVIDUAL PROTEINS AND DISEASE STATES

Part of "13 - The Plasma Proteins"

Prealbumin and Retinol-Binding Protein

Prealbumin (MW 55,000) is a tryptophan-rich tetrameric non-glycosylated protein named for its anodal migration relative to albumin on protein electrophoresis. Synonyms for prealbumin include transthyretin and thyroid-binding prealbumin. The only known function of prealbumin is as a carrier protein for thyroid hormone and, together with retinol binding protein (RBP), for vitamin A. Prealbumin binds approximately 20% of the plasma thyroxine (T_4) and lesser amounts of triiodothyronine (T_3). However, its role in thyroid-hormone transport is relatively minor because most T_4 and T_3 are bound to thyroid-binding globulin, which has a much greater affinity for these hormones. A genetic variant of prealbumin has been described with higher than normal affinity for T_4. This causes an elevation in total T_4, but the patients are clinically euthyroid.

Prealbumin forms a 1:1 stoichiometric complex with the vitamin A carrier protein RBP (MW 21,000). This interaction stabilizes the RBP-vitamin A complex and prevents urinary losses of the vitamin (because free RBP is filtered by the glomerulus).

More than 50 variants in the prealbumin gene have been described. Most of these variants are associated with tissue deposition of amyloid fibrils, especially in heart and neurons (9).

Prealbumin and RBP are synthesized by the liver. Together with albumin and transferrin, these proteins can be used to assess protein status in patients with protein-calorie malnutrition. Both prealbumin and RBP have short half-lives (1.9 days and less than 12 hours, respectively). In addition, the body pool of these two proteins is small compared with albumin. Consequently, a deficiency in protein intake will produce a noticeable fall in prealbumin and RBP levels long before albumin and transferrin are affected. Although albumin and transferrin have traditionally been used to assess protein status, measurement of prealbumin and RBP may provide a better index of acute changes. For this reason, they are also useful in monitoring the adequacy of protein replacement therapy (5,10).

Prealbumin levels are significantly decreased in several liver diseases owing to impaired synthesis. Unlike serum enzymes, which primarily indicate hepatocellular damage, prealbumin levels reflect hepatic synthesis and can serve as an index of liver

function (11). However, prealbumin is also a negative APR and is thus not a specific indicator of liver function.

Prealbumin may be increased in patients receiving steroids, in pregnancy, in chronic renal failure, and in Hodgkin's disease.

Because of the low level of prealbumin in serum, its detection by serum electrophoresis constitutes a quality criterion for serum electrophoresis (9). Prealbumin is a compact molecule that can cross the blood–brain barrier and may also be synthesized by cells of the choroid plexus. This results in a 10- to 40-fold increase in the prealbumin-to-albumin ratio in CSF. Consequently, on CSF electrophoresis, prealbumin appears as a distinct band slightly anodal to albumin. Prealbumin can be measured by a variety of immunoassays, including radial immunodiffusion and nephelometry. The normal reference range is somewhat method dependent but is approximately 195 to 358 mg/L (12).

Albumin

Albumin (MW 60,000) is the most abundant protein in plasma, comprising approximately 60% of the total protein concentration. Although most of the plasma proteins are glycoproteins, albumin contains no carbohydrate and is therefore classified as a simple protein.

Albumin consists of a single peptide chain of 580 amino acids with a molecular mass of 66.3 kDa (13). Amino acid analysis has revealed a high content of glutamic acid, aspartic acid, and lysine, which collectively impart a large number of negatively charged groups to the molecule, rendering it highly soluble at physiologic pH. Additionally, the tryptophan content is low, a feature that has been exploited in some albumin assays. Albumin also contains a high content of cysteine residues, which help maintain the tertiary structure of the molecule through 17 intramolecular disulfide bonds. These bonds support a structure containing nine loops which form a compact elliptical molecule. On storage, albumin may form intermolecular disulfide bonds, giving rise to dimers that may appear as an extra band on SPE.

Albumin is synthesized almost exclusively by the liver, appearing first in the cytoplasm of the hepatocytes as a precursor called proalbumin. Immunoperoxidase studies have demonstrated that only 10% to 35% of the hepatocytes contain albumin at any one time, suggesting that synthesis is either restricted to a limited population of cells or occurs in a cyclical fashion throughout the liver (14). Fasting decreases the rate of synthesis by as much as 60%, whereas refeeding results in a prompt increase in synthesis. Regulation of synthesis is mainly by a negative feedback effect exerted on the hepatocytes by the plasma COP (13). Some cytokines suppress synthesis. The biological half-life of albumin is approximately 19 days.

Albumin has several important and diverse functions, including contributing nearly 80% of the plasma COP. Albumin serves as an amino acid source to a variety of cells and functions as a major transport protein because the large number of negatively charged groups on the molecule are potential binding sites. Bilirubin binds to albumin, and this is the major mechanism for sequestering and transporting bilirubin to the liver from sites of hemoglobin catabolism. Albumin is intimately involved in lipid metabolism, accepting free fatty acids from lipoprotein lipase and transporting them between the liver and peripheral tissues. Several hormones bind to albumin, including T_4, T_3, cortisol, aldosterone, estradiol, and progesterone. Because these hormones preferentially bind to their respective hormone-binding globulins, albumin serves mainly as a high-capacity, low-affinity overflow reservoir (14). Many drugs bind to albumin, including warfarin (97%), phenylbutazone (99%), salicylate (40%), penicillin (65%), and chlorothiazide (89%), among many others. Significant amounts of calcium and lesser quantities of magnesium are also bound.

Only 40% of total body albumin is in the plasma at any one time. The remaining 60% is distributed among virtually every other body fluid compartment, forming an exchangeable pool with plasma (14). Hyperalbuminemia generally results from dehydration. Artifactual increases may occur after prolonged tourniquet application during venipuncture. Conversely, hypoalbuminemia is common. Decreased albumin levels may result from a variety of mechanisms, including impaired synthesis (liver disease), increased catabolism, excessive losses (nephrotic syndrome), or “third spacing” (ascites, effusions). The most severe states of hypoalbuminemia occur in protein-losing syndromes such as the nephrotic syndrome or protein-losing enteropathy. Excessive losses may also occur in hemorrhage, burns, or exudative processes, especially of the skin and gastrointestinal tract. In the acute-phase response, albumin levels are decreased by at least five mechanisms: hemodilution, increased leakage into the extravascular space, increased catabolism, suppression of synthesis directly by cytokines, and, indirectly, by increased osmotic pressure secondary to the synthesis of the positive APRs (13). Increases in immunoglobulin levels also lead to lower albumin levels. In chronic wasting diseases such as tuberculosis and malignancy, both increased catabolism and decreased synthesis are contributing factors. Malnutrition results in decreased synthesis by denying the liver the necessary amino acid substrates needed to make albumin but may also impair liver function by other mechanisms. Malabsorption syndromes produce hypoalbuminemia by mechanisms similar to those in malnutrition. Chronic liver disease of any cause impairs albumin synthesis, although in alcoholic patients this is often accompanied by poor nutritional status. Hypoalbuminemia is also a common finding in pregnancy.

Whatever the cause, hypoalbuminemia decreases the plasma COP. Severe decreases may result in edema. In addition, the binding and transport of a number of substances are correspondingly impaired. Herrman and co-workers (15) reported that patients with low albumin levels at the time of admission exhibited increased mortality, had longer hospital stays, and were readmitted sooner and more frequently than patients with normal albumin levels.

More than 80 genetic variants of albumin have been described (13). These are all inherited in autosomal codominant fashion. The most common type is called albumin A. Variant albumins may result in a wide albumin band on SPE or may give rise to two distinct bands (bisalbuminemia) (8); binding of drugs and metabolites may also change albumin's electrophoretic pattern (13). None of these variants has yet been associated with human disease, although one variant has an increased affinity for T_4 (16). In the rare syndrome analbuminemia, there is a congenital absence of albumin. These patients may have mild edema but are otherwise spared the hemodynamic consequences of severe hypoalbuminemia

owing to compensatory mechanisms, including an increase in plasma globulins, which take over some of the functions of albumin. The major biochemical problem in these patients is a disturbance in lipid metabolism including hypercholesterolemia, increased plasma phospholipids, and lipoproteins (17).

Several methods are available to measure albumin, including electrophoretic, immunochemical, and dye-binding techniques. Albumin may be estimated indirectly by subtracting the measured total globulins from the total protein value. This method exploits the fact that plasma globulins contain substantially more tryptophan residues per gram of protein than does albumin. To estimate albumin, the total protein is first measured by standard methods. Then the globulins are assayed by reaction of their tryptophan residues with glycoxylic acid, forming a purple color, which is measured spectrophotometrically. The value for the globulins is then subtracted from the total protein.

Albumin may also be estimated by quantitative densitometric scanning of the albumin band on SPE. The scan is integrated to provide an estimate of albumin as a percentage of the total electrophoretic protein. Separate measurement of the total serum protein permits the estimate to be converted to mass units. Specificity can be improved by using immunochemical techniques, including radial immunoassay, nephelometry, enzyme-linked immunoassay, and radioimmunoassay, although these methods are relatively expensive to perform (8).

Several dye-binding methods are available to measure albumin (18), and these methods have been widely used. The assays are based on the ability of albumin to bind a number of organic anionic dyes. Binding results in a shift in the absorption maximum of the dye, which can be quantitated spectrophotometrically despite the presence of unbound excess dye. Plasma globulins either do not bind the dyes or do so only slowly. The two most popular dyes are bromocresol green and bromocresol purple. Bromocresol green is not specific for albumin and binds also to globulins, but this occurs slowly and is not a problem if the sample is assayed promptly. Conversely, bromocresol purple does not bind to globulins.

Urine or CSF samples may contain only small amounts of albumin. Traditionally these samples were therefore concentrated before measurement, but some of the newer assays are sufficiently sensitive to permit elimination of the concentration step. Urinary albumin may also be estimated by a screening dipstick method (Albustix Reagent Strips; Bayer Corp., Mishawaka, IN, U.S.A.) that exploits the change in visible color of bromophenol blue from yellow to green to blue with increasing albumin concentration (18). This method is relatively nonspecific, and it does not detect the low levels of albumin seen in microalbuminuria (13).

The reference range for serum albumin varies slightly with age, gender, posture, and the method of assay. Values may be 5 g/L greater in supine versus recumbent patients and are slightly higher in adults than in children. The adult reference range is 31 to 43 g/L (12).

α_1-Antitrypsin

AAT is one of the major plasma proteins, comprising nearly 90% of the α_1 globulin region on SPE. Deficiency of AAT has been associated with pulmonary emphysema and hepatic cirrhosis. The molecule consists of a glycoprotein (MW 52,000) that is synthesized by the liver and released into the plasma. It is also found in a number of other body fluids, including tears, lymph, bile, semen, and amniotic fluid (19). AAT is one of the APRs. Plasma levels may double in states of acute and chronic inflammation, in malignancy, after trauma or surgery, and during pregnancy or estrogen therapy.

AAT belongs to a family of serum proteins collectively known as serpins (*ser*ine *p*roteinase *in*hibitors). Other proteins in this group include α_1-antichymotrypsin, antithrombin, α_2-antiplasmin, C1 inhibitor, and a variety of lesser known proteins. Although 90% of the serum antitrypsin activity is attributable to AAT, significant amounts of trypsin are not found in plasma. Consequently, AAT functions mainly to inhibit nontrypsin proteinases, especially elastase and collagenase, in tissues and body fluids. These enzymes are released by leukocytes and macrophages at sites of inflammation. Thus, AAT is an important component of the body's mechanism to control endogenous proteolysis. Left unchecked, these enzymes would ultimately destroy normal tissues.

AAT exerts its effect by forming a covalent complex with serine-type proteinases. Complex formation inhibits the enzyme, and the complex is subsequently removed by the reticuloendothelial system, primarily in the liver. The genes encoding the AAT protein comprise an autosomal allelic system containing at least 75 codominant genes (20) inherited on a single locus called Pi for *p*roteinase *i*nhibitor. The resultant AAT variant proteins are immunologically similar but can be distinguished based on their electrophoretic mobility (19). Most of these alleles give rise to functionally normal variants of AAT, while some produce abnormal molecules and are associated with AAT deficiency. The most common normal allele is PiM, which encodes for the M protein variant of AAT. In the United States, 95% of the population has the homozygous PiMM genotype. The alleles PiZ and PiS are also relatively common in white populations and produce an abnormal AAT protein (21). The rare null allele Pi produces no protein product. Non-PiM alleles are rare in African Americans (19).

The PiZ allele is the most common variant associated with AAT deficiency. The heterozygote gene frequency of PiZ in the United States is 2% to 3%, while homozygotes comprise one in 3,630 individuals. In homozygous PiZZ subjects, the plasma AAT level is reduced to 10% to 15% of normal, while heterozygotes with the PiMZ phenotype are reduced to 60% of normal. The phenotypes PiSS and PiMS have levels of 63% and 83% of normal, respectively (21).

Deficiency of AAT may be congenital or acquired. Acquired decreases are seen in protein-losing syndromes, malnutrition, severe liver damage, and some respiratory diseases, including neonatal respiratory distress syndrome. Congenital deficiency is associated with pulmonary emphysema and with hepatic cirrhosis. Nearly 70% to 80% of homozygous PiZZ individuals develop chronic obstructive pulmonary disease (COPD) (22), often by the age of 20 to 30 years. The spectrum of COPD in AAT deficiency includes patients with chronic bronchitis and bronchiectasis in addition to those with emphysema. Cigarette smoking rapidly accelerates the progression of pulmonary disease

in homozygous PiZZ individuals and appears also to increase the risk of heterozygotes. Heterozygote PiZM individuals who do not smoke appear to have little if any increased risk of COPD (23). Not all subjects with severe AAT deficiency develop lung disease, and most patients with COPD are not AAT deficient. Possibly as few as 1% of patients with emphysema has AAT deficiency (22).

The phenotypes PiZZ, MZ, and SZ have also been associated with liver disease (24). Fifty percent of homozygous PiZZ patients exhibits intermittent abnormalities in liver function tests, and 10% to 20% develop permanent abnormalities. Nonetheless, most patients with AAT deficiency do not develop chronic liver disease. Patients may present during the neonatal period with jaundice or with neonatal hepatitis. The latter entity may appear similar clinically and morphologically to viral hepatitis or chronic active hepatitis. AAT deficiency should be considered in all infants with jaundice and all children with liver disease (24). Some patients with AAT deficiency develop cirrhosis by 2 to 3 years of age, while in others the onset is delayed until adult-hood.

Homozygous PiZZ AAT deficiency can usually be detected on SPE by noting a marked decrease in the α_1 region. This region is never completely empty because 10% of the α_1 zone is composed of other proteins. Heterozygotes often have a normal-appearing α_1 zone. A semiquantitative estimate of AAT may be obtained by densitometric scanning of the serum electrophoretogram. Traditionally, AAT activity was assessed in a functional assay that measured its ability to inhibit the catalytic activity of trypsin on certain synthetic substrates (trypsin inhibitory activity). These assays were not specific for AAT and have been replaced by immunochemical methods including nephelometric assays (8).

Quantitative assays are useful to detect patients with AAT deficiency but do not provide information about the Pi phenotype. The normal reference range of AAT is 76 to 189 mg/dL (12). Because proteases released from leukocytes or bacteria can form complexes with AAT and interfere with quantification, serum should be separated from the clot as soon as possible and stored aseptically (8). Caution should be exercised in interpreting AAT levels during pregnancy, in patients on estrogen or steroid therapy, and in patients exhibiting an acute-phase response because heterozygotes and some homozygotes can increase their AAT level in these situations. In patients with M protein levels below 50 mg/dL, AAT phenotyping should be performed (8). Methods for phenotyping include isoelectric focusing, starch gel electrophoresis, and two-dimensional immunoelectrophoresis. Polymerase chain reaction–based methods utilizing amniocytes or chorionic villi are available for prenatal diagnosis. For an excellent overview of AAT deficiency, the reader is referred to the recent review by Cox (23).

α_2-Macroglobulin

α_2-Macroglobulin (AMG), haptoglobin, and CER together make up the α_2 region on SPE. AMG consists of four identical subunits and is one of the largest serum proteins (MW 720,000). Because of its large size, it does not diffuse out of the blood vessels and can only be found in the intravascular space. Several functions have been ascribed to AMG. In addition to its long-known function as an inhibitor of endogenous and exogenous proteinases, AMG may also function as a binding, carrier, and targeting protein for cytokines and growth factors (25).

AMG is synthesized by the liver with a half-life in serum of 5 days. The half-life of AMG–proteinase complexes is approximately 8 minutes (26). Owing to the wide spectrum of activity of AMG against a variety of proteinases, it has been difficult to ascribe a specific function to this protein. Although it appears important in defense against proteolytic enzymes, AMG is present in plasma at only one tenth the concentration (on a molar basis) of AAT.

Several endopeptidases, including serine proteinases (trypsin, plasmin, elastase, collagenase), thiol proteases, metal proteases, and the carboxy protease cathepsin D, are inhibited by AMG. Unlike AAT, in which inhibition occurs by formation of a covalent bond with the active site of the proteinase, AMG irreversibly binds proteinases without affecting the active center of the enzyme (1). Once bound, the AMG-proteinase complex is rapidly cleared from the circulation by binding to a receptor expressed on macrophages, hepatocytes, and fibroblasts (27).

The biological significance of AMG as a transport protein is still the focus of ongoing research. AMG can also bind numerous growth factors, hormones, and cytokines. These include TGF-ß1 and TGF-ß2, IL-1ß, IL-6, basic fibroblast growth factor (bFGF), platelet-derived growth factor (PDGF), epidermal growth factor (EGF), nerve growth factor-ß (NGF-ß), TNF-α, insulin, vascular endothelin growth factor (VEGF), placental lactogen, and the inhibin/activin family (27,28). At least three different mechanisms have been described for the binding of these molecules to AMG. The binding of these factors to AMG might play a role in the transport and targeting of these messenger molecules; however, the biological significance of this phenomenon is still the subject of investigation (27).

AMG levels are greatly increased in the nephrotic syndrome and may attain levels greater than 1,000 mg/dL. This occurs because the large size of AMG prevents its loss in the urine through the damaged glomeruli, while the smaller plasma proteins are selectively lost. AMG may also be increased by as much as 70% in cirrhosis despite significant loss of the liver cell mass. This finding is opposite that of most other plasma proteins, although the immunoglobulins may also be elevated in liver disease. Minor increases are seen in diabetes mellitus. Unlike AAT, which behaves as an APR, AMG levels are not altered in states of inflammation. No human deficiency of AMG has been reported, and no disease has yet been identified in which measurement of AMG is of clinical value, although nephelometric and immunochemical assays exist. The reference range is 125 to 215 mg/dL. Values are slightly higher in women and are age dependent.

Ceruloplasmin

CER (MW 132,000) is a glycoprotein synthesized by the liver as a single polypeptide chain to which six to eight copper atoms are attached (29). CER migrates as an α_2 globulin on protein electrophoresis but is not normally visible on routine gels except in situations in which it is significantly increased. The pure protein is blue because of its high copper content. Increased levels may impart a green tinge to plasma samples.

While copper is an essential nutrient, it is also highly toxic to cells. Consequently, efficient means of maintaining copper homeostasis are required. After absorption in the intestine, dietary copper is bound to albumin and transcuprein and is taken up by the liver and to a lesser extent by all cells. The liver functions as a storage site. Copper is incorporated into the CER apoprotein in the liver and is released into the plasma. The majority of plasma copper is bound to CER; the remainder is bound to albumin and transcuprein. A small fraction is bound to small peptides and amino acids (8). The principal route of copper excretion is through the biliary tract, with a small amount being excreted in the urine. Copper homeostasis reflects the balance between intestinal absorption and biliary excretion. The liver is the main recipient of dietary copper and the primary route of excretion.

The principal importance of CER in laboratory medicine is in the diagnosis of Wilson's disease (hepatolenticular degeneration), which typically is associated with low plasma CER levels. Despite intensive investigation, the function of this protein remains controversial. Historically it was thought that CER played a central role in copper transport because CER-bound copper normally constitutes the majority of the total plasma copper. However, radioisotope studies have shown a negligible turnover of CER-bound copper, indicating that copper is neither lost nor gained from the molecule in the circulation (8). Furthermore, the protein deficiency state of hypoceruloplasminemia produces no symptoms of copper overload (30). It is now thought that CER may play a role in copper metabolism by acting as a donor of copper to some key copper-containing enzymes. This is accomplished through uptake and degradation of CER by a number of cells but does not involve transfer of copper from intact CER molecules. However, the primary role of CER appears to be in plasma redox reactions. CER oxidizes Fe^{2+} to Fe^{3+} and thereby allows the incorporation of iron into transferrin without the formation of toxic iron products (8). CER also plays a role in preventing lipid oxidation and free radical formation, both of which are damaging to cells.

Wilson's disease is an autosomal recessive disorder associated with toxic accumulation of copper, particularly in the liver and brain. Although the incidence is low (one in 50,000 to 100,000), it may be the most common cause of chronic liver disease in childhood. It was once believed that Wilson's disease resulted from an inherited defect in CER, but this concept has subsequently been refuted (29). It has been shown that the gene responsible for Wilson's disease resides on chromosome 13, while the gene for CER is on chromosome 3. The Wilson's disease gene is a P-type ATPase; a defect in this gene prevents incorporation of copper into CER and leads to impairment of biliary copper excretion. These defects cause a progressive accumulation of copper in the liver, resulting in chronic liver disease. Ultimately, copper may overflow from the liver and deposit in other tissues, most notably in the basal ganglia of the brain (lenticular degeneration) and in the cornea of the eye. Deposition in the brain causes neurological symptoms, including dysarthria and loss of coordination of voluntary movements. Corneal deposits produce the classic appearance known as Kayser-Fleischer rings. Copper deposition in other tissues may result in renal tubular damage, kidney stones, osteoporosis, arthropathy, cardiomyopathy, and hypoparathyroidism (29).

Hepatic involvement in Wilson's disease produces variable degrees of liver dysfunction. Some patients exhibit only minor impairment in liver function, while others develop severe liver damage as early as 8 years of age. The onset of hepatic failure may be abrupt or may result from progressive cirrhosis over many years. Fortunately, patients with Wilson's disease can be effectively treated with penicillamine, which chelates body copper and promotes its excretion in the urine. If patients are diagnosed before the onset of cirrhosis and neurological symptoms, a normal life span can be expected. For this reason, early diagnosis and treatment are essential, as is the evaluation of the patient's siblings who have a 1-in-4 chance of developing the disease. However, the majority of cases go undetected and ultimately present with advanced disease (8). Consequently, Wilson's disease should be considered in all patients with chronic liver disease and in all patients older than 12 years of age with relevant neurological findings (31). Heterozygotes for the Wilson's disease gene do not develop manifestations of the disease. For siblings of patients with Wilson's disease, DNA linkage analysis can be used to make a diagnosis before the onset of clinical symptoms (32).

The majority (90% to 95%) of patients with Wilson's disease exhibit decreased levels of plasma CER (less than 10 mg/dL) with a concomitant decrease in the total serum copper concentration. However, the non-CER copper is increased, with a resultant increase in urinary copper excretion. Administration of penicillamine augments urinary copper excretion and may be helpful in establishing the diagnosis. While all patients with neurological disease exhibit the classic laboratory findings in addition to the Kayser-Fleischer rings, perhaps as many as 20% of patients with disease limited to hepatic involvement have normal CER levels. Liver biopsy with measurement of the hepatic copper content by atomic absorption is a fairly reliable test for Wilson's disease when the level is greater than 300 mg/g dry weight. However, the hepatic copper content may also be increased in some other types of liver disease. The best test for the diagnosis of Wilson's disease is the demonstration of a negligible incorporation of radioactive copper into CER. In this test, either ^{64}Cu or ^{67}Cu is administered intravenously and the plasma radioactivity counted periodically for 48 hours. Because free radiolabeled copper is rapidly assimilated by the liver, the residual plasma radioactivity reflects copper incorporation into CER (31).

Low levels of plasma CER are not specific for Wilson's disease. Any condition associated with severe liver dysfunction will impair CER synthesis. Decreased levels are also seen in malnutrition and protein-losing states (8). Conversely, plasma CER levels are increased in pregnancy, with oral contraceptive use, and in the acute-phase response. This increase in inflammation is thought to be related to the need for scavenging oxygen radicals released at the site of inflammation by immune cells (29). Biliary tract obstruction, especially primary biliary cirrhosis, may also result in elevated plasma CER.

The reference range for CER varies with age, being highest in young children and somewhat lower in both infants and adults. Values less than 10 mg/dL are suggestive of Wilson's disease, but normal levels do not exclude the diagnosis. The adult range varies with the method of assay but is approximately 27 to 50 mg/dL (12). Most CER assays are based on immunochemical

methods, including nephelometry and radial immunodiffusion (RID). Alternatively, the oxidase activity of CER can be measured using the substrate p-phenylenediamine in a colorimetric assay.

Haptoglobin

Haptoglobin is a glycoprotein synthesized by the liver that migrates as an α_2 globulin. In addition to its well-known function in the binding of free hemoglobin, this APR might also play an important role in the control of local inflammation (8,33,34).

Haptoglobin binds free hemoglobin released by intravascular red blood cell destruction, thus conserving body iron and preventing renal filtration of potentially nephrotoxic hemoglobin. Free hemoglobin dissociates into α-ß dimers that are bound by haptoglobin. The complexes are subsequently removed by the reticuloendothelial system, which degrades the heme groups to iron and bilirubin. Complex formation occurs through the α chain of the hemoglobin dimers. Consequently, haptoglobin can bind hemoglobin types A, F, S, and D but not hemoglobin types that lack an α chain. Methemoglobin and free heme moieties do not bind haptoglobin. The total plasma binding capacity of haptoglobin is approximately 3 g of hemoglobin.

Normally, intravascular destruction accounts for less than 10% of the total red cell turnover. Continued synthesis of haptoglobin by the liver prevents it from being depleted under normal circumstances. In states of accelerated intravascular hemolysis, haptoglobin may become depleted, resulting in free hemoglobin dimers, which are filtered by the kidney. Once filtered, free hemoglobin is degraded by the renal tubular epithelial cells, with heme-iron ultimately being converted to hemosiderin. The kidney can process as much as 5 g of hemoglobin per day, but above this level, free hemoglobin appears in the urine. Free hemoglobin that has not been bound by haptoglobin or processed by the kidneys is oxidized in the plasma to methemoglobin and the heme groups are subsequently released. The free heme groups are taken up by another binding protein, hemopexin, which carries them to the liver, where they are degraded. Like haptoglobin, hemopexin may also become depleted in states of accelerated hemolysis.

The role of haptoglobin in the control of local inflammation is the focus of intensive investigation. Haptoglobin might protect against heme-stimulated oxygen radical damage (33). It also inhibits cathepsin B released by phagocytes (34). Recently, Ghamati and co-workers (35) showed that haptoglobin binds to the CD11b/CD18 receptor, a member of the integrin family. This receptor mediates the adherence and extravasation of phagocytes at sites of inflammation. This suggests a role for haptoglobin in the regulation of chemotaxis.

Three distinct phenotypes of haptoglobin have been identified by electrophoresis and designated Hp1-1 (MW 85,000), Hp1-2 (MW 120,000), and Hp2-2 (MW 160,000). These phenotypes have not been associated with human disease, but they may lead to slightly differing haptoglobin levels (36). Because of the genetic heterogeneity of haptoglobin, it has been widely used in paternity testing and forensic medicine (37). As many as 4% of blacks have congenital anhaptoglobinemia with no detectable haptoglobin in the plasma (37).

Haptoglobin is an APR and may therefore be elevated in the acute-phase response. Therefore, normal haptoglobin levels in patients with an inflammatory process do not rule out hemolysis. Elevated haptoglobin levels are also seen in the nephrotic syndrome in which other proteins are preferentially lost and in patients with increased estrogen levels.

Haptoglobin is decreased in diseases associated with intra-vascular hemolysis, including hemolytic anemias, hemoglobin-opathies, hemolytic transfusion reactions, extensive burns, malaria, disseminated intravascular coagulation, and exercise-induced hemolysis. In severe acute hemolysis, haptoglobin may be totally depleted within 6 to 12 hours, requiring as long as 1 week to return to normal. In chronic hemolytic states such as are seen with mechanical heart valves or hemoglobinopathies, there may be a steady decline in haptoglobin. In these conditions, serial measurements are a better index of ongoing hemolysis than single haptoglobin values. Haptoglobin may also be depressed in ineffective erythropoiesis associated with intramarrow red cell destruction (e.g., vitamin B_{12} and folic acid deficiency) and in liver disease owing to decreased synthesis. Haptoglobin levels in neonates are virtually zero; adult levels are only reached by the age of 4 months (38).

The complex formed by the binding of hemoglobin to haptoglobin shows a slower migration on SPE than haptoglobin alone. The presence of these complexes can indicate acute hemolysis or poor specimen handling (38).

Determination of haptoglobin levels are useful to distinguish hemoglobinuria from myoglobinuria: As long as detectable haptoglobin levels are present in the plasma, a positive dipstick test is consistent with myoglobinuria.

Historically, haptoglobin was measured by determining the hemoglobin-binding capacity of serum. In this assay, a patient sample is first saturated with hemoglobin followed by electrophoretic separation of the free from the bound hemoglobin. The haptoglobin–hemoglobin complex is then quantified by its reaction with benzidine and expressed as milligrams of hemoglobin-binding capacity per deciliter of serum. This cumbersome procedure has now been replaced by immunochemical methods, including nephelometry and RID. The reference range is 16 to 199 mg/dL (12).

Transferrin

Transferrin (siderophilin) appears as a distinct band on SPE and constitutes the major component of the ß-globulin fraction. It functions as the principal plasma protein responsible for the transport of iron. Transferrin is a glycoprotein (MW 79,550) containing two iron binding sites on a single polypeptide chain. Most of the transferrin in plasma is synthesized by the liver, although small amounts are made in a variety of other sites, including the reticuloendothelial system, the gonads, and the submaxillary gland (39). Transferrin binds ferric iron in a reversible ionic bond accompanied by the uptake of one bicarbonate anion per atom of iron. The complex is stable at physiological pH, but dissociation occurs in acidic solution. Transferrin is also capable of loosely binding several other metal ions, including copper, zinc, cobalt, and calcium, but, with the exception of copper, this is not of physiological significance (39).

The body contains approximately 3 to 5 g of iron, but only 3 to 5 mg of iron is found in plasma. Almost all this iron is bound to transferrin, with a small amount bound to other proteins such as albumin. The level of free iron is very low (less than 1 µg/dL). Consequently, the serum iron concentration refers mainly to iron bound to transferrin, whereas the total iron-binding capacity (TIBC), which is also commonly measured, is simply an approximation of the transferrin concentration. Normally only approximately 33% of the iron binding sites on transferrin are occupied. This is sometimes reported as the percentage of saturation of transferrin.

Transferrin is responsible for the transport of iron from its site of absorption in the intestine (or from sites of hemoglobin catabolism) to red cell precursors in the bone marrow or to sites of iron storage in the reticuloendothelial system of the bone marrow, liver, and spleen. The transferrin–iron complex binds to receptors expressed by these cells and is internalized into endosomes. The lower pH in these organelles leads to a dissociation of iron from transferrin. After unloading its iron, transferrin returns to the plasma and is recycled. Radiolabel tracer studies have demonstrated a biological half-life of approximately 8 days. In addition to its transport function, transferrin minimizes the level of free iron in plasma, thus reducing urinary iron losses and preventing the potentially toxic effects of high free iron levels.

The transferrin protein exhibits extensive genetic polymorphism arising from amino acid substitutions in the peptide chain and variations in the carbohydrate chains and in iron content (8). Several variants have been identified on protein electrophoresis. They are named according to their relative electrophoretic mobility when compared with the most common transferrin type, known as transferrin C. For example, transferrin D was the first variant protein discovered. Transferrin D is found in some American blacks and Australian aborigines. Although none of the electrophoretic variants has been associated with human disease, they will occasionally masquerade as M components on serum electrophoresis. Heterozygotes exhibit two transferrin bands, one with a normal mobility and reduced concentration (C protein) and the second migrating either above or below the normal band. Immunofixation techniques will distinguish these variants from true M components. Rarely, congenital atransferrinemia results in very low levels of plasma transferrin in association with iron overload and severe anemia owing to the inability to mobilize the body's iron stores (8). In patients with a history of alcohol abuse, transferrin contains less sialic acid than in control populations. A high-performance liquid chromatography–based method for the quantitation of such carbohydrate-deficient transferrin was described recently by Jeppsson et al. (40). This method appears to be a highly specific and sensitive assay for alcohol abuse.

Decreased levels of serum transferrin may be seen in liver disease and in protein-losing disorders such as the nephrotic syndrome and protein-losing enteropathy. Transferrin levels are depressed in starvation and are a useful marker of protein malnutrition in hospitalized patients (5). Like albumin and ß-lipoprotein, transferrin behaves as a negative APR (8). Consequently, low levels are seen in a variety of acute and chronic inflammatory states and in malignancy. This might serve to deprive bacteria and tumor cells of iron (41). Transferrin levels may rise during the late stages of the acute-phase response (see APRs).

Assay of serum transferrin is most commonly performed in the evaluation of anemia. In iron deficiency anemia, the transferrin level is elevated but the percentage of saturation is low. In the anemia of chronic disease, transferrin is decreased and the percentage of saturation is decreased. Transferrin levels may be increased in the early stages of acute hepatitis, in pregnancy, and with estrogen administration. In idiopathic hemochromatosis and in hemosiderosis owing to excessive blood transfusions, the serum transferrin is normal but the percentage of saturation is markedly elevated (as high as 90%).

The normal serum transferrin level is 191 to 365 mg/dL (1.91 to 3.65 g/L) (12). Increases or decreases in transferrin may be detected on SPE, and marked increases may simulate a paraprotein band.

Several immunochemical assays are available, including RID and rate nephelometry, and transferrin may also be estimated indirectly by determination of the TIBC. The TIBC (in milligrams per deciliter) is then multiplied by 0.70 to give an estimated transferrin value in milligrams per deciliter (8).

ß-Lipoprotein

Lipoproteins constitute a family of molecules composed of lipids and proteins whose function is to transport cholesterol, triglycerides, and phospholipids in the blood. Lipoproteins have been subclassified into four categories designated chylomicrons, very low density lipoproteins (VLDL), low-density lipoproteins (LDL) and high-density lipoproteins (HDL).

On SPE, HDL migrate between the albumin zone and the α region (α-lipoproteins), forming a diffuse background pattern that is not readily discernible. Likewise, the VLDLs migrate in the pre-ß region but are not resolved from other proteins. However, the ß-lipoproteins (LDL) appear in the slow α_2 or ß region as a wavy, broad band with a sharp leading edge and a trailing anodal tail.

Increases or decreases in the ß-lipoproteins can often be recognized on SPE. Elevations may occur in the nephrotic syndrome, hepatobiliary disease, hypothyroidism, diabetes mellitus, and some of the familial hyperlipidemia syndromes.

An increase produces a high-intensity band that may be cathodally displaced; this finding should prompt further investigation. Decreased ß-lipoproteins are associated with a faint band of faster mobility (42).

$ß_2$-Microglobulin

$ß_2$-Microglobulin ($ß_2M$) is a small protein (MW 11,818) comprising the common light chain of the class I major histocompatibility complex antigen on the surfaces of all nucleated cells (43). It consists of a single polypeptide chain that is noncovalently linked to the human lymphocyte antigen (HLA) heavy chain. $ß_2M$ may be necessary for insertion of the HLA molecule into the cell membrane, and its presence appears to stabilize the HLA heavy chain. It also plays a role (currently poorly defined) in the regulation of some lymphocyte functions.

A small amount of $ß_2M$ is normally present in the plasma,

urine, and CSF, presumably because the protein is shed from cell surfaces into the blood (8). Its function in these fluids, if any, is unknown.

Because of its small size, β_2M is filtered by the renal glomerulus, but most is resorbed and degraded by the renal tubular epithelial cells. The plasma level of β_2M is a good index of the glomerular filtration rate. Only trace amounts of the protein are normally present in urine. Elevated levels of urinary β_2M may be seen in patients with renal tubulointerstitial disorders, including heavy metal toxicity, anticancer drug toxicity, aminoglycoside toxicity, upper urinary tract infections, and renal allograft rejection. In contrast, patients with renal failure based on glomerular dysfunction typically show elevated plasma β_2M levels. Plasma β_2M may also be increased in a number of inflammatory diseases [e.g., hepatitis, rheumatoid arthritis, systemic lupus erythematosus, acquired immune deficiency syndrome (AIDS), and sarcoidosis] and in patients with leukemia, lymphoma, and some solid tumors. In some of these disorders, the elevation in the plasma β_2M may reflect increased lymphocyte turnover.

Several reports have indicated that β_2M levels might be an extremely useful prognostic marker for some types of lymphoproliferative disorders. Durie and colleagues (44) reported that serum β_2M is the best available predictor of survival duration in patients with multiple myeloma. Similarly, Kantarjian et al. (45) found that monitoring serum β_2M levels may provide significant prognostic information in patients with adult acute lymphocytic leukemia. Measurement of β_2M also adds relevant prognostic information to CD4 counts in predicting the progression to AIDS in human immunodeficiency virus (HIV)–infected individuals (46,47). β_2M may sometimes be seen on high-resolution electrophoresis. However, because of its low concentration, it is usually measured by immunoassay. The normal upper reference limit is 2.8 mg/L in serum and 0.2 mg/L in urine (12).

C-Reactive Protein

CRP (MW 118,000 Da) was first identified in 1930 as a substance present in the sera of patients with pneumococcal pneumonia that could bind to C-polysaccharide isolated from *Streptococcus pneumoniae*, producing a flocculation reaction. Subsequently, it was found that CRP is elevated in a variety of other acute inflammatory diseases. Thus, CRP was the first recognized APR. It rises rapidly (24 to 48 hours) after the inflammatory stimulus and shows a large increase (as high as 1,000-fold) and a short half-life of 19 hours (48). The rapid, high rise and early return to normal levels after successful therapy make CRP the APR best suited for diagnostic purposes.

CRP is synthesized by the liver and released into the plasma. Small amounts are also made by a subset of peripheral lymphocytes, but this remains bound to the surface of the cells (49). The intact molecule is a nonglycosylated, pentameric protein with five identical subunits arranged in a doughnut-shaped polymer. Each subunit consists of 206 amino acids.

Several functions have been ascribed to CRP, including initiation of opsonization and phagocytosis (8) and activation of complement, neutrophils, and monocyte-macrophages (7). Collectively, these properties imply an important role for CRP in the recognition of microbial organisms and as an immunomodulator in host defense. CRP may also be important to the recognition of necrotic tissues. It can bind to a number of molecules, including phosphate esters, lipids, polyanions (DNA, polylysine), polycations (histones, protamine), and various polysaccharides (8,49).

Small amounts of CRP are normally found in plasma at levels less than 0.8 mg/dL. The majority of CRP values for normal healthy individuals fall below the detection limit of most routinely used methods (48). Levels may increase markedly during the acute-phase response. Although CRP is not ordinarily seen on SPE, it may appear as a distinct band in the mid-β to slow γ region when levels are elevated. This can be confused with a monoclonal M component.

Traditional methods for measuring CRP included precipitation and agglutination assays. These had low sensitivity and provided only qualitative results. The lack of a satisfactory assay caused a loss in popularity of CRP testing in favor of the ESR for detecting patients with inflammatory diseases. Recently, highly specific antibodies to CRP have become available. These antibodies permit the development of rapid, specific, and very sensitive assays for this protein. These newer immunoassays include laser nephelometry (the most popular method), radioimmunoassay, and enzyme immunoassays (7) and have created a renewed interest in CRP testing in various clinical settings. CRP is a more sensitive, rapidly responding indicator of inflammation than the ESR and may someday replace it. Because infections with bacteria lead to higher CRP levels than viral diseases, CRP levels may be helpful to distinguish between these diagnostic possibilities. CRP levels are also used to follow patients with autoimmune diseases.

METHODS OF PROTEIN ANALYSIS

Part of "13 - The Plasma Proteins"

Measurement of Total Protein

Several methods are available to measure the total protein in serum or plasma. Most methods are based on the assumption that protein molecules are pure polypeptides and that different proteins react with chemical reagents in the same way (8). Neither of these assumptions is strictly true. For example, many proteins contain carbohydrate substituents that may constitute a significant part of the molecule but might not be included in the total protein value as measured by standard methods. In addition, the reactivity of different proteins depends on the structure and amino acid composition of the molecule, which are not the same for all proteins. The fact that biological fluids are composed of mixtures of proteins in varying proportions complicates the quantification of proteins in serum, urine, and CSF.

Historically, the Kjeldahl procedure for measuring total nitrogen has been considered a reference method against which other methods for protein could be compared. In this procedure, the sample is first subjected to acid hydrolysis, which converts protein nitrogen to ammonium ions. The ammonia nitrogen is then determined by titration or by nesslerization. Because nitrogen contained in nonprotein compounds is also detected by this procedure, a deproteinized sample must be tested in parallel and the value subtracted from the total nitrogen content of the original specimen. The nitrogen content of proteins varies between

15.1% and 16.8% (3), which imparts a slight error when different proteins or protein mixtures of varying proportions are analyzed. The Kjeldahl procedure is too time-consuming and complex for routine use.

The biuret method is the most widely used procedure for the measurement of total protein in the clinical laboratory. It is both simple and easily automated but lacks sufficient sensitivity for use on low-concentration samples such as CSF or urine. The method is based on the reaction of peptide bonds with copper ions (Cu^{2+}) in alkaline solution, forming a pink to violet complex that is measured spectrophotometrically at 540 nm. The intensity of the color is proportional to the protein concentration. A similar reaction occurs between copper ions and the organic compound biuret from which this method derives its name (8). The biuret reaction will occur with any peptide with three or more amino acids. Common causes of interference are lipemia, hemolysis, bilirubinemia, and turbidity, all of which may contribute to the total absorbance at 540 nm. The appropriate use of serum blanks will correct for these sources of interference.

The Lowry method employs a two-stage reaction sequence and is one of the most sensitive (but not specific) procedures for the measurement of total protein. Proteins are first allowed to undergo the biuret reaction. This is followed by reaction of the tryptophan and tyrosine residues with the Folin-Ciocalteu reagent (phosphotungstic-phosphomolybdic acid), yielding a blue color. Histidine and cysteine residues also react but produce weaker chromogens. The combined absorbance of the copper-peptide bond complexes and the reduced Folin-Ciocalteu chromogens is measured at 650 to 750 nm. The copper-peptide bond complexes account for most of the absorbance.

The Lowry method has not been widely used in clinical laboratories for several reasons. First, a variety of nonprotein compounds and drugs react, causing a positive interference. Second, the tyrosine and tryptophan content of proteins varies widely, which complicates the analysis of mixed protein samples. For example, albumin contains only 0.2% tryptophan by weight, whereas the tryptophan content of the globulins is generally between 2% and 3% (8). The method is best for measuring pure proteins of known composition. The major advantage of the Lowry method is that it is nearly 100 times more sensitive than the biuret reaction. It has been used extensively in research to measure samples of low protein concentration.

The absorption of ultraviolet light by proteins at either 280 nm or at 200 to 225 nm provides a simple method to estimate the concentration of proteins in solution. The absorbance at 280 nm is caused mainly by the aromatic rings of tyrosine and tryptophan. The content of these amino acids varies in different proteins, and free tryptophan and tyrosine cause a positive interference. The peptide bonds of proteins are largely responsible for the absorbance at 200 to 225 nm, but measurements this low in the ultraviolet region are also subject to a number of interferences. Again, the problem of poor specificity limits the usefulness of methods based on ultraviolet light absorption.

A very rapid estimate of the total protein in serum may be obtained by refractometry. This method is based on the principle that dissolved solutes increase the refractive index of water. Although serum contains a number of nonprotein compounds, these normally contribute very little to the refractive index of serum. The total protein value is simply read from a precalibrated scale on the refractometer. A positive interference is seen with azotemia, hyperglycemia, hyperbilirubinemia, and lipemia.

A number of dye-binding procedures are available to measure the total protein in clinical specimens. The principle of these assays is similar to the dye-binding methods used for albumin. The most widely used dyes are Coomassie brilliant blue and amido black. Dye-binding methods are limited by the fact that different proteins exhibit different dye-binding characteristics.

Finally, a number of turbidimetric and nephelometric methods are available for measuring proteins in serum or other fluids. These involve precipitation of the proteins with sulfosalicylic acid or trichloracetic acid. The resulting precipitate will scatter a beam of incident light, which forms the basis of the assay.

The normal reference range for serum proteins is 60 to 80 g/L (6.0 to 8.0 g/dL) (12). The level in infants is somewhat lower than in adults, with adult levels being achieved by 3 years of age. Values in women may be slightly lower than in men.

The measurement of protein in urine or CSF is more difficult than in serum because protein is present in much lower concentrations in these fluids. The most commonly used procedures employ modifications of turbidimetric and dye-binding assays. Among the latter group, Coomassie brilliant blue and Ponceau S are currently popular. Several immunoassays are also available. All methods have the types of problems described previously, and no assay is totally satisfactory.

Serum Protein Electrophoresis

SPE is a common laboratory screening test that may be used to evaluate a variety of disease processes. The most common indication for SPE is the detection of monoclonal paraproteins, although the technique has value in a number of other clinical situations. Electrophoretic methods have also been applied to the analysis of isoenzymes, lipoproteins, and hemoglobinopathies and to the detection of genetic variants of various proteins.

Principles of Electrophoresis

Electrophoresis is a method of separating charged molecules based on their relative mobilities in an electric field. A number of different support media may be used as the stationary phase (e.g., paper, agarose gel, cellulose acetate), while the mobile phase is composed of buffer ions moving under the influence of an electric field. When an electric field is applied, the negatively charged species migrate toward the anode (positive electrode), and positively charged species toward the cathode (negative electrode). Proteins contain both carboxylic acid groups (COOH), which can deprotonate to form negative charges, and amino groups ($-NH_2$), which can accept protons to form positive charges. The net charge on the protein molecule thus depends on its amino acid composition, the presence of charged substituent groups such as carbohydrates, and the pH of the electrophoresis buffer.

The direction of migration of a protein in an electric field depends on the pH of the buffer and the isoelectric point of the protein. The isoelectric point (PI) is defined as the pH at which the sum of all positive and negative charges on the molecule add

up to zero. Thus, a protein at its isoelectric point has no net charge and will not migrate in an electric field. At a pH below the isoelectric point, the protein has a net positive charge and the protein will therefore migrate toward the cathode. At a pH above the isoelectric point, the protein will migrate toward the anode.

Other factors that influence the migration of proteins in an electric field are the size and shape of the molecule, the strength of the electric field, the temperature, the effects of convection and diffusion, and the ionic and pore properties of the electrophoresis medium (50).

The first electrophoresis system was described by Tiselius in 1937. This method used a U-shaped tube filled with an electrolyte solution. The sample was then placed at the bottom of the tube. When a current was run through the system, the proteins were separated into four fractions designated albumin, α, ß, and γ. The fractions could be identified by observing a change in the refractive index of the electrolyte at the boundaries between the fractions. The Tiselius procedure has been called the moving-boundary method.

Several technical innovations have occurred since the original description of Tiselius. Modern electrophoresis methods use a solid supporting medium such as cellulose acetate, agarose gel, or polyacrylamide that is saturated with buffer. The supporting medium is connected to two electrodes by wicks that are also saturated with buffer. A power supply capable of delivering a constant voltage or current is connected to the electrodes. When a protein mixture is placed on the supporting medium, it migrates through the electric field and separates into individual components. Methods that use a solid supporting medium are referred to as zone electrophoresis.

The choice of electrophoresis buffer is one of the most important factors affecting the resolution of the system. The buffer defines the pH, which determines the net charge on the protein. This influences both the speed and direction of movement. In addition, the ionic strength of the buffer affects the rate of protein migration. Increasing the ionic strength increases the ionic association between buffer ions and exposed charge groups on the protein molecule. This has the effect of slowing the rate of movement of the protein through the support medium, producing sharper bands on the electrophoretogram. However, as the ionic strength is increased, so too are the current and heat generated during electrophoresis. This effect may denature heat-labile proteins and cause evaporation of buffer from the support medium. The effects of heating can be minimized by incorporating a cooling unit into the electrophoresis apparatus and by performing the procedure in a closed chamber to reduce evaporation.

Several support media are available for protein electrophoresis. These can be divided into two types. One type separates proteins solely based on net molecular charge. Examples of this type include paper, cellulose acetate, and agarose gel. The second type separates proteins based on both charge and molecular size. The use of two parameters to separate proteins greatly increases the resolution of the system. Examples of the latter category are starch gel and polyacrylamide (51). Both of these media contain minute pores through which the proteins must travel when migrating under the influence of the electric field. Small molecules traverse the pores more easily than larger ones, producing a molecular sieve effect. The rate of migration of proteins through these media is therefore inversely proportional to their molecular size. Proteins of different sizes but with similar size to charge ratios are not separated on cellulose acetate or agarose, whereas separation may be achieved on starch gel and polyacrylamide.

Some electrophoretic support media adsorb hydroxyl ions from the buffer, creating fixed negative charges. This property is responsible for an effect called electroendosmosis. The fixed charges are surrounded by a cloud of mobile positive ions in the buffer. When a current is applied to the system, the ionic cloud is free to move toward the cathode, whereas the fixed negative charges are immobilized on the support medium. The migrating positive ions are hydrated and therefore carry a shell of solvent molecules with them. The movement of positive ions and solvent toward the cathode creates the endosmotic effect. Proteins migrating toward the anode must move against this force. At the usual pH of electrophoretic buffers, most proteins contain enough negative charges to overcome electroendosmosis. However, proteins that bear relatively few negative charges, such as the γ-globulins, may remain still during electrophoresis or be swept backward toward the cathode by the force of endosmosis. This is the reason that most of the γ-globulins are found cathodal to the origin on standard electrophoretograms. Endosmosis is most pronounced with support media containing a large number of fixed negative charges (paper, cellulose acetate, and conventional agarose) but is less important with systems using polyacrylamide or purified agarose gels.

SPE on paper was once popular but has been largely replaced by methods using cellulose acetate or agarose gels. Polyacrylamide and starch gel electrophoresis provide the greatest resolution of protein bands but are generally used only in specialized situations or in research where high resolution is desirable.

Some laboratories are using an improved method for SPE known as high-resolution electrophoresis (HRE). HRE separates serum into 13 zones instead of the usual five. In addition, several protein bands are detected on HRE that are not well resolved by conventional SPE. HRE may be particularly useful to detect faint monoclonal paraproteins but is of limited benefit for analyzing other serum protein patterns.

Agarose Gel Electrophoresis

Both cellulose acetate and agarose have been widely used in the clinical laboratory. The use of agarose gels for SPE is becoming increasingly popular. A comparison of the two methods was reported by Aguzzi et al. (52). The remainder of this section is devoted to a discussion of agarose gel electrophoresis.

A distinction should be made between agar and agarose. Agar contains both agarose and agaropectin. Agaropectin is a highly charged substance with both carboxylic and acid sulfate groups, which create a significant endosmotic effect. In addition, agar tends to retain protein stains, which decreases the clarity of the electrophoresis pattern. Agarose is a purified form of agar in which most of the agaropectin has been removed. Partially purified agarose preparations showing moderate endosmosis are commercially available and provide a virtually transparent gel. This facilitates visualization and quantitation of the protein bands.

A standard method for serum protein agarose gel electrophoresis was described as a proposed selected method in 1979 (42). Several variations of this procedure have been used. Electrophoresis kits containing premade gels, staining solutions, and all the necessary apparatus are available from several suppliers. Most systems use a 0.5% to 1.0% gel with either a barbital or tris-boric acid-ethylenediaminetetraacetic acid buffer at pH 8.6. Fully automated electrophoresis systems have been developed that eliminate many of the manual manipulations normally required in the procedure.

For a detailed description of a standard agarose gel electrophoresis method, the reader is referred to Jeppsson et al. (42). In this method, a serum sample is first spotted on the surface of the agarose support medium and allowed to absorb into the gel. The gel is then connected to two electrodes by filter paper wicks saturated with buffer. The electrodes are attached to a constant-voltage power supply. A current is then passed through the gel to separate the proteins. When electrophoresis is complete, the gel is removed from the electrode apparatus and placed in a bath of fixing solution. The fixing solution usually contains acetic acid, which denatures the proteins and immobilizes them in the gel. This prevents diffusion of the protein bands in the support medium. Proteins in an agarose gel are invisible. Therefore, a staining solution is needed to visualize the bands. A variety of stains have been used, including Coomassie brilliant blue, amido black, and Ponceau S, which differ in their protein-binding affinities, sensitivity, and absorbance spectra. Coomassie brilliant blue is the most sensitive of the commonly used stains. Highly sensitive silver staining methods are available but are seldom necessary in routine practice. After the proteins have been stained, the gel is placed in a destaining solution (usually a mixture of methanol, acetic acid, and water) that clears the excess stain from the gel, producing sharp protein bands on a transparent background. The finished gel is quite fragile and is not suitable as a permanent record. However, the gel may be dried on a sheet of clear plastic, creating a durable record that can be stored for subsequent review.

Standard agarose gel electrophoresis systems separate serum into five zones designated albumin, α_1, α_2, ß, and γ (Fig. 13.2). The albumin band is relatively homogeneous, whereas the other zones are composed of mixtures of different proteins. Interpretation of electrophoretograms requires practice and a thorough understanding of the disease processes that may affect the protein pattern. The most useful information is obtained by careful visual inspection of the stained gel. In addition, the individual fractions can be quantitated by scanning the electrophoretogram with a densitometer. Most commercial kits are equipped with a scanning densitometer that produces a two-dimensional tracing of the electrophoresis pattern (Fig. 13.1). An integrator is used to calculate the area under each of the zones, which can then be expressed as a percentage of the total area under the tracing. If the total protein concentration is known, the results can be converted into grams per liter because the area under each zone is proportional to its concentration. The reference range for each of the five zones depends on the electrophoresis method and also varies somewhat with the age, gender, and race of the reference population and various other minor factors. A typical reference range is shown in Table 13.4.

TABLE 13.4. TYPICAL SERUM PROTEIN ELECTROPHORESIS REFERENCE PATTERN

Zone	Percentage of Total Protein (mean)
Albumin	45.8–68.2 (57.0)
α	2.1–6.1 (4.1)
α_2	8.8–18.0 (13.4)
ß	7.8–13.0 (10.4)
γ	7.8–22.6 (15.2)

Quantitative densitometric scanning provides a good measurement of the total serum albumin and a crude estimate of the γ-globulin fraction. It is not reliable for measuring other serum proteins because each zone is composed of a mixture of proteins that are not easily resolved by the densitometer. For this reason, densitometric scanning is of limited value. Most laboratories rely primarily on visual interpretation of the overall electrophoresis pattern.

The proteins that make up each of the five electrophoretic zones are listed in Table 13.2. Most of these were discussed in the preceding sections.

Albumin migrates to the far anodal region of the gel. The mobility of the albumin band may be affected by both congenital and acquired alterations in the molecule. Individuals who are heterozygotes for genetic albumin variants may show a widening of the band, or it may be split into two separate bands (bisalbuminemia). Hyperbilirubinemia, bound fatty acids, and some drugs (aspirin, penicillin) may increase the mobility of albumin (43). The significance of increases and decreases in the serum albumin was discussed previously. Sometimes a faint prealbumin band may be seen anodal to albumin.

The α_1 zone contains mostly AAT with minor contributions from several other proteins (Table 13.2). Although AAT cannot be reliably quantitated on serum electrophoretograms, a marked decrease in the AAT band or a band with altered mobility may suggest AAT deficiency and prompt further investigation.

The principal components of the α_2 zone are AMG and haptoglobin. The ß-lipoprotein band (LDL) may be found in the α_2 region on some systems but more commonly appears in the ß zone. AMG and haptoglobin usually merge into a composite fraction but are sometimes resolved into two separate bands. In patients with hemolytic anemia, haptoglobin may be noticeably decreased.

The major components of the ß zone are transferrin, the third component of complement (C3), and the ß-lipoproteins. In iron deficiency anemia, the transferrin band may be increased, although this finding is not specific. The C3 band is quite labile and is often absent unless fresh serum is used. The conversion product of C3 (C3c) sometimes appears in the fast ß region. A true decrease in C3 may occur in a variety of inflammatory disorders, including systemic lupus erythematosus. The ß-lipoproteins appear as a wavy band with a sharp leading edge and a smudgy tail. Increases in this band are seen in various conditions, including liver diseases, the nephrotic syndrome, malnutrition, and some familial hyperlipidemia syndromes.

The immunoglobulins are found in the γ region. This heterogeneous group of proteins is distributed throughout the entire

region, and they do not ordinarily produce discrete bands. The finding of a band in the γ-region should alert the observer to the possibility of a monoclonal paraprotein. Other proteins including CRP, lysozyme, and traces of fibrinogen may also produce bands in this zone. Normally, fibrinogen is not seen on SPE, but occasionally a faint band appears in patients receiving anticoagulant therapy. Marked elevations in CRP occur in the acute-phase response, which may produce a faint band in the γ region.

The most important use of SPE is for the detection of monoclonal paraproteins. The analysis of paraproteins is discussed in detail in Chapter 44, Chronic Lymphoproliferative Disorders.

In addition to the abnormalities in the individual aforementioned proteins, certain diseases produce alterations in more than one protein, which affect the overall electrophoretic pattern. These patterns may be recognized on visual examination of the electrophoretogram (Fig. 13.2) or by quantitative densitometric scanning.

The acute inflammatory pattern produces predictable changes based on the behavior of the positive and negative APRs. The albumin band is normal or decreased, while the α_1 and α_2 globulins show a marked increase. The transferrin band is usually decreased, and the γ-globulins, at least initially, appear normal. The acute-phase response may in time evolve into a chronic inflammatory pattern. Albumin continues to be decreased, and the α-globulins remain elevated. In addition, a diffuse polyclonal increase in γ-globulins appears, reflecting increased antibody synthesis. This feature serves to distinguish the acute from the chronic inflammatory pattern.

Cirrhosis produces a decrease in both albumin and the α-globulins owing to defective hepatocellular synthesis. This is associated with an increase in the β-lipoprotein band and a polyclonal hypergammaglobulinemia. The increase in γ-globulins may result in fusion of the β- and γ-globulin zones, which has been called β-γ bridging.

The electrophoretic pattern in patients with acute liver disease is quite variable. The most common pattern is a decrease in albumin with increases in the α- and γ-globulins. However, other patterns may be seen, including an isolated decrease in albumin, diffuse hypergammaglobulinemia, or a normal pattern.

Patients with nonselective protein-losing syndromes such as burns or exudative disorders of the skin, lung, and gastrointestinal tract may have decreases in all the serum proteins. This pattern may be superimposed on acute or chronic inflammatory changes, which cause an increase in some bands in the α and γ regions. Malnutrition and malabsorption result in generalized hypoproteinemia.

In the nephrotic syndrome, there is a selective protein loss. The α_2 region is often noticeably increased because of the increase in AMG. The β-lipoprotein band may be prominent, reflecting the hyperlipidemia seen in nephrotic syndrome. However, albumin is markedly decreased, as are most of the other fractions. The large increase in AMG differentiates this pattern from that seen with nonselective protein losses.

An alteration in the electrophoretic pattern is often seen in pregnancy and in some patients taking oral contraceptives. Pregnancy is associated with an increase in the plasma volume, which causes a dilutional decrease in albumin. In late pregnancy, an increase in α_1-glycoprotein may elevate the α_1 globulins, while the α_2 and β-globulins are elevated by increases in CER, β-lipoprotein, and transferrin. The γ-globulins may also be slightly elevated.

Capillary Electrophoresis

In capillary electrophoresis, electrophoretic separation of analytes is carried out in fused-silica capillaries with an inner diameter of 50 μm (range, 20 to 200 μm), an outer diameter of 375 μm, and an effective length of approximately 50 cm (range, 7 to 100 cm) (53). The high surface-to-volume ratio leads to fast dissipation of heat generated by the electric current. This allows the use of electric fields at least 50-fold more powerful than with traditional slab gels and therefore the rapid separation of proteins. Separation times of 5 to 15 minutes can be achieved. The small size of the capillary also allows sample sizes in the nanoliter range (54). Detection methods used for capillary electrophoresis include ultraviolet absorbance, fluorescence, and conductivity (53). The method can be fully automated, and commercial systems equipped with automated sample loading and data analysis are available (8). In a comparison of this method with conventional electrophoresis, it was found to be a reliable and reproducible technique with the advantages of high separation efficiency, on-line data analysis, quick separation, and ease of automation (55).

Other Methods

Various methods are available for the study of proteins. These include methods that separate proteins based on molecular size (ultrafiltration, density gradient centrifugation, molecular exclusion chromatography), solubility (isoelectric precipitation, salt precipitation, solvent fractionation), electric charge (electrophoresis, isoelectric focusing, ion exchange chromatography), and selective adsorption (affinity chromatography). Several of these methods have been used extensively in research. For further details, the reader is referred to standard biochemistry texts. Only a few of these techniques are used with any frequency in the clinical laboratory and are described briefly in the following sections.

Ion Exchange Chromatography

All forms of chromatography contain a stationary phase and a mobile phase. The mobile phase provides the driving force to separate molecules, while the stationary phase may partially or completely retard their movement, thus permitting separation. Molecules that differ in their size, charge, or other physical parameters can be resolved by selecting appropriate stationary and mobile phases.

Ion exchange chromatography uses a charged supporting medium to separate molecules based on their acid-base properties. The supporting medium consists of minute particles (“beads”) to which fixed charges are attached. The beads are loaded in a plastic or glass column. The sample is then placed on top of the column and allowed to percolate through the medium. Some molecules bind tightly to the fixed charges, some

bind only weakly, and others do not bind at all. Those that do not bind pass unhindered through the column and are collected in a fraction tube. Molecules that remain bound to the beads are then sequentially eluted with buffers of increasing ionic strength or varying pH, which displace the bound molecules from the column.

Ion exchange resins are often derivatives of cellulose. Carboxymethylcellulose contains fixed negative charges at pH 7.0 and thus is called a cation exchange resin because cationic molecules bind by exchanging for sodium ions bound to the negative charges on the resin. In contrast, diethylaminocellulose contains fixed positive charges at neutral pH and is used as an anion exchange resin. By selecting an appropriate resin, sample pH, and elution buffers (pH and ionic strength), complex protein mixtures can be separated for further study.

Ion exchange methods have been used to separate isoenzymes, glycosylated hemoglobins, and genetic variants of some proteins. Minicolumns for separating certain proteins are available in kit form.

Affinity Chromatography

Affinity chromatography is based on the principle that some proteins will bind by noncovalent forces to specific ligands. Examples of protein ligand combinations include hormone-receptor pairs, enzyme-coenzyme pairs, and antigen-antibody pairs. Because the ligand may be highly specific to a particular protein, affinity chromatography permits a high degree of purification of the protein, often in a single step.

The method of affinity chromatography is similar to other column methods. The ligand is first immobilized to an inert supporting medium. This is then loaded into a column and the sample added to the top of the column. As the sample percolates through, the protein binds to its specific ligand. Other proteins flow through the column and are collected in a fraction tube. Then an eluting buffer is added to the top of the column. The eluting buffer may be a high ionic-strength salt solution or a solution containing a high concentration of some other compound that competes with the bound protein for the ligand. The protein is thus displaced from its ligand-binding site and is eluted from the column.

Affinity chromatography is one of the most powerful techniques used to purify proteins but to date has been used only on a limited basis in the clinical laboratory.

Isoelectric Focusing

Isoelectric focusing (IEF) or electrofusing is a type of electrophoresis that separates proteins based on their isoelectric point (PI) (see earlier discussion). A protein at its isoelectric point will not migrate in an electric field, and this feature has been exploited in IEF. IEF is performed in a glass tube filled with a gel supporting medium that contains a pH gradient extending from the top to the bottom of the tube. The sample is loaded on the top of the column and a current is applied to the apparatus. The proteins in the sample migrate through the pH gradient until they reach their isoelectric point, where they stop (focus). Because the isoelectric point of proteins changes with even minor changes in amino acid composition, complex mixtures of proteins can be separated by IEF.

Like affinity chromatography, IEF is a powerful method to separate proteins. However, its use in the clinical laboratory is currently restricted to specialized situations (such as detecting AAT variants), and most laboratories are not equipped to perform this technique.

REFERENCES

1. Laurell C, Jeppson J. Protease inhibitors in plasma. In: Putnam FW, ed. *The plasma proteins*, 2nd ed. New York: Academic Press, 1975:229–263.

2. Schreiber G. Synthesis, processing and secretion of plasma proteins by the liver and other organs and their regulation. In: Putman FW, ed. *The plasma proteins*. New York: Academic Press, 1987:293–349.

3. DeCresce R. Total serum protein: review of methods. *Am Soc Clin Pathol Check Sample Series* 1988;4:1–7.

4. Rose B. The total body water and the plasma sodium concentration. In: Rose B, ed. *Clinical physiology of acid base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989:211–247.

5. Haider M, Maider SQ. Assessment of protein-calorie malnutrition. *Clin Chem* 1984;30:1286–1299.

6. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med* 1999;340:448–454.

7. Hart W. C-reactive protein: the best laboratory indicator for monitoring disease activity. *Cleve Clin J Med* 1989;56:126–130.

8. Johnson AM, Rohlfs EM, Silverman LM. Proteins. In: Burtis CA, Ashwood ER, eds. *Tietz textbook of clinical chemistry*, 3rd ed. Philadelphia: Saunders, 1999:477–540.

9. Bienvenu J, Jeppsson J-O, Ingenbleek Y. Transthyretin (prealbumin) and retinol binding protein. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*. Scarborough: Foundation for Blood Research, 1999:9.0-1–9.01-7.

10. Moskowitz SR, Pereira G, Spitzer A, et al. Prealbumin as a biochemical marker of nutritional adequacy in premature infants. *Pediatrics* 1983;102:749–753.

11. Hutchinson DR, Halliwell RP, Smith MG, et al. Serum “prealbumin” as an index of liver function in human hepatobiliary disease. *Clin Chim Acta* 1980;114:69–74.

12. Kratz A, Lewandrowski KB. MGH case records: normal reference laboratory values. *N Engl J Med* 1998;339:1063–1073.

13. Johnson AM, Guder WG. Albumin. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*, 1st ed. Scarborough: Foundation for Blood Research, 1999:6.00-1–6.00-12.

14. Peters T. Serum albumin. In: Putnam FW, ed. *The plasma proteins, volume 1*. New York: Academic Press, 1975:133–173.

15. Herrmann FR, Safran C, Levkoff SE, et al. Serum albumin level on admission as a predictor of death, length of stay, and readmission. *Arch Intern Med* 1992;152:125–130.

16. Ruiz M, Rajatanavin R, Young RA, et al. Familial dysalbuminemic hyperthyroxinemia. *N Engl J Med* 1982;306:635–639.

17. Keller H, Morell A, Noseda G, et al. Analbuminemie: Pathophysiologische Untersuchungen an einem Fall. *Schweiz Med Wochenschr* 1972;306:635.

18. Gendler S. Proteins. In: Kaplan L, Pesce A, eds. *Clinical chemistry: theory, analysis, and correlation*. St. Louis: Mosby, 1989:1029–1065.

19. Morse JO. Alpha1-antitrypsin deficiency. *N Engl J Med* 1978;299:1045–1048.

20. Brantly M, Nukiwa T, Crystal RG. Molecular basis of alpha-1-antitrypsin deficiency. *Am J Med*1988;84(Suppl 6A):13–31.

21. Bearn A, Hartwig C. Genetic variations of plasma proteins. In: Stanbury JB, Wyngaarden JB, Fredrickson D, eds. *The metabolic basis of inherited disease*, 2nd ed. Philadelphia: Saunders, 1972:814–885.

22. Kobzik L. The lung. In: Cotran RS, Kumar V, Collins T, eds. *Robbins pathologic basis of disease*, 6th ed. Philadelphia: Saunders, 1999:697–755.

23. Cox DW. Alpha-1 antitrypsin deficiency. In: Scriver CR, Beaudet AL, Sly WS, et al., eds. *The metabolic and molecular bases of inherited disease. volume 3*. New York: McGraw-Hill, 1995:4125–4158.

24. Ruebner B, Cox K. Liver disease in infancy. In: Peters R, Craig J, eds. *Liver pathology*. New York: Churchill-Livingstone, 1986:37–60.

25. Borth W. Alpha-2-macroglobulin, a multifunctional binding protein with targeting characteristics. *FASEB J* 1992;6:3345–3353.

26. Ohlson K. Elimination of 125-I-trypsin alpha-macroglobulin complexes from blood by reticuloendothelial cells in dog. *Acta Physiol Scand*1971;81:269–272.

27. Davis AEI. Alpha-2-macroglobulin. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*. Scarborough: Foundation for Blood Research, 1999:8.02-1–8.02-8.

28. Feige J-J, Negoescu A, Keramidas M, Souchelnitskiy M, et al. Alpha-2-macroglobulin: a binding protein for transforming growth factor-beta and various cytokines. *Horm Res* 1996;45:227–232.

29. Linder MC, Hazegh-Azam M. Copper biochemistry and molecular biology. *Am J Clin Nutr* 1996;63:797S–811S.

30. Logan JI, Harveyson KB, Wisdom GB, et al. Hereditary caeruloplasmin deficiency, dementia and diabetes mellitus. *Q J Med* 1994;87:663–670.

31. Danks DM. Disorders of copper transport. In: Scriver CR, Beaudet AL, Sly WS, et al., eds. *The metabolic and molecular bases of inherited disease, volume 2*, 7th ed. New York: McGraw-Hill, 1995:2211–2235.

32. Maier-Dobersberger T, Mannhalter C, Rack S, et al. Diagnosis of Wilson's disease in an asymptomatic sibling by DNA linkage analysis. *Gastroenterology* 1995;109:2015–2018.

33. Gutteridge JMC. The antioxidant activity of haptoglobin towards haemoglobin-stimulated lipid peroxidation. *Biochim Biophys Acta* 1987;917:219–223.

34. Snellman O, Sylven B. Haptoglobin acting as a natural inhibitor of cathepsin B activity. *Nature* 1967;216:1033.

35. Ghamati SME, Hoeyveld EMV, Van Strijp JAG, et al. Identification of haptoglobin as an alternative ligand for CD11b/CD18. *J Immunol* 1996;156:2542–2552.

36. van Rijn HJ, vander Wilt W, Stroes JW. Is the turbidimetric immunoassay of haptoglobin phenotype dependent? *Clin Biochem* 1987;20:245–248.

37. Jeppsson J-O. Haptoglobin. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*, 1st ed. Scarborough: Foundation for Blood Research, 1999:7.04-1–7.04-6.

38. Keren DF. *High resolution electrophoresis and immunofixation*, 2nd ed. Boston: Butterworth-Heinemann, 1994.

39. Putnam FW. Transferrin. In: Putnam FW, ed. *The plasma proteins, volume 1*, 2nd ed. New York: Academic Press, 1975:265–315.

40. Jeppsson J-O, Kristenson H, Fimani C. Carbohydrate-deficient transferrin quantified by HPLC to determine heavy consumption of alcohol. *Clin Chem* 1993;39:2115–2120.

41. Jeppsson J-O, Aguzzi F. Transferrin. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*. Scarborough: Foundation for Blood Research, 1999:9.02-1–9.02-8.

42. Jeppsson J-O, Laurell C-B, Franzen B. Agarose gel electrophoresis. *Clin Chem* 1979;25:629–638.

43. Messner RP. Beta-2-microglobulin: an old molecule assumes a new look. *J Lab Clin Med* 1984;104:141–145.

44. Durie BGM, Stock-Novack D, Salmon SE, et al. Prognostic value of pretreatment serum beta-2-microglobulin in myeloma: a southwest oncology group study. *Blood* 1990;75:823–830.

45. Kantarjian HM, Smith T, Estey E, et al. Prognostic significance of elevated serum beta-2-microglobulin levels in adult acute lymphocytic leukemia. *Am J Med* 1992;931:599–604.

46. Lifson AR, Hessol NA, Buchbinder SP, et al. Serum beta-2-microglobulin and prediction of progression to AIDS in HIV infection. *Lancet* 1992;339:1436–1440.

47. Fahey JL, Taylor JM, Detels R, et al. The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. *N Engl J Med* 1990;322:166–172.

48. Bienvenu J, Whicher JT, Chir B, et al. C-reactive protein, CRP. In: Ritchie RR, ed. *Serum proteins in clinical medicine, volume 1*. Scarborough: Foundation for Blood Research, 1999:7.00-1–7.00-5.

49. Gotschlich EC. C-reactive protein: a historical overview. *Ann N Y Acad Sci* 1988;553:9–15.

50. Harrison H, Levitt M. Serum protein electrophoresis: basic principles, interpretations, and practical considerations. *Am Soc Clin Pathol Chem Check Sample Series* 1987;3:1–16.

51. McManamon T, Lott J. Serum protein electrophoresis. In: Kaplan L, Pesce A, eds. *Clinical chemistry*, 2 ed. St. Louis: Mosby, 1989:1054–1057.

52. Aguzzi F, Jayakar SD, Merlini G, et al. Electrophoresis: cellulose acetate vs. agarose gels, visual inspection vs. densitometry. *Clin Chem* 1981;27:1944–1945.

53. Landers JP. Clinical capillary electrophoresis. *Clin Chem* 1995;41:495–509.

54. Karger BL, Chu Y-H, Foret F. Capillary electrophoresis of proteins and nucleic acids. *Annu Rev Biophys Biomol Struct* 1995;24:579–610.

55. Kim JW, Park JH, Park JW, et al. Quantitative analysis of serum proteins separated by capillary electrophoresis. *Clin Chem* 1993;39:689–692.

14

Diagnostic Enzymology and Other Biochemical Markers of Organ Damage

Alan H. B. Wu

Enzymes play a vital role in catalyzing the biochemical reactions necessary for normal human growth, maturation, and reproduction. Diagnostic enzymology involves the measurement of enzymes in body fluids for the diagnosis of disease. In most cases, serum or blood levels are the most useful, although urine, cerebrospinal, and extracellular fluid levels are sometimes important.

This chapter focuses on the analytical aspects and clinical significance of important enzymes that are measured for diagnostic purposes. The major emphasis is on the use of these enzymes for the diagnosis of diseases involving the liver, myocardium, skeletal muscle, pancreas, and bone. In the case of myocardial infarction and skeletal muscle injury, nonenzyme markers such as myoglobin and troponin are also discussed.

BASIC CONCEPTS FOR DIAGNOSTIC ENZYMOLOGY

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Pathologic Role of Enzymes in Blood

Most enzymes that are used for diagnostic purposes have no direct physiologic role in the blood. Their presence under normal circumstances is the result of natural aging and turnover of cells. Enzymes released into the blood are usually cleared from the circulation by metabolism within the reticuloendothelial system, although some are cleared by the kidneys (e.g., amylase). The normal level of activity of these enzymes in the serum is a function of the rate of release and clearance. The half-lives of these enzymes vary greatly from an estimated 2 hours for the BB isoenzyme of creatine kinase (CK) to 170 hours for placental alkaline phosphatase (ALP).

High levels of enzymes in the blood can indicate increased cellular turnover and tissue necrosis caused by disease. This was first demonstrated by LaDue et al. (1) in 1954, when they observed that aspartate aminotransferase activity increased after acute myocardial infarction (AMI). In subsequent studies using carbon tetrachloride–induced liver injury in rats, these investigators showed that high serum enzyme levels were caused by leakage from the damaged hepatocytes. Once released from cells, enzymes can either pass directly into the blood (if the tissue is highly vascular and there is no obstruction to flow) or reach the bloodstream through slower lymphatic drainage. The subcellular origin plays a role in how readily an enzyme will appear in the blood. Soluble cytoplasmic enzymes pass through one set of membranes, whereas mitochondrial enzymes must also pass through a second set of membranes.

An equally important and largely overlooked cause of high enzyme levels in blood is tissue synthesis of new enzymes that occurs in response to disease, induction by various drugs, and carcinogenesis (2). For example, in obstructive liver disease, hepatic ALP is increased largely because of increased synthesis owing to abnormal intracanalicular pressure. Phenobarbital and several other drugs induce microsomal enzyme production that leads to increased synthesis of γ-glutamyltransferase (GGT). In addition, enzymes normally absent in serum can appear as the result of a neoplastic process.

FACTORS IN ENZYME ASSAY DEVELOPMENT

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Enzyme Kinetics and Substrate Concentration

Enzymes (E) are biological catalysts that combine with substrates (s) to form intermediate complexes (ES) leading to the formation of product (P) and restoration of the original enzyme (3). Reactions involving one substrate exhibit the following kinetics:

$S + E \rightarrow [ES] \rightarrow P + E$

Amylase and lipase are examples of enzymes of this scheme, but most other enzymes catalyze more complex reactions in which there are two or more substrates that form multiple products.

The reaction rate is a function of several factors, including substrate and enzyme concentration, temperature, pH, and the presence of cofactors, coenzymes, activators, and inhibitors. The rate equations for enzyme-catalyzed reactions were first developed by Michaelis and Menten in 1913 after the empirical observation that reaction velocity is proportional to substrate concentration at relatively low values (3). At high concentrations, the reaction rate is maximized. This behavior is shown in Fig. 14.1. In the region where reaction velocity is directly proportional

to substrate concentration, the reaction is termed first order with respect to substrate. At a high substrate concentration, the velocity is independent of substrate concentration, and the reaction is termed zero order with respect to substrate. The Michaelis-Menten constant K_m is defined as the substrate concentration that corresponds to one half of the maximum reaction rate.

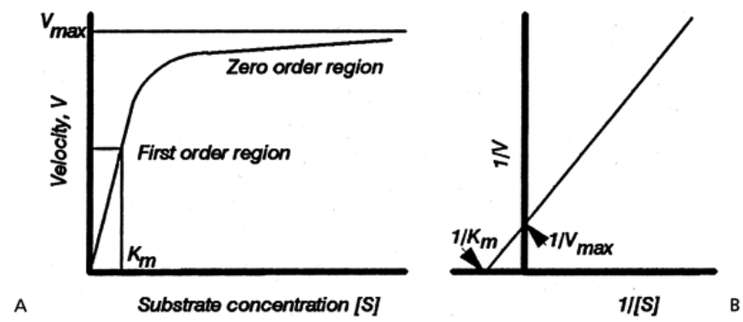


FIGURE 14.1. A: General form of the relationship between substrate concentration and the velocity of an enzyme-catalyzed reaction. V_{max} denotes maximum velocity and K_m denotes the Michaelis-Menten constant (see text). **B:** Reciprocal plots of substrate versus velocity (Lineweaver-Burke plots).

In diagnostic enzymology, the goal is the measurement of the activity or concentration of the enzyme itself. Therefore, test conditions are chosen such that the reaction rate will depend only on the amount of enzyme present in the sample to be measured. The substrate concentration should be in the zero-order region, and all other factors that influence the reaction rate should be optimized and controlled.

Selection of Other Reaction Conditions

Many parameters must be considered when developing a reagent formulation for measuring enzymes. Most enzymes operate optimally within a narrow pH range. Values above or below this range can inactivate enzymes by altering the tertiary structure of the active site or by denaturing the protein. The pH can also influence the direction of the reaction, such as for lactate dehydrogenase (LDH) for which the conversion of lactate to pyruvate is favored at alkaline pH, whereas the reverse reaction is favored at physiologic pH. The specific buffer chosen to maintain pH can also be important, as some buffers participate in the reaction and increase its rate. For example, for ALP, the reagent 2-amino-2-methyl-1-propanol (AMP) acts both as a buffer and as a phosphate group acceptor.

Cofactors and coenzymes not present in the sample or supplied by the buffer must also be added to the reaction mixture to ensure optimal activity. Pyridoxal phosphate is a necessary coenzyme for both aspartate aminotransferase (AST) and alanine aminotransferase (ALT), whereas magnesium is a cofactor needed for ALP. Other substances that are needed to enhance activity are called activators and are also added to the reaction mixture. For example, *N*-acetyl-L-cysteine is used to activate CK.

The temperature used for the reaction must be tightly controlled because reaction rates generally double with every 10°C that the temperature increases. High temperatures can denature many proteins and enzymes. By convention, most enzymes are assayed at 37°C. Although this choice is somewhat arbitrary, having a standard temperature facilitates the comparison of test results from one laboratory to another.

Inhibitors

Inhibitors are substances that can decrease the rate of enzyme-catalyzed reaction. They interact with the enzyme in a number of ways. *Competitive* inhibitors compete with substrate for binding to the active site on the enzyme. The effect of a competitive inhibitor is to increase K_m while not affecting the maximum velocity (V_{max}). Increasing the substrate concentration counteracts the effect of a competitive inhibitor. *Noncompetitive* inhibitors decrease the reaction rate by interacting with the enzyme without directly affecting the enzyme-substrate binding. V_{max} is decreased with no effect on K_m. In this case, the effect cannot be counteracted by increasing the substrate concentration. *Uncompetitive* inhibitors interact with the enzyme-substrate complex, prohibiting the formation of product. Both V_{max} and K_m are altered.

Inhibitors play an important role in diagnostic enzymology. Inhibitors are used to selectively eliminate activity from specific isoenzymes, providing a means for detecting other isoenzymes. Tartrate, for example, is used to inhibit some acid phosphatase isoenzymes, whereas dibucaine is used to detect the presence of genetic variants of serum cholinesterase.

ENZYME ASSAY PROCEDURES

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Enzyme measurements have traditionally used kinetic assays to determine activity. Newer assays have been developed, however, that measure protein (mass) concentrations. Both types of assay are discussed in this section.

Single-Reagent Kinetic Assays

Figure 14.2 illustrates a typical enzyme-catalyzed reaction for a single-reagent kinetic assay. The reaction is monitored at a wavelength at which the reaction product absorbs light. The initial lag phase is caused by the time necessary for the enzyme to be fully

activated by cofactors and coenzymes added to the reagent. A delay of 30 to 120 seconds is usually necessary before absorbance readings become meaningful. After the lag phase, the reaction begins the linear phase, characterized by a constant rate of change in absorbance. The reaction is monitored by taking readings at regular intervals for 2 to 5 minutes after the lag phase.

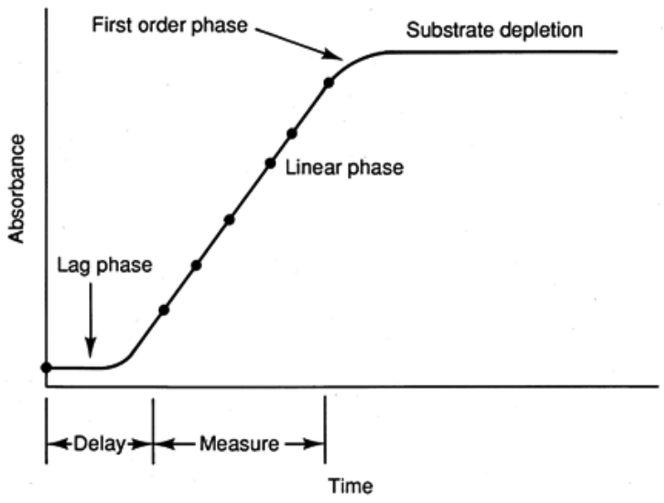


FIGURE 14.2. Absorbance versus time curve for a single-reagent assay system. The lag phase is the time necessary for activation of the enzyme. The linear phase (zero-order kinetics) is where the reaction rate is proportional to enzyme activity. The end of the first-order phase is where the substrate is depleted and the reaction ends.

If the reaction proceeds long enough, the substrate concentration becomes limiting, and the reaction enters a first-order phase. When the substrate is used up, the reaction stops. Samples with very high enzyme activity can deplete the substrate even before the first analytical reading is taken. In this case, falsely low results may be obtained (Fig. 14.3A), and the sample may appear the same as one with little or no enzyme activity (Fig. 14.3B). Enzyme analyzers must be able to detect samples with high enzyme activity. As shown in Fig. 14.3, a large difference in absorbance between the zero time point and the first reading may be used to detect the presence of a sample that has depleted the substrate. When this has occurred, the enzyme activity in these samples is obtained by dilution.

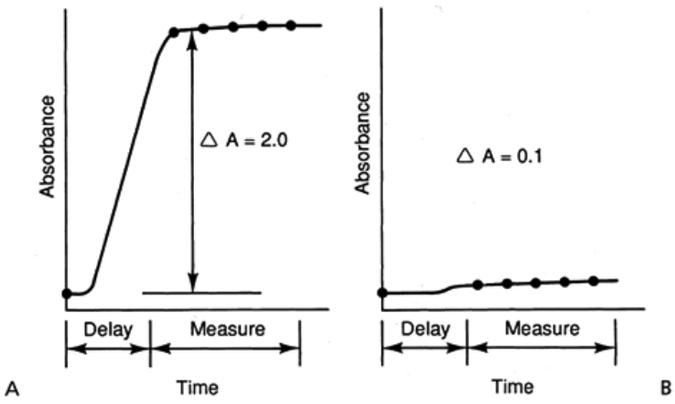


FIGURE 14.3. Absorbance versus time curves for samples with very high (**A**) and very low (**B**) enzyme activity. Because the measurement window is fixed, both samples appear to have low activity. Calculation of the difference in absorbance between the baseline and the first reading permits the distinction between samples with high enzyme activity (curve A) and those with low activity (curve B).

Start Reagent Activity Assays

Certain enzyme assay designs require the sequential addition of two reagents. The first includes the serum sample, which contains all reagents except one of the substrates. Incubation of this reagent allows both the activation of the enzyme and substrate-independent side reactions to occur, which, when present, erroneously contribute to the apparent activity of the enzyme. The analytical reaction of interest is triggered by the addition of the substrate (sometimes termed the *start reagent*) after an appropriate incubation period, usually 5 minutes. Figure 14.4 illustrates the reaction rate versus time profile for a start reagent assay. The analysis of ALT, for example, is improved by using the start reagent system, because high concentrations of pyruvate (which may occur in some serum specimens) will produce an apparent consumption of nicotinamide adenine dinucleotide (NADH) that is not dependent on ALT activity. Note that in Fig. 14.4, absorbance is shown decreasing as the reaction proceeds, indicating that the rate of consumption of a reactant is being followed.

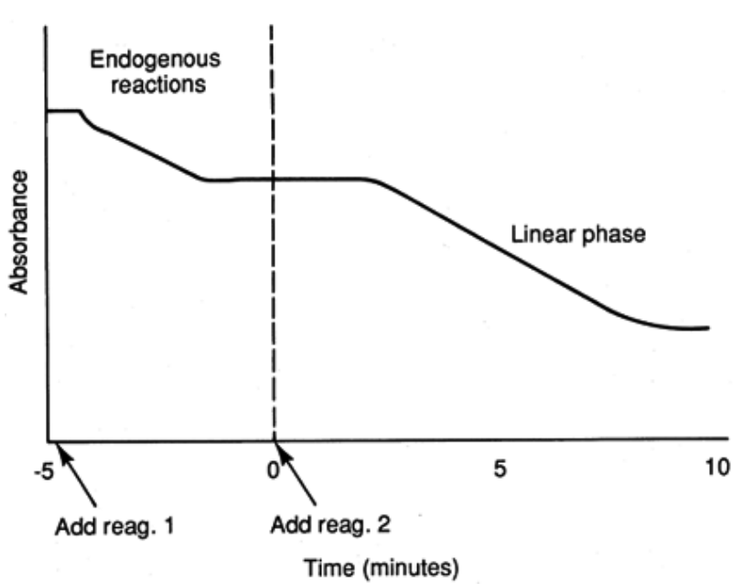


FIGURE 14.4. Absorbance versus time curve for the two-addition reagent assay system. The first reagent contains all reagents minus the substrate. The second reagent, the start or trigger reagent, initiates the reaction of interest, which is measured after a delay period.

Enzyme Activity Calculations

Enzyme activity can be expressed in either units per liter or nanokatals per liter. An International Unit is defined as the amount of enzyme necessary to catalyze the conversion of one micromole of substrate per minute. The corresponding SI unit is the amount of enzyme that will catalyze the conversion of one nanomole of substrate per second (nanokatal). Because Beer's law states that concentration is proportional to absorbance, enzyme activity can be determined from the change in absorbance per unit of time. The rate of change in absorbance is multiplied by factors to account for the molar absorptivity of the product or

substrate that is being monitored, the cell path length, and the dilution of the sample by the reagents in the reaction mixture:

$$\text{activity (U/L)} = \frac{\text{A/min} \times \text{TV}}{\epsilon \times \text{b} \times \text{SV}} \times 1{,}000$$

where A/min is the change in absorbance versus time, TV is the total volume (in microliters), SV is the sample volume (in microliters), ε is the micromolar absorptivity (in liters per millimoles · centimeters), and b is the path length (in centimeters). As an example, if an enzyme produces NADH (millimolar absorptivity = 6.3) with a ΔA/min of 0.046, and if the total and sample volumes are 300 and 5 µL, respectively, and a 1-cm path length is used, the activity (in U/L) is calculated as:

$$\text{activity} = \frac{0.046 \times 300}{6.3 \times 1 \times 5} \times 1{,}000 = 438 \text{ U/L}$$

Mass Measurements

The development of immunoassays has permitted the measurement of enzyme concentration rather than activity. These assays are particularly useful for isoenzyme analysis, as antisera can be directed toward specific isoenzymes, isoforms, or subunits. Studies have shown that for the detection of diseases, enzymes do not need to be present in an active form in blood, and mass measurements are equivalent to enzyme activity assays in diagnostic efficiency and are sometimes preferred. The enzyme concentration is determined from a calibration curve. Mass measurements are also used to measure the concentration of protein markers (e.g., myoglobin) that do not possess enzymatic activities. Specific immunoassays for enzymes and proteins are discussed in more detail under specific examples.

LIVER DISEASES

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

The enzymes ALT and AST, ALP, LDH, GGT, and (to a lesser extent) 5′-nucleotidase (5′NT) are commonly measured for the assessment of liver function. The first four of these are typically included in comprehensive chemistry profiles. None of these markers is specific for any single liver disorder. This section covers liver dysfunction, describes individual liver enzymes and their laboratory measurement, and explains how patterns of liver enzyme data can be used to aid in the diagnosis of liver diseases.

Acute Hepatocellular Injury

Viral Hepatitis

Acute viral hepatitis is a common cause of hepatocellular injury. The isolation and characterization of specific viral agents in recent years, coupled with the development of specific serologic markers, have resulted in the definitive diagnosis of these infections. A more complete description of the virology, etiology, epidemiology, and relevant serology of hepatitis types A, B, C, and D can be found in Chapter 59. Also described in this section are the other viral causes of hepatitis such as Epstein-Barr virus, herpesviruses, cytomegalovirus, and coxsackieviruses, which generally produce milder forms of hepatocellular necrosis. The features of acute viral hepatic inflammation include infiltration of monocytes and hyperplasia of parenchymal hepatocytes, often described as ballooning degeneration. There is active hepatocellular destruction and cell damage, leading to a variable degree of biliary obstruction. The liver has considerable reserve capacity for bilirubin metabolism, however, and patients are often only mildly jaundiced. The acute phase of the disease typically lasts 3 to 6 months. The enzyme markers that are the most useful for acute hepatocellular necrosis are AST and ALT.

Acute Liver Failure

Acute liver failure is characterized by extensive hepatocellular necrosis and a very high mortality; it has several etiologies. One is the progression to fulminant hepatitis, most frequently seen with hepatitis B. Toxic hepatitis is also a major cause. Table 14.1 lists various substances that can produce acute liver failure. Clinical manifestations include metabolic dysfunctions such as encephalopathy and water and electrolyte imbalance, coagulation abnormalities, hypoglycemia, respiratory failure, and renal failure.

TABLE 14.1. CAUSES OF TOXIC HEPATITIS

Solvents	Carbon tetrachloride
	Trichloroethylene
Mushrooms	*Amanita phalloides*
Metals	Yellow phosphorus
Drugs	Acetaminophen
	Halothane
	Isoniazid
	Rifampicin
Metabolic	Reye's syndrome
	Wilson's disease
	Galactosemia
	α_1-Antitrypsin deficiency

Acetaminophen overdose is another cause of acute hepatic failure. With therapeutic uses of acetaminophen, the drug is conjugated within the liver to inactive metabolites and excreted into the urine. Accidental or intentional overdose results in a saturation of the hepatic conjugation capability. As a result, highly reactive neutrophilic intermediates form that bind to hepatocytes, causing liver damage over the ensuing 2 to 3 days. Management of acetaminophen overdoses begins with the recognition of the problem. Initially, results for liver function tests are normal. Nevertheless, acute overdoses should be treated by restoring glutathione stores with administration of *N*-acetyl-L-cysteine to prevent liver damage. If fulminant hepatic failure does develop, liver transplantation is the only curative measure. Measurement of acetaminophen involves hydrolysis with an aryl acylamidase to produce *p*-aminophenol and acetate. The *p*-aminophenol is converted to indophenol with *o*-cresol. As with hepatocellular disease, acute liver failure is characterized by large increases in serum AST and ALT.

Cholestatic Liver Diseases

Cholestasis is suppression of bile flow and can be either extrahepatic or intrahepatic in origin. Delineation between these

types is not possible solely based on enzyme testing, although some patterns of enzyme activities favor one form over the other. Morphologically, the presence of bile pigments in hepatocytes defines this condition. Serum levels of liver enzymes and bilirubin are abnormal in cholestasis. Because of the reserve capacity of the liver for metabolism and excretion, however, these enzymes and salts increase in the serum only when more than 80% of the hepatocellular functional capacity has become impaired. Chapter 21 describes the role of conjugated and unconjugated bilirubin in cholestatic disease. Enzyme markers of obstruction include ALP, GGT, 5′-NT, and leucine aminopeptidase.

Intrahepatic Obstruction

Intrahepatic cholestasis often results from cirrhosis or hepatitis, but administration of some drugs such as chlorpromazine, estrogens, and anabolic steroids can also induce cholestasis. The mechanism may include the precipitation of bile salts, which impairs the flow of bile. Idiopathic familial intrahepatic jaundice is also seen in the third trimester of some otherwise uncomplicated pregnancies. This condition is relatively benign but may recur in future pregnancies. Other congenital disorders associated with intrahepatic obstruction include Dubin-Johnson syndrome and Rotor's syndrome.

Extrahepatic Obstruction

Extrahepatic cholestasis is usually the result of mechanical obstruction of the common bile duct or hepatic duct. Common etiologies of this disorder include cholelithiasis and carcinoma of the tip of the pancreas, papilla of Vater, or common bile duct. These processes lead to dilation of the ductal tree immediately above the obstruction. Inflammatory processes such as pancreatitis, and bile duct lesions seen in sclerosing cholangitis can also produce extrahepatic cholestasis.

Chronic Liver Diseases

Chronic Hepatitis

Chronic hepatitis is a disease with diverse etiologies, including viral, autoimmune, and lupoid causes. In viral causes, some 5% to 10% of hepatitis B patients and as many as 50% of hepatitis C patients develop this recurring disease. The active form is characterized by monocyte infiltration and hepatic necrosis. Chronic active hepatitis commonly leads to disruption of lobular architecture and fibrosis and can progress to liver cirrhosis. In contrast, chronic persistent hepatitis is a milder disorder that rarely progresses to cirrhosis and fibrosis. The liver enzyme results tend to be more elevated in the active form; however, differentiation is best made by performing a liver biopsy.

Liver Cancer

Primary hepatocellular and cholangiocellular carcinomas often occur in patients with longstanding hepatitis B infections and liver cirrhosis. The diagnosis is aided by computed tomography, ultrasonography, nuclear scans, and α-fetoprotein levels in blood. Primary liver cancer is especially prevalent among Oriental and black African males. Metastatic liver disease is very common in the United States, particularly from tumors that originate in the colon, lung, breast, and skin. Enzyme levels in primary and secondary cancer are usually elevated; this is caused by the destruction of surrounding normal tissue and obstruction of normal architecture by these space-occupying tumors.

Alcoholic Liver Disease

Alcoholic liver disease is associated with acute and chronic alcoholism. There are three recognized entities associated with excess use of ethanol (defined by some as a daily intake of 80 g/d): fatty liver, alcoholic hepatitis, and cirrhosis. These entities may coexist with the same patient, or occur independently. The diseases do not necessarily follow a pathophysiologic sequence.

Fatty Liver

Alcoholic fatty liver is characterized by hepatomegaly, in which the liver contains large triglyceride-containing vacuoles. The accumulation of triglycerides results from the inhibition of ß-fatty acid oxidation within the mitochondria. In addition to alcoholism, fatty liver can also occur in nutritional imbalances and metabolic disease and during pregnancy. The concentrations of traditional liver enzymes are typically within the normal range. Definitive diagnosis is made by a liver biopsy. The disease is considered benign and can be effectively reversed by immediate abstinence from ethanol intake.

Alcoholic Hepatitis

Alcoholic hepatitis is characterized by irreversible hepatocellular injury and degeneration. The major feature is hepatic infiltration by neutrophils and lymphocytes causing acute inflammation. The diagnosis is made by liver biopsy. The finding of Mallory bodies is suggestive of alcoholic hepatitis. The disease is chronic and progressive and can be fatal. Gastrointestinal bleeding and hepatic encephalopathy can occur in severe forms of the disease. Alcoholic hepatitis is considered the precursor to cirrhosis in many cases. In addition to abstinence from alcohol, treatment with corticosteroids can be beneficial. Liver enzyme concentrations are increased, notably AST and ALT, commensurate roughly with the degree of acute hepatocellular injury.

Cirrhosis

Chronic liver disease that leads to an increase in diffuse fibrosis and progression toward a nodular architecture defines cirrhosis. Some of the etiologic factors for cirrhosis are liver damage secondary to drug use, chronic hepatitis infections, biliary obstruction, and metabolic disease (such as Wilson's disease, hemochromatosis, and α_1-antitrypsin deficiency). Alcoholic cirrhosis is very prevalent in the United States and western Europe, whereas cirrhosis owing to chronic active hepatitis B infections is more common in Asia and Africa. Liver enzyme levels in cirrhosis are variably elevated and can be normal during the terminal stages of the disease.

Primary biliary cirrhosis is characterized by impairment of bile excretion and has been linked to an autoimmune process. It affects women more often than men. Inflammation and destruction of bile ducts are observed in this disorder. Elevations in ALP and aminotransferases are expected along with high titers of antimitochondrial antibody.

MARKERS OF LIVER DISEASES

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Alkaline Phosphatase (ALP) (EC 3.1.3.1)

The phosphatases are a collection of sialoproteins located mostly within the cell membrane. They function to transport inorganic phosphate from donor to receptor molecules. Phosphatases with catalytic activity at a pH of 10.0 and higher make up the collection of ALP (orthophosphoric acid monoester phosphohydrolase).

Measurement of Total ALP Activity

Because of the heterogeneous nature of the enzyme and because a single physiologic substrate has not been identified, all analytical methods for ALP make use of synthetic substrates. The most common substrate is *p*-nitrophenylphosphate (*p*-NPP), which is converted to *p*-nitrophenoxide (*p*-NP) and inorganic phosphate:

$$p\text{-NPP} + H_2O \xrightarrow{ALP} PO_4^{-3} + p\text{-NP}$$

Because *p*-NPP is colorless, the reaction is monitored by the increasing absorbance of *p*-NP at 404 nm. The pH of the reaction is controlled at 10.3 by AMP buffer, which also acts as a phosphate acceptor to facilitate the phosphatase action of the enzyme. A magnesium salt is also part of the reagent formulation and acts as an activator of the enzyme. In other procedures, such as those recommended by the International Federation of Clinical Chemistry, zinc ions are added along with *N*-hydroxy-ethylenediaminetetraacetic acid to maintain the precise ratio of Mg^{2+}/Zn^{2+} necessary for optimal activation. The millimolar absorptivity of *p*-NP is 18.75 L/mmol · cm.

The reference range for ALP is dependent on age and relates to different stages of bone growth and demineralization during life. Figure 14.5 illustrates this dependency. Children have higher reference ranges than adults, particularly during adolescence. Older populations have increasing ALP reference ranges, reflecting increases in the incidence of osteoporosis in this population.

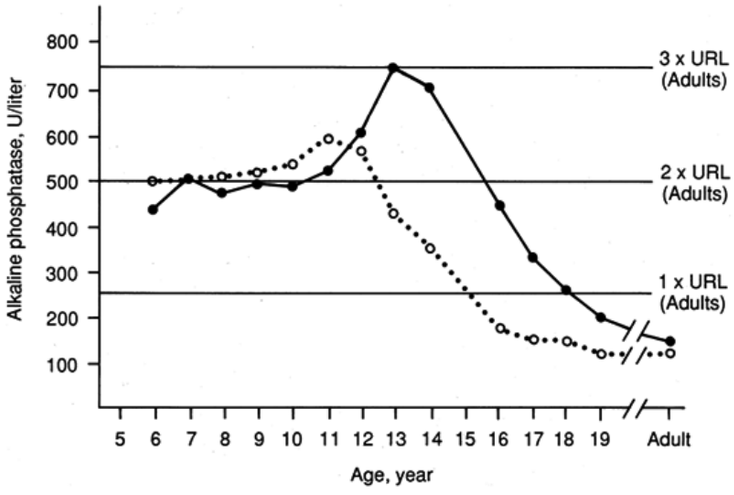


FIGURE 14.5. Plasma alkaline phosphatase activity as a function of age and sex (·--, men; o…o, women). Horizontal lines refer to multiples of the adult upper reference limit. (From McComb RB, Bowers GN Jr, Posen S. *Alkaline phosphatase*. New York: Plenum Press, 1979:532, with permission.)

ALP Isoenzymes

ALP is widely distributed throughout the tissues of the body. Among those with the highest relative activity are the adrenals, placenta, gastrointestinal tract, bone, kidney, and liver. In normal nonpregnant adults, however, most of the serum ALP originates from the bone and liver. For patients with high ALP levels, measurement of ALP isoenzymes is useful for differentiating between bone and liver sources. However, because of technical difficulties with zonal electrophoresis and the increasing use of more specific enzyme markers for hepatic obstruction, such as GGT, the clinical need for ALP isoenzymes is rare, and most laboratories do not offer routine fractionation. Analytical methods for ALP isoenzymes include selective chemical inhibition, such as with phenylalanine, heat inactivation, electrophoresis, and isoelectric focusing (IEF).

Zonal electrophoresis has been extensively studied and will produce bands for the placental, intestinal, and bone, liver, and kidney fractions. There are varying degrees of overlap between ALP isoenzymes from tissues of the last group, depending on the electrophoresis support medium. ALP isoenzymes from bone, liver, and kidney originate from a single gene expression that has been postsynthetically modified. Improvement in resolution between the bone and liver fractions is possible by treating either of the electrophoresis supports with wheat-germ lectin or by treating the samples themselves with neuraminidase. Electrophoretic studies have also demonstrated a new isoenzyme fraction migrating anodic to the liver fraction. The isoenzyme, termed the α or fast-liver fraction, is seen in hepatobiliary disorders and may be a useful marker for liver cancer. More definitive assessment of ALP isoenzymes is possible with IEF.

GGT (EC 2.3.2.2)

GGT (g-glutamyl-peptide:amino acid γ-glutamyltransferase) catalyzes the transfer of γ-glutamyl groups between donor and acceptor molecules. GGT is located in the cell membrane of nearly all human cells and tissues and functions to transport amino acids into the cell. GGT has two subunits: a light hydrophilic fraction, where the active site is located, and a heavy hydrophobic site, which anchors the enzyme to the membrane. It does not have isoenzymes as defined by separate genetically encoded forms, but there are posttranslational modifications that have been identified and measured by electrophoresis. The clinical utility of these isoforms, however, is not established (4).

Most commercial analytical methods for GGT make use of one of two synthetic substrates: γ-L-glutamyl-*p*-nitroanilide and γ-L-glutamyl-3-carboxy-4-nitroanilide. The carboxy substrate is currently preferred because of a higher solubility and because hydrolysis of the substrate does not occur spontaneously. The older unsubstituted substrate, however, is more widely used. The reactions of both are shown below:

$$\gamma\text{-glutamyl-}p\text{-nitroaniline} + \text{glycylglycine} \xrightarrow{\text{GGT}} \gamma\text{-glutamylglycylglycine} + p\text{-nitroaniline}$$

$$\gamma\text{-glutamyl-3-carboxy-4-nitroanilide} + \text{glycylglycine} \xrightarrow{\text{GGT}} p\text{-nitroaniline} + \text{5-amino-2-nitrobenzoate}$$

Either reaction can be monitored by following the increasing absorbance of the respective product at 405 nm. Tris and glycylglycine are used as buffers to maintain a pH of 8.0. Magnesium ions are added when *p*-nitroaniline is used to facilitate solubility. The millimolar absorptivities are 9.9 L/mmol · cm for *p*-nitroaniline at 405 nm and 7.9 L/mmol · cm for 5-amino-2-nitrobenzoate at 410 nm. The reference ranges at 37°C are 9 to 50 for men and 8 to 40 U/L for women. Although most of the enzyme originates from the liver, a small amount of GGT is released from the prostate, which may explain why men have slightly higher values than women.

5′-NT (EC 3.1.3.5)

5′-NT (5′-ribonucleotide phosphohydrolase) is a cytosolic membrane-bound enzyme that hydrolyzes the 5′-phosphate esters of nucleotides. Although 5′NT is found throughout the body, increased activity in the serum reflects hepatobiliary diseases. Methods for 5′NT analysis involve the use of adenosine-5′-monophosphate as a substrate:

$$\text{adenosine-5′-monophosphate} + H_2O \xrightarrow{5'\text{NT}} \text{adenosine} + \text{phosphate}$$

The reaction is monitored by measuring phosphate, using conventional methods. Because the substrate is also hydrolyzed by ALP in the serum sample, the total activity observed is the sum of 5′NT and ALP activities. The reaction is repeated with the addition of nickel ions, which selectively inhibit 5′NT, thus measuring only ALP activity. This number is then subtracted from the total to give a measure of 5′NT alone. Other inorganic ions used in the reaction mixture are manganese, to activate 5′NT, and copper, to accelerate color development.

AST (EC 2.6.1.1) and ALT (EC 2.6.1.2)

Although ALT (L-alanine:2-oxoglutarate aminotransferase) is found in a variety of human tissues, measurement of ALT is most clinically useful in the evaluation of liver diseases. ALT is found in the cytoplasm and only one isoenzyme is found in serum. Interpretations of ALT levels in these disorders are usually made in conjunction with levels of AST (L-aspartate:2-oxoglutarate aminotransferase). Unlike ALT, however, AST is also useful in the evaluation of AMI and skeletal muscle diseases. AST has cytosolic and mitochondrial isoenzymes that appear in the serum.

Analytical Measurements

ALT catalyzes the transfer of the amino group from alanine to α-ketoglutarate to form pyruvate and L-glutamate. AST catalyzes the transfer of the amino group from aspartate to α-ketoglutarate to form malate. Because none of the reactants or products of either of these reactions absorbs in the ultraviolet or visible region of the spectrum, they must be coupled to indicator reactions for analysis. ALT requires LDH, whereas AST requires malate dehydrogenase:

$$\text{L-alanine} + \alpha\text{-oxoglutarate} \xrightarrow{\text{ALT}} \text{pyruvate} + \text{L-glutamate}$$

$$\text{pyruvate} + \text{NADH} \xrightarrow{\text{LD}} \text{lactate} + \text{NAD}^+$$

$$\text{L-aspartate} + \text{-oxoglutarate} \xrightarrow{\text{AST}} \text{oxaloacetate} + \text{L-glutamate}$$

$$\text{oxaloacetate} + \text{NADH} \xrightarrow{\text{MDH}} \text{L-malate} + \text{NAD}^+$$

The reactions are monitored by following the decrease in the absorbance at 340 nm as NADH is consumed in the reaction. Pyridoxal-5′-phosphate (P5′P) is needed as a coenzyme for both the ALT and AST reactions. Normally, serum contains adequate amounts of P5′P and additional coenzyme is not necessary. However, patients with a vitamin B6 deficiency or those undergoing renal dialysis have decreased P5′P concentration, which will produce falsely low serum aminotransferase levels unless P5′P is added to the reagent formulation.

Reference Ranges

The adult reference range for both AST and ALT is roughly 10 to 40 U/L when measured at 37°C. Although men have slightly higher values than women, most laboratories use a single range for both genders.

PATTERNS OF LIVER ENZYMES AND MARKERS FOR INTERPRETATION OF DISEASE

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Unlike some disorders such as acute pancreatitis and myocardial infarction, for which there are enzyme markers that are primarily used for one disorder and have high diagnostic efficiencies, there are no enzyme markers that are specific for any single liver

disease. When evaluating these disorders, therefore, it is appropriate to consider a panel of markers, sometimes called liver function tests (LFTs). Through common usage, this term has come to mean a group of tests that usually includes bilirubin, AST, ALT, ALP, and sometimes GGT and 5′NT. However, the term is a misnomer, because although these tests can reflect various disease processes in the liver, they do not reflect hepatic reserve for synthesis and metabolic functions. Table 14.2 lists some markers for both liver function and disease.

TABLE 14.2. MARKERS FOR LIVER FUNCTION AND DISEASE

Disease or Function	Markers
Functional evaluation	
Normal synthesis capacity	Albumin, retinol binding protein, prealbumin, prothrombin time, cholinesterase
Excretory function	Bilirubin, bile acids, rose bengal and sulfobromophthalein excretion
Metabolic function	Ammonia, amino acids, lipids, serum protein electrophoresis, globulins
Drug metabolism	Antipyrine breath and clearance test $^{14}CO_2$ breath test
Pathologic evaluation	
Hepatocellular injury	Aspartate aminotransferase, alanine aminotransferase, lactate dehydrogenase
Obstruction	Bilirubin, alkaline phosphatase, γ-glutamyltransferase, 5′-nucleotidose, bile acids
Infections	Viral and autoimmune serologies
Malignancies	Carcinoembryonic antigen, α-fetoprotein

Hepatocellular Versus Obstructive Liver Diseases

Acute Injury

The most useful enzyme markers of acute hepatocellular injury are ALT and AST. Other markers such as LDH and glutamate, isocitrate, and malate dehydrogenases exhibit elevations secondary to hepatocellular necrosis, but they are not as sensitive or as specific as the combination of AST and ALT and are therefore rarely used for diagnostic purposes.

The highest absolute levels of ALT and AST are seen after acute hepatitis, either viral or toxic. Values exceeding 1,000 U/L are commonly seen during the early phases of these diseases. In toxic hepatitis, such as with acetaminophen overdoses, levels of ALT and AST rise within a few hours after exposure and remain high for many days or weeks. Unfortunately, the extent of ALT and AST elevations do not reflect the severity of the disease, nor can they be correlated to patient prognosis. In acute hepatitis A, ALT is one of the first markers to be elevated, usually 3 to 4 weeks after infection. Levels return to normal within 8 to 12 weeks. In acute hepatitis B, the preclinical incubation phase is longer and ALT and AST may remain normal for 2 to 6 months. In acute hepatitis B, ALT and AST return to normal within 2 to 3 months, and a similar course is observed in hepatitis C. In chronic active hepatitis, enzyme levels are elevated five- to 10-fold, depending on the stage of the disease. Similar elevations are also seen in acute hepatocellular injury secondary to mononucleosis. In end-stage liver disease, enzyme levels return to normal or subnormal levels as hepatocytes become depleted of enzyme content.

In contrast, levels of ALP, GGT, and 5′NT are not as markedly elevated in these disorders. Levels of these enzymes generally do not exceed two to three times the upper limit of normal. Therefore, a disproportionate increase in ALT and AST relative to ALP and GGT favors a diagnosis of hepatocellular necrosis rather than liver obstruction. The relationship of these enzymes in hepatocellular liver disease is summarized in Fig. 14.6 (5).

Cholestasis

The best markers for intrahepatic and extrahepatic cholestasis are ALP, GGT, and 5′NT. The largest elevations (four- to 10-fold) of ALP are typically seen in obstruction owing to gallstones or malignancy and in biliary cirrhosis. The source of ALP in malignancy may be obstruction of hepatic architecture in either primary or secondary liver cancers, but other causes of ALP elevations must be ruled out. Obstructive liver disease as a cause of ALP elevations can be confirmed by measurement of either GGT or 5′NT. Increases in these enzymes are also expected in cholestasis, but these assays are more specific because although bone disease can produce elevations of ALP in the same range expected for obstructive liver disease, GGT and 5′NT are not found in the bone. Measurement of total and direct bilirubin are also important in making the diagnosis of obstructive jaundice.

Cholestasis can be readily distinguished from hepatocellular injury based on a disproportionate increase of ALP and GGT relative to AST and ALT (Fig. 14.6). The latter two enzymes are generally only slightly elevated in cholestasis, rarely more than 500 U/L.

De Ritis Ratio (AST/ALT)

Further differentiation of specific liver diseases is aided by calculating the ratio of AST to ALT levels (6). Although there is considerable overlapping of values within a given diagnosis, the de Ritis ratio can give a general indication of whether a disorder is acute or chronic and whether it is intra- or extrahepatic in origin. Values for these ratios are most useful when standard AST and ALT assays are used. When using the assays recommended by the International Federation of Clinical Chemistry, the de Ritis ratio is normally approximately 1.15. Figure 14.7 illustrates results of the de Ritis ratio from patients with various liver diseases.

Acute Versus Chronic Liver Diseases

Acute disorders of the liver, such as acute viral hepatitis and infectious mononucleosis, generally have higher values of ALT relative to AST, and the de Ritis ratio is less than 1.0. In Reye's syndrome,

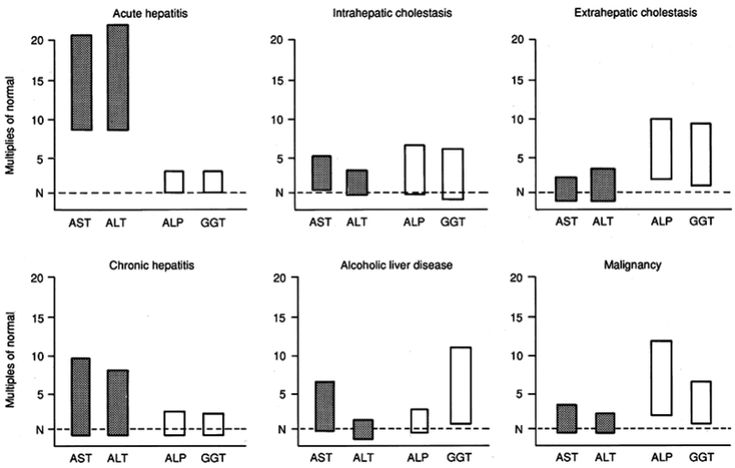


FIGURE 14.6. Relationship of alanine aminotransferase and aspartate aminotransferase to alanine phosphate and γ-glutamyltransferase in several liver diseases. (Adapted from Zimmerman HJ. Function and integrity of the liver. In: Henry JB, ed. *Clinical diagnosis and management*, 17th ed. Philadelphia: Saunders, 1984.)

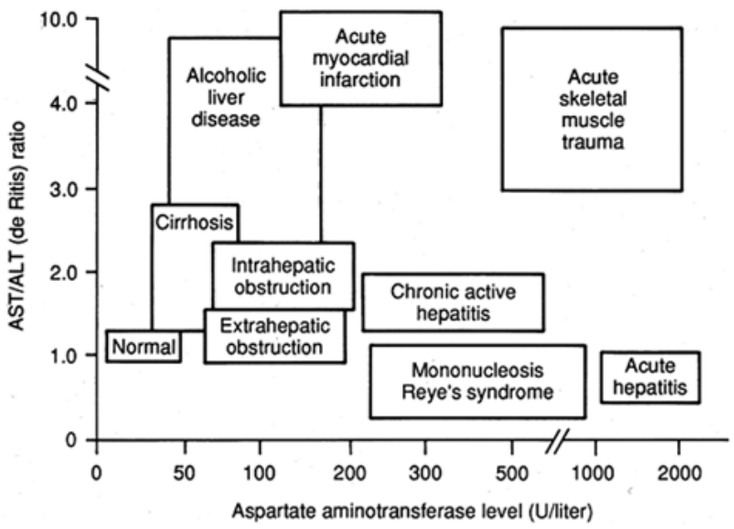


FIGURE 14.7. Ratio of aspartate aminotransferase:alanine aminotransferase versus aspartate aminotransferase levels in various liver diseases.

in which there is an acute onset of encephalopathy, the de Ritis ratio is also less than 1 in most cases. It has been postulated that elevations in ALT are seen before elevations in AST in acute liver processes because AST includes mitochondrial isoenzymes, and more time is needed for these enzymes to pass through a second set of membranes to reach the circulation. Indeed, chronic disorders such as alcoholic liver disease, postnecrotic cirrhosis, and chronic active hepatitis have de Ritis ratios greater than 1. Although this is a convenient way of remembering this relationship, the mitochondrial fraction of AST actually contributes only a small percentage of the total AST and is insufficient to significantly lower the de Ritis ratio (Fig. 14.7). Patients with chronic persistent hepatitis will have a normal or slightly elevated aminotransferase level, reflecting the mild clinical course of this disease.

Intrahepatic Versus Extrahepatic Obstruction

The de Ritis ratio may be helpful in differentiating between intra- and extrahepatic cholestasis (Fig. 14.7). Ratios of 1.5 or greater suggest intrahepatic cholestasis, whereas values less than 1.5 suggest an extrahepatic process. This difference is also consistent with the notion of acute versus chronic liver disorders. Extrahepatic obstruction is often caused by the acute passage of a stone, whereas intrahepatic causes such as biliary cirrhosis and malignancy develop over a period of time. Other laboratory test results such as high conjugated bilirubin (greater than 75%), high amylase, and the presence of occult blood in the stool favor extrahepatic rather than intrahepatic obstruction. In addition, ALP is more often elevated in extrahepatic than in intrahepatic obstruction.

Alcoholic Liver Disease

Disproportionate increases of AST relative to ALT are observed in alcoholic liver disease, and the de Ritis ratio can exceed 6.0 (Fig. 14.7). Decreased levels of ALT relative to AST are caused in part by deficiencies of dietary pyridoxal phosphate in alcoholics, a component that is more important for normal hepatic ALT synthesis than for AST.

Muscle Disease

The largest increases of AST relative to ALT are seen in myocardial and skeletal muscle diseases because of the different distribution of the enzymes in these tissues. A ratio greater than 10 can be seen in these disorders (Fig. 14.7), which are discussed more fully in the next section.

Markers of Alcohol Use

Ethanol metabolism follows zero-order kinetics, with a constant rate of clearance of approximately 20 mg/dL per hour. Ethanol is metabolized largely by alcohol dehydrogenase, with contributions from the microsomal ethanol oxidizing system, and catalase. Mildly intoxicated subjects (100 to 150 mg/dL) will have positive blood ethanol concentrations for only 5 to 8 hours after ingestion. For medical (and insurance) purposes, it is important to know whether an individual chronically consumes alcohol. The most widely used marker for alcoholism is GGT. GGT is highly sensitive to alcohol use, in part because ethanol is a potent inducer of GGT synthesis by the microsomal P-450 enzyme system. Significant elevations of GGT are also observed after prolonged drinking episodes of several months or more. Unfortunately, it is not a specific test, as other diseases and drugs can produce increases in GGT. Other markers currently being studied include carbohydrate-deficient transferrin and fatty acid ethyl esters (7). Excessive alcohol inhibits the posttranslational glycoconjugation of transferrin, leading to high concentrations of carbohydrate-deficient transferrin. This marker is thought to be a useful marker for chronic alcohol use. Fatty acid ethyl esters are produced as a metabolite of fatty acid and ethanol. Fatty acid ethyl esters may be responsible for ethanol-induced liver and pancreas damage. High concentrations may be clinically useful as an acute indicator of ethanol use within the previous 24 hours, even after serum ethanol concentrations have returned to normal.

Liver Transplantation and Markers of Rejection

The clinical indications of liver transplantation in adults includes end-stage chronic liver disease, cirrhosis (e.g., biliary, alcoholic), nonresectable hepatic tumors, and mushroom and acetaminophen poisoning. In children, liver transplantation is also performed in biliary atresia, hepatitis, Wilson's disease, tyrosinemia, and other forms of metabolic liver diseases. The suitability of the donor liver can be assessed after the infusion of lidocaine and the monitoring the rate of metabolism to monoethylglycinexylidide. High concentrations indicate a good metabolic capability and a higher rate of graft survival (8). Liver function tests such as AST, ALT, γ-GT, bilirubin, bile acids, and coagulation tests are typically used to determine the viability of the liver. Increased concentrations of these markers may indicate allograft rejection of the organ. The sensitivity of these markers, however, has not been high, and the rationale for using these tests can be questioned. Recently, new markers including glutathione S-transferase (9) and serum fructose-1-6-bisphosphatase (10) have been studied as markers for early liver transplantation rejection. These tests may become available in laboratories whose hospitals have an active liver transplantation program.

MYOCARDIAL AND SKELETAL MUSCLE DISEASE

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Unstable Angina and AMI

The development of AMI is the result of a sequence of events that takes place over many years. The process begins with atherosclerosis, which is characterized by the narrowing of coronary arteries by formation of lipid-filled plaques. Plaques that have a thick fibrous cap overlying the lipid layer are stable and not at risk for rupturing (Fig. 14.8A). Processes such as acute inflammation and monocyte infiltration will erode the cap, making it vulnerable to rupture by shear stresses that exist within arterial vessels (Fig. 14.8B). Rupture of this plaque leads to the syndrome of unstable angina, the disease that immediately precedes an AMI. The exposure of plaque to circulating blood leads

to the aggregation of platelets and the activation of thrombin, which converts fibrinogen to an insoluble fibrin clot. Subocclusive clots (not completely blocking the artery) produce unstable angina. A classification scheme for unstable angina was established by Braunwald (11). A totally occluded clot produces AMI. Both conditions, collectively known as acute coronary syndromes, can produce chest pain at rest, lead to specific changes in the electrocardiogram, and a mild or severe release of enzymes and proteins into blood.

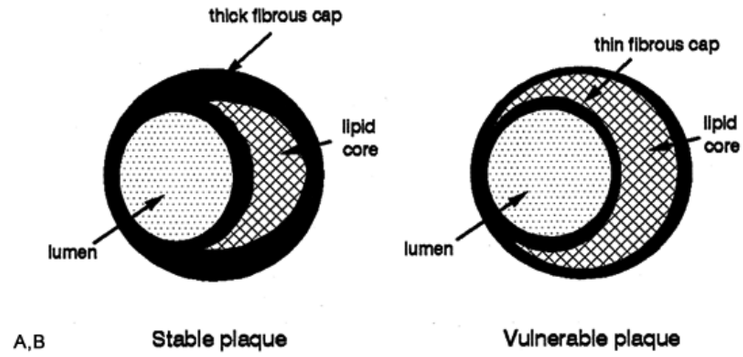


FIGURE 14.8. A: Stable coronary artery plaque not vulnerable for rupturing. **B:** Unstable plaque that is vulnerable for rupturing by shear stresses.

The most dramatic increases in biochemical markers occur in patients with an AMI. The immediate consequence is lack of oxygen delivery to the myocardium and tissue anoxia. This leads to an energy deficit and a shift toward anaerobic metabolism. Electrolytes are released into the interstitial space as the result of adenosine triphosphate (ATP)–dependent ion pump failure. If coronary circulation is restored before the onset of irreversible damage, either spontaneously or by therapeutic intervention, jeopardized myocardial tissue may be salvaged and substantial irreversible damage is avoided. Prolonged deprivation of blood flow (more than 4 hours) will cause irreversible myocardial damage, which leads to leakage of macromolecules such as enzymes and proteins. Figure 14.9 shows the temporal relationship between the release of these materials as a function of disease progression. The appearance of proteins and enzymes in the serum is prolonged because this clearance is dependent on slow drainage from the cardiac lymphatic circulation (Fig. 14.10). The presence of cardiac-specific enzymes and isoenzymes provides the basis for the biochemical diagnosis of AMI (12).

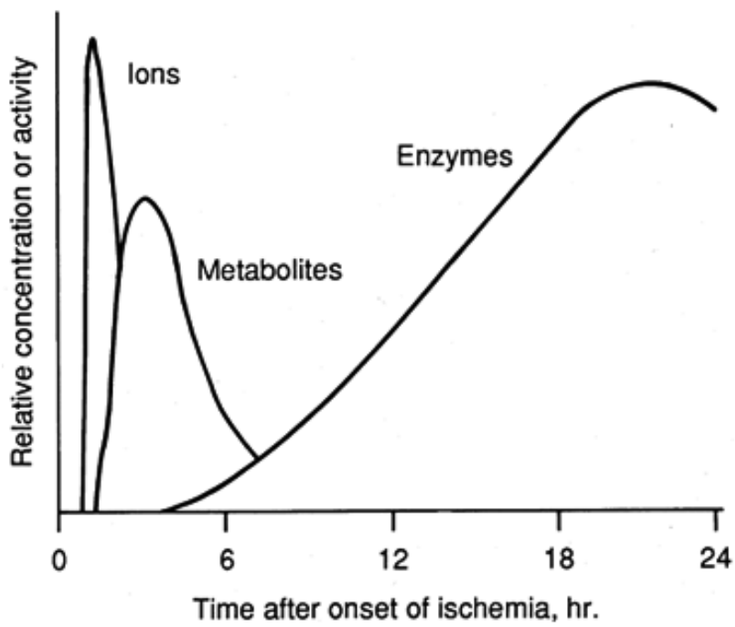


FIGURE 14.9. Release of ions, metabolites, and proteins after acute myocardial infarction. (Adapted from Hearse DJ. Cellular damage during myocardial ischaemia: metabolic changes leading to enzyme leakage. In: Hearse DJ. *Enzymes in cardiology. Diagnosis and research*. New York: John Wiley and Sons, 1979.)

Congestive Heart Failure

Heart failure is a very common disease, with more than 400,000 new cases each year in the United States. Heart failure results when the myocardium is unable to meet the blood supply demands of the tissues and organs. Failure can be classified as either acute or chronic. Primary acute failure can be caused by massive AMI or valve rupture. Chronic congestive heart failure (CHF) is seen in cardiomyopathies and valve disease. Factors that can precipitate CHF include infections, hypertension, endocarditis and myocarditis, arrhythmias, and pulmonary embolism. Physical exertion, emotional stress, and increased blood volume secondary to excessive sodium intake or discontinuation of diuretics

can also cause CHF. The New York Heart Association (NYHA) classified heart failure according to four classes based on functional capacity and on an objective assessment. Although the laboratory has played only a minor supportive role in CHF, new markers are being developed, such as brain natriuretic peptide (BNP), that might have a role in the management and monitoring of these patients. Preliminary studies have shown that there is an incremental increase in mean BNP concentrations with each progressive NHYA stage. It may be possible that BNP could be used to monitor the effects of drug therapy, such as with angiotensin-converting enzyme (ACE) inhibitors.

Skeletal Muscle Disorders

Evidence of high CK activity in the serum is useful for the diagnosis and evaluation of both acute and chronic skeletal muscle diseases. Acute release of muscle enzymes occurs in surgery, trauma, crush injuries, excessive muscular contractions (as in long-distance running or uncontrolled convulsions), hyperthermia, viral myositis, and exposure to toxins. Chronic progressive muscular fiber dystrophies include Duchenne, Becker, limb-girdle, and myotonic dystrophy. These genetic disorders generally affect young children. Inflammatory disorders of skeletal muscle include dermatomyositis and polymyositis, which are also associated with progressive muscle weakness. Neurogenic disorders of the lower motor neurons include amyotrophic lateral sclerosis and spinal muscular atrophies, whereas myasthenia gravis is caused by a disorder at the neuromuscular junction. The measurement of enzymes and isoenzymes is useful in the diagnosis of muscle disorders and, in the case of Duchenne muscular dystrophy, for detection of the carrier state in women (13).

Enzymes and Proteins Useful in Myocardial Disease

CK (EC 2.7.3.2)

The mitochondrial isoenzyme of CK (ATP:creatine *N*-phosphotransferase) functions in muscle cells to catalyze the transfer of a high-energy phosphate bond from ATP to creatine to form creatine phosphate. During active muscle contractions, cytoplasmic CK catalyzes the reverse reaction, thereby providing myocytes an immediate source of ATP. Cytoplasmic CK consists of dimeric combinations of two subunits, M and B. This enables three possible isoenzymes, MM, MB, and BB. In addition, posttranslational modifications of isoenzymes will produce three MM and two MB isoforms. Each of these forms has a combined molecular weight of approximately 80 kDa.

Measurement of Total CK Activity

Nearly all analytical methods for total CK enzyme activity make use of the formulation involving creatine phosphate and adenosine diphosphate (ADP). The product of the reverse reaction, ATP, is coupled with hexokinase and glucose-6-phosphate dehydrogenase (G6PD) to form NADH:

$$\text{creatine phosphate} + \text{ADP} \xrightarrow{\text{CK}} \text{creatine} + \text{ATP}$$

$$\text{ATP} + \text{glucose} \xrightarrow{\text{hexokinase}} \text{glucose-6-phosphate} + \text{ADP}$$

$$\text{glucose-6-phosphate} + \text{NAD}^+ \xrightarrow{\text{G6PD}} 6\text{-phosphogluconate} + \text{NADH} + \text{H}^+$$

The reaction is monitored by the increase in absorbance of NADH measured at 340 nm. A lag phase of 90 to 120 seconds is required for CK before a linear reaction rate is initiated. The pH of the reaction is maintained at 7.0 with imidazole buffer. *N*-acetyl-L-cysteine is added as an activator to maintain a supply of reduced sulfhydryl groups necessary for the complete activation of CK. The presence of adenylate kinase (AK) in the serum will produce a positive interference in this assay. AK catalyzes the production of ATP from ADP alone, which will falsely increase the apparent CK activity:

$$2\,\text{ADP} \xrightarrow{\text{AK}} \text{ATP} + \text{AMP}$$

Adenylate kinase originates from skeletal muscle, red cells, liver, and kidney. Most commercial assay systems use adenosine monophosphate and diadenosine pentaphosphate to inhibit AK activity. Alternatively, a blank can be used by measuring AK activity

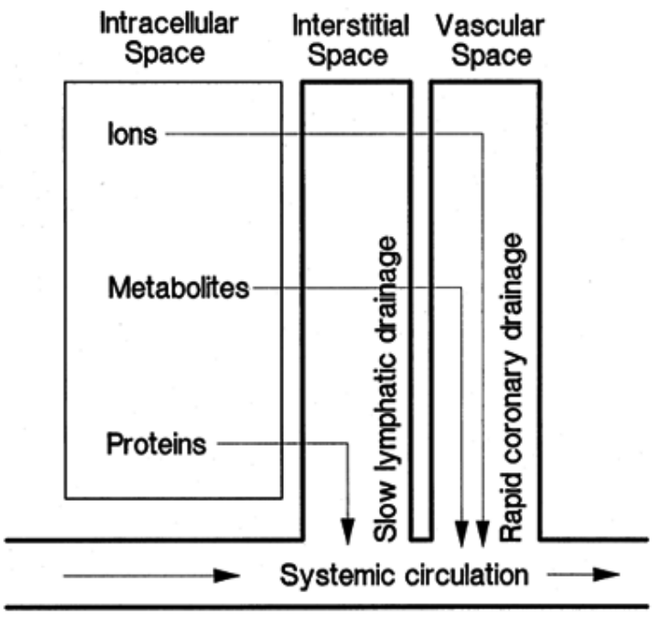


FIGURE 14.10. Route of clearance for ions, metabolites, and enzymes after acute myocardial necrosis. (Adapted from Hearse DJ. Cellular damage during myocardial ischaemia: metabolic changes leading to enzyme leakage. In: Hearse DJ. *Enzymes in cardiology. Diagnosis and research*. New York: John Wiley and Sons, 1979.)

with creatine phosphate omitted. The reference range at 37°C for total CK is 38 to 174 U/L for adult men and 96 to 140 U/L for adult women.

CK Isoenzymes, Isoforms, and Variants

Isoenzyme Tissue Distribution. The clinical utility of CK isoenzyme measurements stems from the fact that skeletal and myocardial muscle tissue have different distributions of CK isoenzymes. As shown in Fig. 14.11, CK-MM is the predominant isoenzyme in both tissues (14), but in skeletal muscle, only trace levels of CK-MB are present. In contrast, the myocardium contains a much higher percentage of CK-MB. CK-BB is found throughout the brain and smooth muscle.

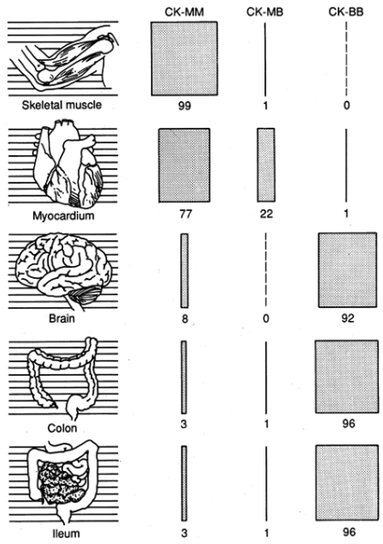


FIGURE 14.11. Tissue distribution of creatine kinase isoenzymes. (Adapted from Lott JA, Wolf PL. *Clinical enzymology. A case-oriented approach*. New York: Year Book, 1986.)

CK Isoforms. Examination of human muscle tissue isoenzymes reveals the presence of single pure gene products for CK-MM and MB. However, when released from the tissues these enzymes are slowly converted to multiple CK isoforms (MM_1, MM_2, and MB_1) through the successive cleavage of the C-terminal amino acids by serum carboxypeptidase. The isoform pattern of normal patients primarily consists of converted CK isoforms, with only trace quantities of tissue forms. High activities of the unmodified isoforms (MM_3 and MB_2) indicate the presence of an acute process.

CK Variants. In addition to the CK isoenzymes and isoforms that are present in normal sera, two atypical isoenzymes have been identified in approximately 2% of hospitalized patients. Macro CK type 1 has been identified to be CK-BB linked with an immunoglobulin (Ig). The IgG form migrates between CK-MM and MB by electrophoresis, whereas the IgA form migrates with MB. Macro CK type 2 is a polymeric aggregate of mitochondrial CK. Macro CK type 1 had no clinical significance, whereas type 2 is seen in metastatic cancer. Assays such as immunoinhibition and electrophoresis (especially with the type 1 IgA form) will produce false-positive results for CK-MB in the presence of macro CK.

Analytical Measurements. There are several analytical methods currently in use or CK isoenzyme analysis, including ion-exchange chromatography, electrophoresis, immunoinhibition, radioimmunoassay, and enzyme immunoassay (15). Electrophoresis for isoenzymes and isoforms is still used, but its popularity is diminishing. Isoenzymes are resolved on agarose gel after electrophoresis at 100 V for 20 minutes. Bands are observed fluorometrically after incubating the gels with CK reagents. The CK activity attributed to each band can be determined by scanning densitometry. Figure 14.12 illustrates the electrophoretic migration of CK isoenzymes, isoforms, and all known CK variants. Most normal samples will exhibit a single band owing to CK-MM. Patients with AMI will have high concentrations of CK-MM and MB. Increasing the electrophoresis time or the applied voltage permits the resolution and analysis of CK isoforms. Figure 14.13 illustrates the major CK-MM and MB isoforms. Electrophoresis is less sensitive than immunoassay methods, although it is the only method that permits visualization of all CK isoenzymes.

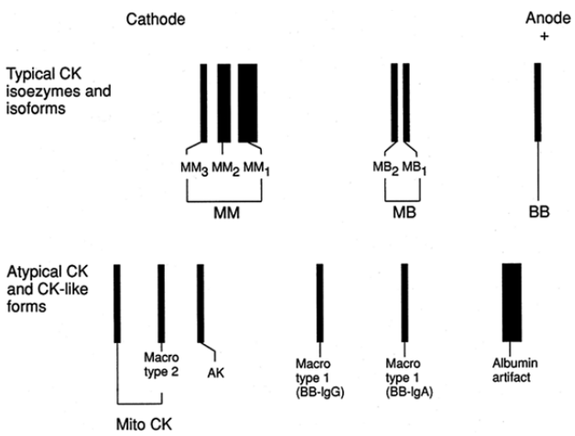


FIGURE 14.12. Electrophoretic migration of typical and atypical isoenzymes and isoforms of creatine kinase.

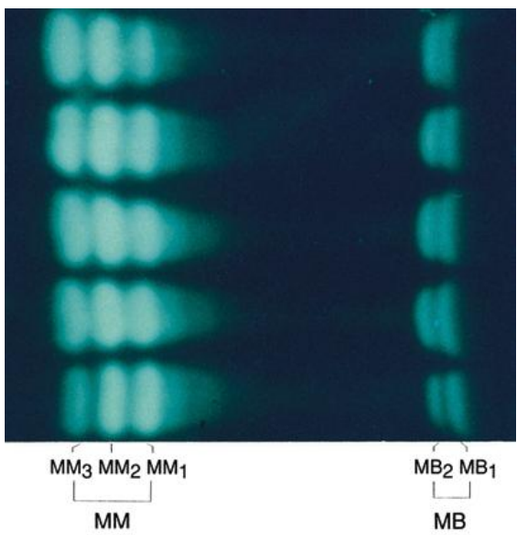


FIGURE 14.13. Creatine kinase isoforms as measured by agarose electrophoresis. (CK-isoforms, courtesy of Helena Labs., Beaumont, TX.)

In the immunoinhibition method, serum is analyzed for CK after the addition of specific antibodies. The technique, illustrated in Fig. 14.14, involves inhibiting CK-M subunit activity

with anti-M antibodies and measuring the residual B-subunit activity. If CK-BB and atypical CK forms are absent, the MB content of the sample can be determined by multiplying the B-subunit activity by 2. However, if CK-BB or macro CK is present, a usually high percentage of CK-MB/total CK will occur, and the analysis should be confirmed by a more specific CK-MB assay. The immunoinhibition technique is very useful for screening because it is inexpensive and can be linked to automated chemistry analyzers.

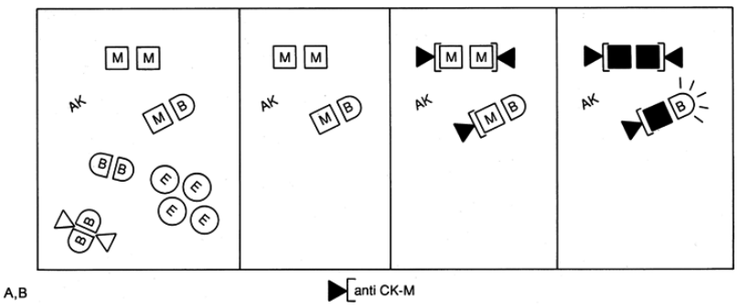


FIGURE 14.14. The immunoinhibition technique for creatine kinase (CK) B. **A:** Normal and atypical isoenzymes of CK and adenylate kinase. **B:** Only CK-MM, MB, and adenylate kinase are assumed to be present. **C:** Addition of anti-CK-M antibodies. **D:** Measurement of residual B-subunit activity after M-subunit inactivation (solid symbols). Inhibitors to adenylate kinase minimize the contribution of this interferent. The CK-MB activity is calculated by multiplying the residual activity by 2. (From Wu AHB. *Clinical chemistry*. Bethesda, MD: Health Education Resources, 1991, with permission.)

The most commonly used method for CK-MB are nonisotopic immunoassays, which measure enzyme concentration instead of activity. All commercial immunoassays make use of a two-site sandwich technique. The first antibody recognizes a particular determinant on the CK-MB molecule and is linked to a solid phase such as a tube, bead, magnetic particle, or fibrous surface. Monoclonal anti-bodies to CK-MB have been used to capture MB from the sample in this first step. The second antibody recognizes a different MB determinant and is conjugated to a label, such as an enzyme, fluorophore, or chemoluminescent tag. The concentration of CK-MB from the sample is determined from a standard curve. Figure 14.15 illustrates the Dade Stratus II method as one example. Immunoassays are very sensitive and rapid. Improvements in analytical precision and clinical performance have been reported over enzyme-activity assays, particularly when total CK levels are low.

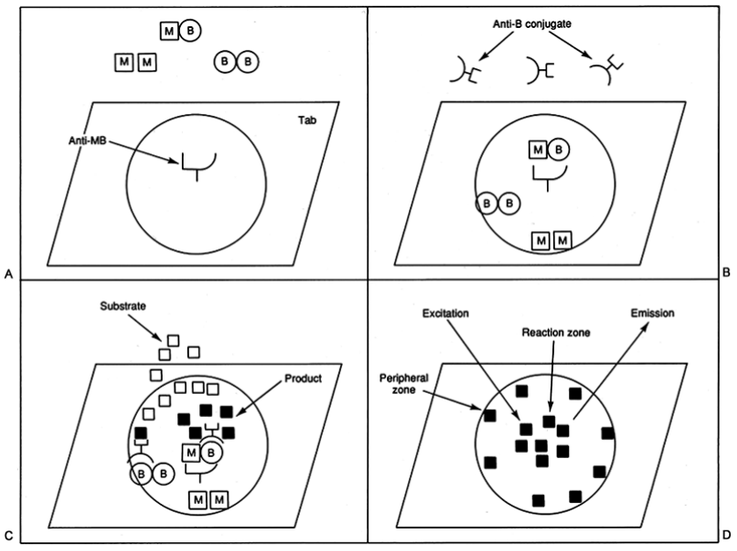


FIGURE 14.15. Radial partition immunoassay for creatine kinase (CK)-MB. **A:** A serum sample with CK isoenzymes is added to the center of a tab containing anti-CK-MB antibodies linked to glass fiber paper. The MB isoenzyme becomes immobilized. **B:** Anti-CK-B antibodies linked to alkaline phosphatase are added. These bind both to the immobilized CK-MB isoenzymes and to BB. **C:** Substrate to ALP is added to the center of the tab to produce a fluorescent product. CK-MM and BB are not bound to anti-MB and are washed outward from the reaction zone. **D:** The fluorescence intensity is proportional to (MB) measured at the center of the front surface of the tab. The fluorescence of conjugate (of CK-BB) migrating outward from the periphery of the reaction zone is not measured.

Reference Values

Reference values for CK-MB are dependent on the analytical method used and are often expressed as a fraction of total CK activity. Reference values from a healthy population should not be used as decision limits for the diagnosis of AMI, however, because there would not be adequate discrimination between the AMI and non-AMI patient (e.g., unstable angina). Cutoff values for electrophoresis and immunoinhibition are usually set around 5% of total CK or 10 U/L. For immunoassays, decision limits are expressed in mass quantities and are approximately 5 to 10 ng/mL. CK-BB is normally absent in adult serum. CK-MB and CK-BB levels are higher in children, but there are no established reference ranges, as CK isoenzyme analysis is not usually performed for this population.

There are two ways by which CK-MB results can be expressed: absolute activity (U/L) or mass concentration (ng/mL) and a percentage of total CK (relative index). The relative index is defined as:

$$\text{Relative index: } \frac{\text{CK-MB (activity or mass)}}{\text{total CK (activity)}} \times 100$$

As a rule, absolute limits are most useful when the total CK is low (e.g., less than 1,000 U/L). Because CK-MB is found in skeletal muscles, however, CK-MB levels will exceed absolute reference limits in patients with severe skeletal muscle injury or disease. In this case, a relative limit (CK-MB expressed as a percentage of total CK) is more useful when the total exceeds 1,000 U/L.

Lactose Dehydrogenase (LDH) (EC 1.1.1.27)

LDH (L-lactate:NAD^+ oxidoreductase) is a ubiquitous enzyme found in the cytoplasm of nearly all cells; it catalyzes an important step in glycolysis. LDH has a molecular weight of approximately 134 kDa and is a tetramer of H and M subunits. There are five possible isoenzyme results from these combinations. Measurement of LDH is useful for several different diseases involving the blood, liver, kidneys, skeletal muscle, and myocardium.

Measurement of Total LDH Activity

LDH catalyzes the reversible oxidation of L-lactate to pyruvate, with simultaneous reduction of NAD^+:

$$\text{lactate} + NAD^+ \xrightarrow{\text{LD}} \text{pyruvate} + NADH$$

Both the forward (lactate-to-pyruvate) and reverse (pyruvate-to-lactate) reactions are currently in use in clinical laboratories. In the United States, the forward assay is most widely used. The advantage of this method is that the reaction rate is slower and is held more constant than for the reverse assay. The reaction is monitored by measuring the absorbance of NADH at 340 nm. The optimum pH range for this reaction is 8.8 to 9.8. The reaction

rate for the reverse assay is higher than for the forward, and the precision is therefore somewhat improved. The pH range for the reverse reaction is 7.4 to 7.8. The reference interval for the forward reaction is approximately 80 to 200 U/L at 37°C, and approximately 200 to 400 U/L for the reverse reaction.

LDH Isoenzymes

Tissue Distribution. The LDH isoenzyme distribution of various tissues is shown in Fig. 14.16 (14). Requests for LDH isoenzymes are most frequently made in conjunction with the diagnosis of AMI, although measuring LDH isoenzymes can be used to determine other sources of tissue injury. Because the heart is richest in LDH_1, some analytical methods have focused on this isoenzyme alone. The decision limit most widely used is that LDH_1 activity in excess of 40% of total LDH activity supports the diagnosis of AMI. LDH isoenzyme fractionation is rarely of any value in the diagnosis of other conditions (e.g., liver disease) because there are other more specific markers available and there is overlap in the tissue distribution of LDH isoenzymes from other organs.

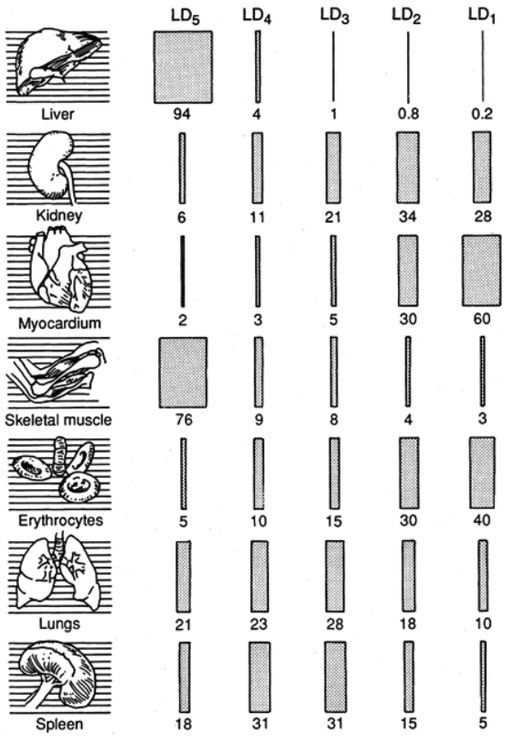


FIGURE 14.16. Tissue distribution of lactate dehydrogenase isoenzymes. (From Lott JA, Wolf PL. *Clinical enzymology. A case-oriented approach*. New York: Year Book, 1986, with permission.)

Analytical Methods. Given the decreased utilization of LDH isoenzymes, only a brief description of methods is given. The reader is referred to the first edition of *Clinical Laboratory Medicine* for details. The only routinely used procedure that allows analysis of all LDH isoenzymes is electrophoresis. Bands are visualized by densitometrically scanning the fluorescence of the NADH product or reacting NADH with a tetrazolium salt to form a colored product that can be measured at 600 nm. The serum of some patients will exhibit the presence of other LDH forms such as macro-LDH (immunoglobulin bound), but the significance of these other forms is unknown. The LDH_1 isoenzyme can also be measured directly by the Roche immunoprecipitation method. This assay uses a goat anti-LDH-M antibody LDH isoenzymes LDH_2 through LDH_5, and a second precipitating (donkey anti-goat) antibody, which recognizes and binds the first antibody. After centrifugation, only the LDH_1 isoenzyme remains, as it does not contain any M subunits. The extracted sample is then measured for residual LDH activity. A variation of this method is the use of selective chemical inhibition. The Roche LDH_1 assay makes use of guanidine thiocyanate, which acts as a competitive inhibitor for LDH_1 and a noncompetitive inhibitor for LDH_5. The degree of inhibition is greatest for LDH_5, intermediate for LDH_4, LDH_3, and LDH_2, and least for LDH_1. The reagent is formulated to optimize recovery of LDH_1 activity. The chemical inhibition method gives results that are equivalent to immunoprecipitation and has the advantage that it does not require a centrifugation step to isolate LDH_1; therefore, the assay can be automated to high-volume clinical chemistry analyzers.

Myoglobin

Myoglobin is a low-molecular weight protein (less than 18 kDa) found in all skeletal muscle and myocardial tissues. Myoglobin is an oxygen-binding protein, serves as a reserve for oxygen, and facilitates the movement of oxygen within muscle cells. Myoglobin released from the heart is indistinguishable from that released from skeletal muscle tissues. It has been known since the mid-1970s that myoglobin was an early marker for AMI (16). However, because early assays for myoglobin were based on a manual radioimmunoassay, it was not practical for a laboratory to produce stat myoglobin results for real-time management of cardiac patients. The long assay time and expense for the testing negated the advantage of early diagnosis that myoglobin had over CK-MB. In the late 1980s, nonisotopic nephelometric and turbidimetric assays were developed. This renewed the clinical interest for myoglobin as a cardiac marker. The current myoglobin assays are based on immunoassays, are available on a wide variety of automated immunoassay analyzers, and are practical for stat testing. Reference limits for serum myoglobin vary but are generally less than 100 ng/mL. Although myoglobin is filtered by the glomerulus, it is reabsorbed and catabolized by the tubules and there is little myoglobin that appears in the urine of normal individuals. Myoglobin is also increased in patients with skeletal muscle disease or injury, and in patients with acute or chronic renal failure.

Troponin

Troponin is a regulatory protein complex located on the thin filament of striated muscles and consists of three isotypes. Troponin-T

(cTnT) has a molecular weight of 38 kDa and binds the troponin complex to tropomyosin. Troponin-I (cTnI) has a molecular weight of 24 kDa and functions to inhibit actomyosin ATPase. Troponin-C has a molecular weight of 18 kDa and regulates cTnI activity by binding calcium. cTnT and cTnI are better diagnostic markers for AMI than existing enzymes such as CK-MB because cardiac isotypes are distinctly different from their corresponding skeletal muscle isotypes. Monoclonal antibodies have been developed to cardiac troponin that do not cross-react with skeletal muscle forms. There are no specificity advantages for measuring cardiac troponin C (cTnC) as the skeletal muscle forms have homology to the cardiac forms.

Point-of-Care Testing Devices

Emergency departments (ED) in the United States and worldwide are changing their patient triaging practices toward a rapid rule out of chest pain patients who do not have acute coronary syndromes. This aggressive approach involves testing blood on a more frequent basis than before, e.g., every 3 to 4 hours for the first 12 hours instead of at 8- or 12-hour intervals. After AMI has been ruled out, an exercise stress test can be performed to rule out other acute coronary diseases. Negative results for these tests enables ED physicians to discharge patients earlier than before. To meet these changing clinical demands, laboratories must perform testing for cardiac markers on a stat or as-needed basis. The National Academy of Clinical Biochemistry (NACB) has recommended a reporting turnaround time for cardiac markers of 1 hour or less (17). Many laboratories are unable to comply with this objective because of delays in receiving specimens from the ED. Hospitals that do not have an ED laboratory or pneumatic tube transport system to directly connect the ED to the laboratory must hand-deliver specimens to the laboratory. To address these problems, qualitative and quantitative point-of-care devices are now available for myoglobin, CK-MB, and cardiac cTnT and cTnI (18). These assays use whole blood and can produce results within 10 to 20 minutes. They are designed to be used at ED bedside and obviates the need for delivering the sample to the laboratory. Although the benefits for rapid turnaround time testing have been documented for EDs who have an aggressive triaging protocol, there have been no studies to determine whether bedside testing improves clinical outcomes compared with stat testing performed in a central laboratory. Point-of-care testing for cardiac markers is more expensive, and there are regulatory requirements for proficiency and quality control for the nonlaboratory personnel who perform the testing. Nevertheless, point-of-care testing is very attractive for hospitals who need rapid triage for their chest pain patients.

Use of Cardiac Markers in the Diagnosis of AMI

The diagnosis of AMI is defined by the World Health Organization (WHO), based on clinical signs and symptoms (e.g., chest pain), specific changes in electrocardiographic recordings, and the presence of elevated serum activities and concentrations of enzymes such as total CK, CK-MB, LDH, and LDH isoenzymes (19). With the development of other tests such as myoglobin, and the cardiac troponins, the WHO definition of AMI should be expanded to include enzymes and proteins because these markers are not enzymes.

The optimal use of enzyme markers for diagnosis of AMI requires careful selection of decision limits and correct timing of blood samples. Because false-positive results can occur, the best information is obtained when serial collections are analyzed to construct enzyme activity or cardiac marker concentration versus time curves.

Early Diagnosis: Myoglobin and CK Isoforms

An important factor that affects the efficiency of biochemical markers for AMI diagnosis is the sampling time from the onset of chest pain. Figure 14.17 shows the activity-time curves for cardiac enzymes and proteins after AMI. As shown, myoglobin

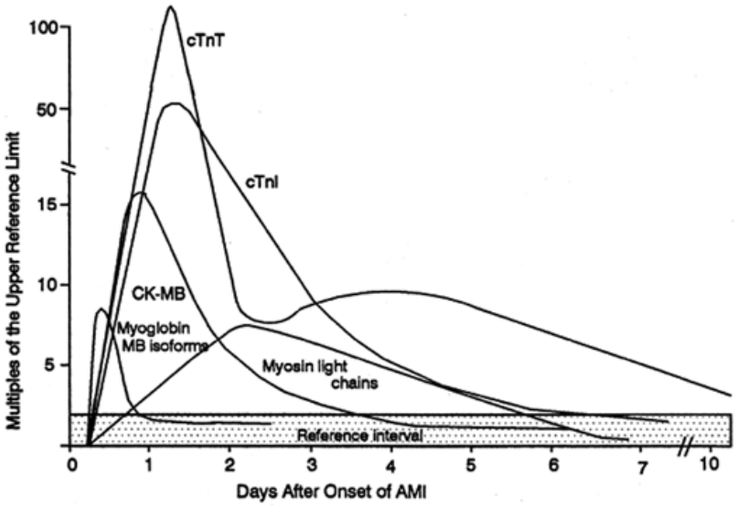


FIGURE 14.17. Activity versus time curves in serum for biochemical markers of acute myocardial infarction.

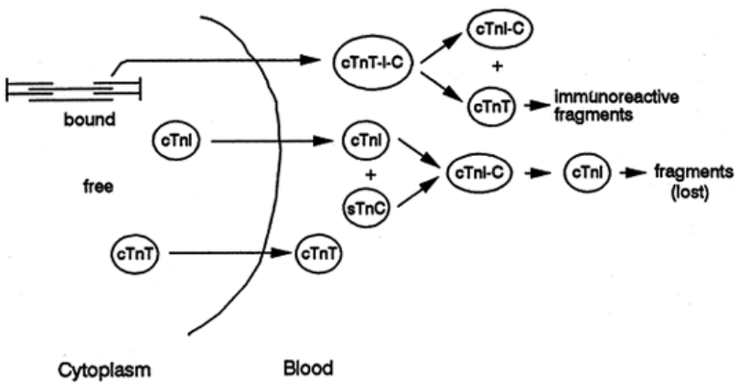


FIGURE 14.18. Release of cardiac troponin from damaged myocytes. (Modified from Wu AH. Biochemical differences between cTnT and cTnI and their significance for diagnosis of acute coronary syndromes. *Eur Heart J* 1998;19(Suppl N):25--29.

and CK-MB isoforms have the earliest diagnostic windows of the markers shown. Myoglobin is increased early after AMI because of its small size. CK isoforms appear early because serum levels normally contain only trace levels of CK isoforms, and a low cutoff can be used to detect the very earliest release of tissue isoforms after myocardial necrosis. Estimates for clinical sensitivity at various time intervals after chest pain and clinical specificity are shown in Table 14.3. Early diagnosis is important in the management of AMI, particularly if thrombolytic therapy is to be used. The clinical specificity of myoglobin is not as high as for CK-MB isoforms. Utilization of MB isoforms for routine use, however, warrants improvements in the methodology.

TABLE 14.3. ESTIMATES OF CLINICAL SENSITIVITY AND SPECIFICITY OF DiAGNOSTIC TESTS FOR ACUTE MYOCARDIAL INFARCTION

	Sensitivity				
Marker	**2–8 h**	**8–24 h**	**24–72 h**	**>72 h**	**Specificity**
Myoglobin	95	75	0	0	70
CK-MB isoforms	90	60	0	0	90
CK-MB	60	95	98	50	95
LDH_1	40	85	95	90	85
Troponin	75	95	98	98	90

CK-MB, creatine kinase-MB; LDH_1, lactate dehydrogenase 1.

Definitive Diagnosis of AMI: CK-MB and Troponin

The CK-MB isoenzyme is the most widely used enzyme for the diagnosis of AMI. Levels begin to rise 6 to 10 hours after onset, peak at 18 to 24 hours, and return to normal within 3 days (Fig. 14.17). The sensitivity of CK-MB is approximately 95% to 98% (Table 14.3). The specificity of CK-MB when sampled at the optimum time after the onset of AMI ranges from 92% to 99%. CK-MB is more specific than total CK because the total can increase in a variety of nonmyocardial diseases. False-positive results do occur when myocardial tissue is damaged by other means, such as trauma to the heart or open heart surgery. CK-MB of skeletal muscle origin can also be observed in serum with dermatomyositis, muscular dystrophies, and training for long-distance running. In each of these conditions, the skeletal muscle may contain a higher percentage of MB than normal because of increased turnover and alterations in isoenzyme expression during regeneration.

The cardiac troponins (T or I) are the most efficient markers available today for diagnosis of AMI. Because troponin has a higher tissue distribution in myocytes than CK-MB or myoglobin, the degree of increase after AMI (relative to the normal range) is much higher than for the other cytosolic markers (Fig. 14.17). The early release of cTnT and cTnI are owing to release of the free cytosolic pool of these proteins, estimated to be approximately 6% to 8% and 2.8% to 4.1%, respectively (Fig. 14.18). Because of the higher cytosolic content, cTnT may be increased slightly before cTnI after AMI. In addition, cTnT exhibits a biphasic pattern because of the significant free cytosolic cTnT component. (The second peak is owing to release of the structural elements of cTnT, as discussed in the next section.) Cardiac cTnI exhibits a largely monophasic pattern because of the lower free cTnI release and because some of the cTnI putatively binds to circulating skeletal muscle troponin C to form a binary complex (Fig. 14.18).

The specificity of troponin is also increased over CK-MB or myoglobin because false-positive results are not observed in patients with skeletal muscle disease or injury. Estimates of the clinical sensitivity and specificity for troponin are listed in Table 14.3. Because of the improved performance in clinical studies, cardiac troponin has been recommended by the NACB as the preferred marker for the definitive diagnosis of AMI, replacing CK-MB (17).

Late Diagnosis of AMI: LDH and Cardiac cTnT and cTnI

The diagnosis of AMI in patients who present to the ED with a 3- to 4-day history of chest pain may be difficult with myoglobin or CK-MB because these markers might have already returned to normal. In this situation, use of a cardiac marker that remains increased for a week or more after AMI onset is useful for making a retrospective diagnosis. Isoenzymes of LDH have classically been used for this purpose. The appearance and clearance of LDH is delayed because this enzyme is considerably larger than either myoglobin or CK.

With the development of structural markers such as cTnT and cTnI, the NACB has recommended discontinuance of

LDH isoenzymes (17). Cardiac cTnT and cTnI have a prolonged release pattern after myocardial injury owing to the slow degradation of muscle fibers by proteolytic and lysosomal enzymes. As shown in Fig. 14.17, cTnI remains increased for approximately 7 days, whereas cTnT is abnormal for as long as 10 days. The differences in the clearance rate between cTnT and cTnI is related to how the troponin complex from myocytes is released from tissues and degraded in blood and the recognition of the various troponin forms by commercial assays. Figure 14.18 shows that structurally bound troponin is released into blood as a ternary complex of cTnT-I-C, which subsequently degrades into a binary complex of cTnI-C and free cTnT. The binary complex further degrades into free subunits. Gel filtration studies have shown that the commercial assay for cTnT recognizes complexed, free, and degradative fragments of free cTnT (20). In contrast, there is not a large, free cTnI component in blood, and immunoreactive fragments of cTnI do not appear to be detectable by existing assays. A possible reason is that cTnI binds to proteins and other materials that may affect its recognition by antibodies used in commercial cTnI assays. As a result, the apparent clearance of cTnI after AMI is faster than for cTnT.

Experimental Markers for Early AMI

Cardiac troponin is gradually receiving acceptance as the standard marker for the retrospective diagnosis of AMI. Given the need by ED physicians to triage chest pain patients earlier and earlier and the sensitivity and specificity limitations of myoglobin and CK-isoforms, there is ongoing research into new early markers of AMI and acute coronary syndromes. The current research efforts are focused on using markers to identify the early pathophysiologic events that are known to occur in an evolving AMI. Table 14.4 summarizes the biochemical attributes of investigational early AMI markers. Laboratory tests for C-reactive protein and amyloid protein A can be used to indicate the presence of acute inflammation. Markers of thrombosis include soluble fibrin monomers and thrombus precursor protein. Soluble and membrane-bound P-selectins are markers of platelet aggregation, whereas glycogen phosphorylase BB is thought to be released during periods of reversible ischemia. Two markers of cellular necrosis have been proposed to augment the early diagnostic utility of myoglobin. Carbonic anhydrase III is released in patients with skeletal muscle injury, thus a normal result with a high myoglobin suggests acute cardiac injury. Fatty acid binding protein has a higher myocardial tissue content than myoglobin. The use of a ratio between the binding protein and myoglobin can be helpful in differentiating between myocardial and skeletal muscle injury, which would have a different ratio between these two markers.

TABLE 14.4. SUMMARY OF EARLY BIOCHEMICAL MARKERS FOR ACUTE CORONARY SYNDROMES

Marker	Biochemical Function	Size (kDa)	Clinical Utility
Markers of inflammation			
C-Reactive protein	Acute phase reactant	~120	Nonspecific markers of inflammation
Amyloid protein A	Acute phase reactant	12.5 (mono), 220–235 (poly)	
Coagulation factors and proteins			
Soluble fibrin monomers	Soluble protein precursor	??	Early detection of thrombus form
Thrombus precursor protein	Insoluble fibrin	??	
Platelet function			
Soluble P-selectin	Platelet activation	140	Platelet aggregation
Ischemic marker			
Glycogen phosphorylase BB	Enzyme of glycogenolysis	~200	Reversible injury
Biochemical markers			
Carbonic anhydrase III	Converts HCO_3^- to H_2CO_3	28	Skeletal muscle protein (used with myoglobin)
Fatty acid binding protein	Cytosolic fatty acid carrier	15	Nonspecific early acute myocardial infarction marker

From the National Academy of Clinical Biochemistry. *Standards of laboratory practice monograph: recommendations for the use of cardiac markers in coronary artery diseases*, 1999, with permission.

Early clinical studies showed that none of these markers by themselves has the specificity needed to be used as an early AMI marker. C-reactive protein is increased in any conditions associated with infection or inflammation. Hemostatic markers will be positive in patients with other thrombotic diseases such as stroke, pulmonary embolism, and deep vein thrombosis. Markers of reversible ischemia may be able to rule in acute coronary syndromes but cannot differentiate between AMI and unstable angina, as the reduction in coronary artery blood flow occurs in both diseases. Given these limitations, a multivariate approach may be necessary to utilize the incremental advantage that each marker might offer.

Risk Stratification of Patients with Unstable Angina

Total occlusion of a coronary artery after plaque rupture will produce large areas of myocardial necrosis and extensive release of cardiac markers. In unstable angina in which plaque rupture leads to a nonocclusive clot, minor myocardial injury can occur in subset of these patients, presumably owing to blockage of small vessels and collateral arteries. Highly sensitive cardiac markers such as cardiac cTnT or cTnI can detect the presence of minor myocardial injury, despite normal concentrations of other markers. This has been demonstrated in patients undergoing elective coronary angioplasty, in which the presence of side

branch occlusions and distal thrombi was associated with increased cTnT relative to preangioplasty values (21).

The significance of detecting minor myocardial injury in patients with unstable angina has been demonstrated by numerous investigators. Figure 14.19 shows that patients with a positive cTnT at presentation had a significantly higher short-term cardiac morbidity and mortality than patients with a normal cTnT. In a metaanalysis combining results of several outcome studies, patients with unstable angina and a positive cTnT or cTnI have an odds ratio of 5 to 1 for untoward events (22). It is possible that a slightly elevated troponin result identifies the presence of a microinfarct that might indicate a more advanced stage of the disease compared with a patient who has not yet exhibited evidence of irreversible damage. Unstable angina patients at high risk can be treated with low molecular weight heparin or a glycoprotein IIb/IIIa receptor inhibitor, each of which has been shown to reduce the incidence of future cardiac events in these patients (23). The use of cardiac troponin has the dual role of diagnosis of AMI and risk stratification in non-AMI patients who have unstable angina.

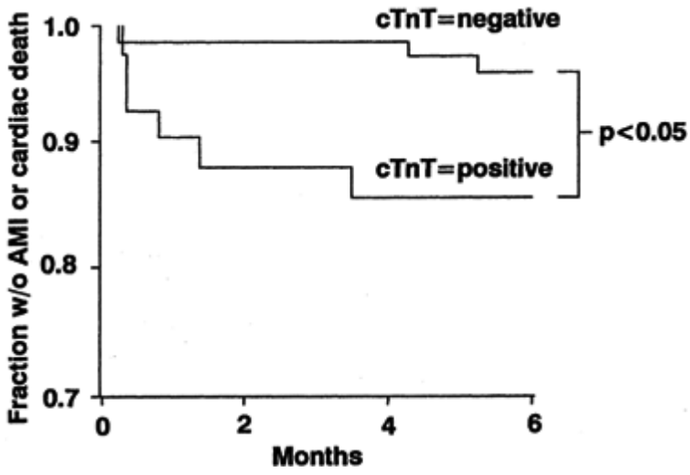


FIGURE 14.19. Kaplan-Meier survival curves for patients with unstable angina and cardiac troponin T. (Modified from Ravkilde J, Horder M, Gerhardt W, et al. Diagnostic performance and prognostic value of serum troponin T in suspected acute myocardial infarction. *Scand J Clin Lab Invest* 1993;53:677–685.)

Enzymes and Proteins in Skeletal Muscle Disease

The important enzyme markers of skeletal muscle necrosis include total CK, CK-MB, and, to a lesser extent, AST and aldolase. Elevations of total CK with a normal relative index for CK-MB indicate skeletal muscle disease, necrosis, or injury. When both total CK and CK-MB are elevated, the diagnosis of skeletal muscle injury may be difficult. Patients with muscular dystrophies can have a higher than normal proportion of CK-MB in their muscle tissue because of increased muscle turnover. In addition, individuals who run long distances can exhibit significant levels of noncardiac CK-MB in serum. Table 14.5 shows typical results of CK and CK-MB levels after a marathon race (24). In most cases, the clinical history is sufficient for differentiating acute skeletal muscle damage from acute myocardial injury. In cases of concomitant skeletal muscle injury and AMI, measurement of troponin allows differentiation between skeletal muscle injury and myocardial damage. This is because the specificity of cardiac troponin is very high, with no cross-reactivity of commercial assays to the presence of skeletal muscle troponin.

TABLE 14.5. CREATINE KINASE (CK) LEVELS BEFORE AND AFTER A MARATHON RACE

Prerace CK (U/L)	Postrace CK[a] (U/L)	Postrace MB[a] (U/L)	% MB/Total CK
202	1,293	85	6.6
322	1,654	108	6.5
232	2,261	138	6.1
340	1,768	85	4.8
147	683	80	11.7

[a] Values taken 24 hours after the race.
Data from ref. 24.

Extensive skeletal muscle injury, as characterized by extremely high levels of activity of total CK in serum and concentrations of myoglobin in serum and urine, is observed in patients with rhabdomyolysis. The causes of rhabdomyolysis include crush injuries, excessive muscular contractions, drugs and toxins, and immobilization accompanied by hypothermia. A serious complication of extensive skeletal muscle injury is rhabdomyolysis-induced acute renal failure. There are two theories as to the cause of renal failure: direct toxicity from the heme group of myoglobin and precipitation of myoglobin within renal tubules causing obstruction. Patients with rhabdomyolysis are treated with osmotic diuresis with mannitol to increase renal blood flow and alkalinization to stabilize the solubility of myoglobin. The complications of diuretic therapy include CHF caused by fluid overload. Not all patients with rhabdomyolysis develop renal failure.

Use of the myoglobin clearance test or the ratio of serum to urine myoglobin concentrations may be helpful in predicting which patients would progress to acute renal failure and who would benefit from interventional therapy. Figure 14.20 illustrates two patients who presented with similar histories. Both patients had very high concentrations of total CK and serum and urine myoglobin. In the first case (Fig. 14.20A), the urine myoglobin exceeded the serum myoglobin, indicating that the clearance was high, suggesting that the rate of renal removal exceeded the rate of production from necrotic skeletal muscles. Acute renal failure did not occur in this patient. In the second patient (Fig. 14.20B), the serum myoglobin was initially higher than the urine, indicating that the rate of clearance was inadequate, and acute renal failure occurred on the next day (as shown by the creatinine clearance result) (25).

Use of quantitative serum and urine myogloin tests should replace the qualitative urine myoglobin dipstick test. This latter assay is not as sensitive as an immunoassay for myoglobin and can produce false-positive results in the presence of hemoglobinuria. Furthermore, finding myoglobin in the urine alone is not predictive of acute renal failure. High urine concentrations, in fact, may indicate that the kidneys were able to excrete the high plasma myoglobin load.

PANCREATIC DISEASES

Enzymes of Pancreatic Origin

Amylase (EC 3.2.1.1)

Amylase (α-1,4-glucan,4-glucanohydrolase) catalyzes the hydrolysis of complex carbohydrates at the α-1,4 linkages of adjacent glucose residues. The amylases have a molecular weight of approximately 50 kDa and are small enough to pass through the glomerulus and are among the few serum enzymes that appear in urine. Amylase is found in the highest concentrations in the acinar cells of the pancreas, but other isoenzymes are found in the salivary gland, testis, and ovary. Each isoenzyme is encoded by a separate gene. Serum amylase consists primarily of the pancreatic and salivary isoenzymes. Macroamylase is a complex of immunoglobulin (usually IgG) with amylase (usually the salivary isoenzyme). Because of the high molecular weight of this complex, macroamylase is not cleared by the kidneys, and high amylase levels persist in the serum, often without any clinical disorder. Macroamylasemia is estimated to account for 2% to 5% of all patients with hyperamylasemia (26).

Analytical Measurements

Numerous methods have been developed for measuring serum amylase. Viscometric, nephelometric, and iodometric assays have been largely replaced with saccharogenic and chromogenic assays that are very precise and can be adapted to automated chemistry analyzers. Saccharogenic assays use small oligosaccharide substrates.

$$\text{maltopentaside} \xrightarrow{\alpha\text{-amylase}} \text{maltotrioside} + \text{maltose}$$

$$\text{maltotrioside} \xrightarrow{\alpha\text{-glucosidase}} \text{maltose}$$

$$\text{maltose} \xrightarrow{\alpha\text{-glucosidase}} \text{glucose}$$

The reaction can be monitored by following the production of glucose as conventionally measured, for example with glucose oxidase and an O_2-specific electrode, or hexokinase coupled with G6PD and NAD^+. These assays must correct for the presence of variable amounts of endogenous glucose in serum.

Chromogenic assays make use of synthetic substrates containing a chromogen that is liberated directly by amylase or after coupling with α- or ß-glucosidase. These assays eliminate the additional steps needed to measure glucose. A common chromogen is *p*-nitrophenol (PNP).

$$p\text{-nitrophenylmaltoheptaoside} \xrightarrow{\alpha\text{-amylase}} p\text{-NP-maltotetraoside} + \text{PNP-maltotrioside}$$

$$p\text{-NP-maltotrioside} + H_2O \xrightarrow{\alpha\text{-glucosidase}} \rightarrow p\text{-NP} + \text{glucose}$$

In this example, the reaction rate is monitored by measuring the absorbance of *p*-nitrophenol at 405 nm. The assay has an optimum pH of 7.0 and requires calcium and anions as cofactors. The reference interval for amylase is dependent on the type of assay and the substrate used. For the chromogenic assay, the reference interval is 20 to 160 U/L at 37°C.

Amylase Isoenzymes and Isoforms

Human serum contains two amylase isoenzymes originating from the salivary gland (S type) and pancreas (P type). Each isoenzyme exists as one of three isoforms. The pure gene product for the salivary and pancreatic isoforms are S_1 and P_2, respectively. These are converted to S_2 and S_3, and to P_3, respectively, by posttranslational deglycosidation, and deamidation of asparagine and glutamine residues. The P_1 isoform of amylase is a genetic isoenzyme variant. Isoenzyme measurements are used to differentiate the causes of hyperamylasemia, such as those of salivary and pancreatic origin, or the presence of macroamylase. Amylase isoenzymes can be measured by electrophoresis, ion-exchange chromatography, IEF, and selective inhibition and precipitation with lectins

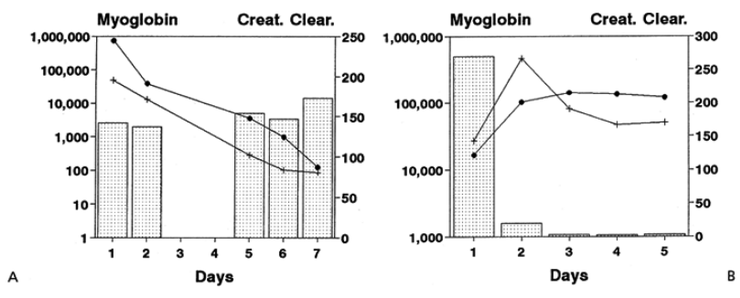


FIGURE 14.20. Use of the myoglobin clearance test to predict acute renal failure in male patients with rhabdomyolysis. Serum (+) and urine (•) myoglobin concentration (in nanograms per milliliter, left ordinate), and creatinine clearance (bars in milliliters per minute, right ordinate) plotted as a function of time after admission. Upper limit of normal in men for serum myoglobin, 0–62 ng/mL; urine, 6.8 ng/mL. Reference range for creatinine clearance, 97–137 mL/min for men. (Creatinine clearance was higher than the normal range owing to creatinine release from skeletal muscles.) **A:** Patient who did not develop acute renal failure. **B:** Patient who developed renal failure.

and monoclonal antibodies. Using electrophoresis, amylase isoenzyme bands can be made visible by use of a dyed-starch suspension (e.g., Phadebas tablets, Pharmacia Diagnostics, Piscataway, NJ, U.S.A.), which produces a color upon hydrolysis by amylase. Quantitative measurement of each band can be made scanning bands with a densitometer. In normal individuals, there are roughly equal amounts of S-type and P-type isoamylases in serum. Other methods are available to determine the total P-type amylase isoenzyme content of serum. Salivary amylase is approximately 90% inhibited by lectins from wheat germ. The estimation of pancreatic amylase is made by measuring the residual amylase activity after incubation with the lectin. Macroamylase of the S type is also inhibited by this technique. Monoclonal antibodies directed toward either salivary or pancreatic amylase can also be used to measure amylase isoenzymes in a variety of formats such as selective immunoinhibition, immunoprecipitation, immunoextraction (Fig. 14.21), or enzyme immunoassay. Immunoassays are very specific and do not cross-react with other isoamylase forms. However, they cannot be used on general chemistry analyzers, thus limiting their widespread application.

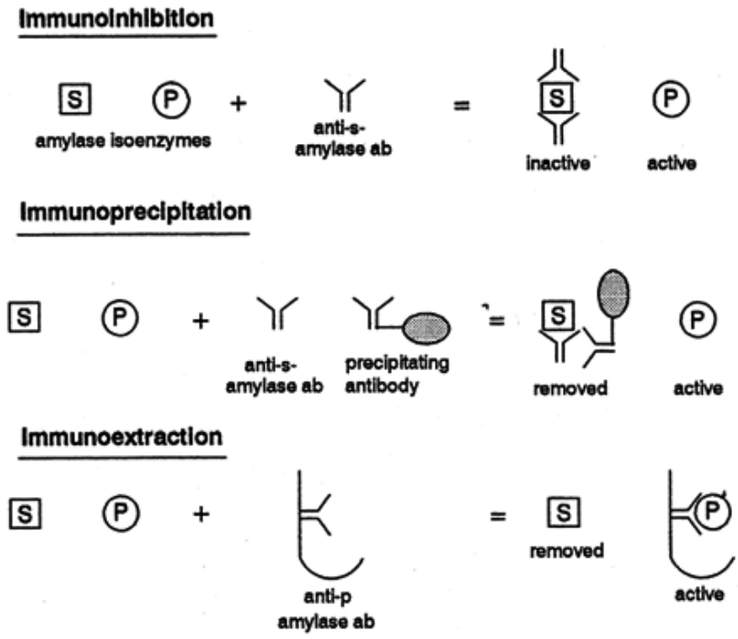


FIGURE 14.21. Analytical methods for measurement of the pancreatic amylase isoenzyme activity, using an appropriate amylase substrate. Immunoinhibition involves the inhibition of salivary isoenzymes with specific salivary (s) amylase antibodies. The residual activity is owing to pancreatic (P) amylase. The immunoprecipitation method involves the removal of salivary amylase with specific antibodies, followed by use of a precipitating antibody. The immunoextraction assay uses antibodies to pancreatic amylase to isolate it from the other forms before enzymatic analysis. (From the American Association for Clinical Chemistry, with permission.)

Lipase (EC 3.1.1.3)

Lipase (triacylglycerol acylhydrolase) catalyzes the hydrolysis of triglycerides sequentially into ß-monoglyceride and two free fatty acids. Lipase activity depends on the substrate being present as an emulsion in contrast to related enzymes such as lipoprotein lipase and other hydrolases, which are active with short fatty acid chain substrates that are water soluble. The most commonly used assay for lipase involves measuring the clearing of a substrate emulsion by the action of lipase. Measurements can be made by either nephelometry or turbidimetry:

$$\underset{\text{(cloudy emulsion)}}{\text{Triolein}} \xrightarrow{\text{lipase}} \underset{\text{(clear solution)}}{\text{glycerides + free fatty acids}}$$

The optimum pH for this reaction is 8.8. In this assay, both lipase and lipoprotein lipase are measured. However, if colipase and a bile salt such as sodium deoxycholate are included, the reaction rate and analytical sensitivity of pancreatic lipase is increased, whereas that for lipoprotein lipase is eliminated (27). Colipase, aided by the addition of bile salts, binds to lipase to form a complex. This association produces a conformational change in lipase, such that the latter can now more efficiently bind to the substrate. The mechanism of how colipase and bile salts interact with the enzyme in vivo is shown in Fig. 14.22. The reference range for lipase depends on the substrate and whether colipase is used. The upper reference limit is 200 U/L at 37°C when triolein is used as the substrate in the presence of colipase and bile salts.

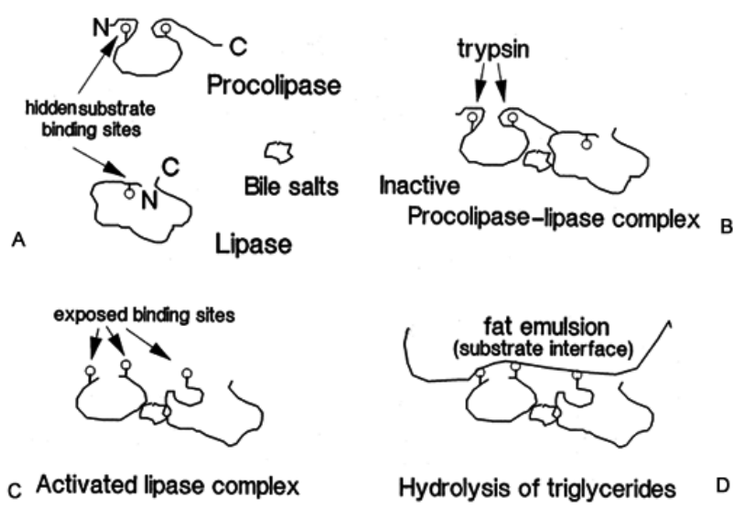


FIGURE 14.22. Mechanism for the hydrolysis of fat in the intestine by lipase–colipase. **A:** Lipase and procolipase are secreted as separate proteins from the pancreas. The binding sites for the triglyceride substrates are hidden for both enzymes. **B:** In the intestine, the two proteins associate to form an inactive complex. The binding is favored by the addition of mixed-chain bile salts. **C:** Trypsin converts procolipase to colipase by removal of the N- and C-terminal residues, exposing the substrate binding sites. This also produces a conformational change for lipase, thereby exposing its substrate binding site. **D:** The lipase-colipase can now attach to the fat particles, facilitating the hydrolysis of triglycerides. (Adapted from Borgstrom B. *Lipase*. Amsterdam: Elsevier, 1984, with permission.)

Electrophoresis of lipase reveals the presence of at least two true isoforms in serum, labeled L_1 and L_2. A third form labeled L_3 also appears on electrophoresis and is thought to be pancreatic carboxyester lipase.

Trypsin (EC 3.4.21.4)

Trypsin is a pancreatic proteolytic enzyme that specifically cleaves peptide bonds at carboxyl groups of lysine and arginine. Trypsin originates from inactive trypsinogens produced by pancreatic acinar cells. Activation of trypsin occurs in the intestinal tract by an enterokinase. Measurement of serum trypsin is seldom performed because amylase and lipase measurements are readily available and provide much the same information. However, immunoreactive trypsin from dried blood spots is being used as a neonatal screen for cystic fibrosis (28). Because of the existence of antiproteases in the blood, trypsin must be measured immunologically by using specific antisera rather than by the use of enzymatic activity.

Clinical Uses of Pancreatic Enzymes

Acute Pancreatitis

Pancreatitis results when digestive enzymes of the exopancreas find their way into the endocrine parenchyma and autodigestion of pancreatic tissue ensues. Key among these digestive enzymes is trypsin, which, together with other bile components, activates other proteolytic enzymes that are collectively responsible for pancreatic necrosis. The most consistent clinical symptom of acute pancreatitis is abdominal pain. Risk factors include excessive alcohol use, the presence of gallstones, and hyperlipidemias. The disease can take either of two forms: the edematous form, which is relatively mild and occurs in approximately 80% of all cases, and the hemorrhagic form, which is more serious and results in most of the fatalities attributed to acute pancreatitis (29).

The diagnosis of acute pancreatitis is made based on clinical presentation and medical history, in addition to radiographic and laboratory studies. Laboratory findings include an elevated serum amylase, amylase clearance, and lipase. Amylase and lipase levels rise within a few hours after onset and remain elevated for 36 to 48 hours. The extent of amylase and lipase elevations is not strongly correlated to the severity of the disease. The clinical sensitivities of serum amylase and lipase are both approximately 90% when blood is sampled in the first 36 hours, and highly sensitive assays are used. False negatives result when pancreatitis occurs in patients who have had recurrent attacks and have little or no functional pancreatic tissue remaining.

The clinical specificity for the diagnosis of acute pancreatitis is only 40% for amylase and 60% for lipase. The specificity for amylase is low because of the existence of nonpancreatic sources of the enzyme such as the salivary gland. The specificity of lipase is influenced by the design of the assay used. When serial amylase levels are measured in patients presenting with acute abdominal pain and a positive history of alcohol use or cholestatic disease, the specificity is increased. However, hyperamylasemia is observed in several other pancreatic as well as nonpancreatic disorders, and the diagnostic value of a single elevated amylase level without regard to other clinical findings is limited. Automated P-amylase isoenzyme determinations are helpful in differentiating pancreatic from salivary and macromolecular hyperamylasemia and may preclude the need for a total amylase measurement.

With improvements in methods and better standardization, a serum lipase measurement is the laboratory test of choice for acute pancreatitis. Turbidimetric assays have been adapted to automated clinical chemistry analyzers for routine and stat determinations. Further improvements in diagnostic efficiencies are obtained with the measurement of amylase and lipase isoforms. Among the markers studied, the P_3 amylase and the L_2 lipase isoforms have the highest clinical specificity for the diagnosis of acute pancreatitis. Whether the increase in clinical efficiency justifies the use of more labor-intensive electrophoretic assays remains to be determined.

The diagnosis of acute pancreatitis can be improved with the use of expert systems or decision algorithms. Rather than using the upper limit of the reference range, a flowchart makes use of higher cutoff values to increased the specificity of individual tests, whereas the use of multiple tests maintains the sensitivity. In the algorithm developed by Panteghini and Pagini (30), two times the upper reference limit for amylase was used as the entry point to the algorithm (Fig. 14.23). The diagnosis of acute pancreatitis required a fivefold increase in lipase. Using this algorithm, a diagnostic accuracy of 95.8% was produced. A neural network for use of pancreatic enzymes was also developed to improve the diagnostic efficiency for acute pancreatitis (31). This neural network produced its own algorithm based on laboratory data and final diagnoses received during the input or “training” phase. During the “test” phase, in which laboratory data but no final diagnoses were inputted, the neural network had an accuracy rate of 84%.

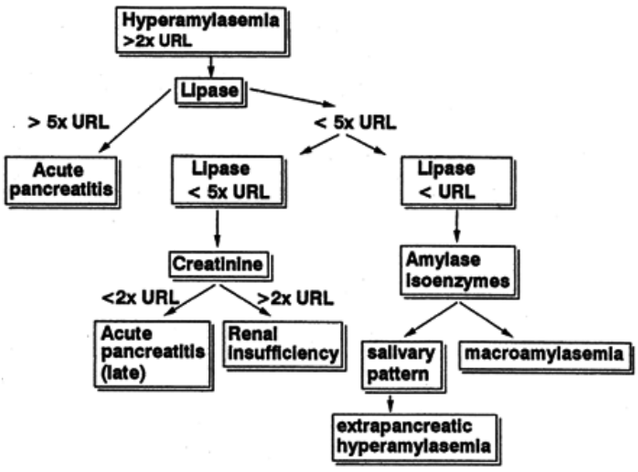


FIGURE 14.23. Algorithm for the diagnosis of acute pancreatitis from extrapancreatic hyperamylasemia, and other nonpancreatic conditions affecting results for amylase and lipase. URL, upper reference limit. (Data from Panteghini M, Pagani F. Clinical evaluation of an algorithm for the interpretation of hyperamylasemia. *Arch Pathol Lab Med* 1991;115:355—358.)

Chronic Pancreatitis

Chronic inflammation of the pancreas can be caused by either recurrent episodes of acute pancreatitis or continual subclinical pancreatic damage occurring over many months. Chronic pancreatitis often presents in patients with a history of acute pancreatitis together with diabetes in adults, and cystic fibrosis in children. Gradual degradation of exocrine function often leads to steatorrhea. Approximately 40% of chronic pancreatitis patients have vitamin B_{12} malabsorption. Examination of fecal fat and administration of the D-xylose tolerance and secretin stimulation tests are useful in determining the presence of these complications. Serum enzyme levels in chronic pancreatitis are less useful than in the acute presentation. Amylase and lipase levels are very often within normal limits and may actually be low during the end stages of the disease because of the small amount of functional pancreatic tissue. Patients with low-grade disease will also not be detectable by enzyme studies. The most useful procedure for diagnosis appears to be endoscopic retrograde choledochopancreatography, which permits direct viewing of pancreatic ducts.

Pancreatic Enzymes in Other Disorders

The optimum use of enzymes for the diagnosis of pancreatitis requires an understanding of other pancreatic and nonpancreatic causes of enzyme release. Several pancreatic injuries and diseases have the potential for releasing enzymes into the blood. These include

pancreatic trauma, abdominal surgery, carcinoma, diabetes mellitus and ketoacidosis, and injury by viruses (such as hepatitis B) and drugs (such as tetracycline, thiazides, and furosemide). Injury to or obstruction of surrounding tissue, such as a perforated ulcer or peritonitis, may also cause compression of the pancreas, resulting in enzyme release. Elevations of amylase, immunoreactive trypsin, and, to a lesser extent, lipase also occur in other nonpancreatic disorders. Because amylase and trypsin are largely cleared by glomerular filtration, high values are seen in patients with renal failure. The existence of parotid lesions and inflammation will produce salivary hyperamylasemia. Amylase isoenzyme measurements can be used to detect this form. Macroamylasemia can also cause high serum amylase values, resulting in a false-positive diagnosis for pancreatitis. Amylase isoenzyme and clearance measurements can be used to detect the presence of this form.

MISCELLANEOUS ENZYMES

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

Acid Phosphatase (EC 3.1.3.2)

The acid phosphatases (orthophosphoric-monoester phosphohydrolase) are a collection of enzymes that catalyze the hydrolysis of phosphate from a variety of natural and synthetic substrates. The optimal pH is on the acid side, usually between 5 and 6. Acid phosphatase is found in the liver, red cells, platelets, bone marrow, and prostate. Several isoenzymes of acid phosphatase can be identified in serum by electrophoresis or IEF. Measurement of acid phosphatase was primarily used for monitoring patients with prostatic cancer. However, with the development of prostate-specific antigen (PSA), assays for acid phosphastase, if used at all, are only used on rare occasions. As such, the reader is referred to the first edition of *Clinical Laboratory Medicine* for details on the enzymatic and immunoassays for acid phosphatase. Discussion of PSA in prostate cancer is discussed in Chapter 26 (Tumor Markers).

Cholinesterase

Analytical Measurement

The cholinesterases are enzymes that hydrolyze acetylcholine. True cholinesterase (acetylcholine acetylhydrolase, EC 3.1.1.7) is found in the red cells, lungs, spleen, and brain. In the serum, however, only pseudocholinesterase [(PCHE) acylcholine acylhydrolase, EC 3.1.1.8] is found. These enzymes differ in their specificity for substrates and sensitivities to inhibitors. Of particular clinical interest is the response of serum cholinesterase genetic variants to inhibition by dibucaine and fluoride.

Serum cholinesterase is measured by monitoring the hydrolysis of acetylthiocholine esters to form thiocholine:

$$\text{acetylthiocholine} + H_2O \xrightarrow{\text{PCHE}} \text{thiocholine} + \text{acetate}$$

The enzyme rate is monitored by reacting thiocholine with a disulfide agent that forms a chromogen after rearrangement and can be measured at 340 nm:

thiocholine + 4,4′-dithiopyridine → → chromogen

Clinical Significance

Serum levels of cholinesterase can be used in three different areas. Low levels of PCHE are found in patients with various forms of liver disease, including acute and chronic hepatitis, cirrhosis, and metastatic liver disease. Nonhepatic disorders, such as AMI, infections, and pulmonary embolism, have also been shown to decrease PCHE levels. Because of the existence of more specific markers for the latter disorders, serum cholinesterase is not used for detecting or monitoring these processes.

The activity of PCHE is also important in monitoring industrial and agricultural workers who use and are exposed to organophosphate insecticides. Commonly used agents such as parathion and malathion inhibit both erythrocyte and serum cholinesterase activity, leading to neurologic and neuromuscular deficits. Levels of these enzymes are useful in determining the extent of exposure to these toxic pesticides.

The third area of clinical application of cholinesterase measurements involves the use of succinylcholine (suxamethonium) as a muscle relaxant in anesthesia. Because of a genetic variant of cholinesterase, there is a substantial difference in the rate of suxamethonium clearance by PCHE in some individuals. Patients who have the atypical cholinesterase gene can experience prolonged apnea (because of extended relaxation of respiratory muscles) after administration of the drug. Recognition of these individuals is important and can be accomplished by measuring serum cholinesterase activity with and without inhibitors such as dibucaine or fluoride. Because atypical PCHE is more resistant to inhibitors than the normal PCHE, measurement of the percentage of inhibition (e.g., the “dibucaine number”) provides a measure of the amount of the atypical enzyme that is present. Normal individuals exhibit 80% to 90% inhibition. Those who are heterozygous for one of the several possible variant genes exhibit 70% to 80% inhibition, and persons with two variant genes exhibit 10% to 30% inhibition. The individuals of the latter group are most prone to breathing complications after administration of suxamethonium.

ACE (EC 3.4.15.1)

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

ACE (peptidyl-dipeptide hydrolase) is produced by the lungs and catalyzes the conversion of the decapeptide angiotensin I to the octapeptide angiotensin II. It is a vital enzyme in the control of aldosterone secretion and regulation of blood pressure. ACE can be measured by the use of synthetic peptide substrates or by immunoassay.

Although ACE is an important step in the renin-aldosterone metabolic pathway, serum measurements of ACE are not useful for studying the etiologies of hypertension. However, it has been found that ACE is elevated in approximately 50% to 80% of cases of sarcoidosis, a multisystem granulomatous disease that most commonly involves the lungs. ACE is useful for differentiation between this disease and other granulomatous disorders and between active and dormant sarcoidosis. Unfortunately, the specificity of ACE measurements is not high because several other conditions can produce elevated activities. These include Gaucher's disease, leprosy, active histoplasmosis, pulmonary embolism, Hodgkin's disease, alcoholic cirrhosis, and chronic hepatitis.

G6PD (EC 1.1.1.49)

Part of "14 - Diagnostic Enzymology and Other Biochemical Markers of Organ Damage"

G6PD is an important erythrocyte enzyme in the hexose monophosphate shunt for the utilization of glucose. The in vivo reaction of G6PD is used in its laboratory analysis:

$$\text{glucose-6-phosphate} + \text{NAD}^+ \xrightarrow{\text{G6PD}} \text{6-phosphogluconate} + \text{NADH}$$

The reaction is monitored at 340 nm by following the production of NADH. Measurements of G6PD are clinically useful for the detection of individuals who are deficient for the X-linked G6PD gene. G6PD deficiency affects black men and whites of Mediterranean descent. These individuals develop varying degrees of hemolytic anemias when exposed to particular drugs, such as sulfa drugs, aspirin, and antimalarial medications.

REFERENCES

1. LaDue JS, Wroblewski F, Karmen A. Serum glutamic oxaloacetic transaminase activity in human acute transmural myocardial infarction. *Science*1954;12:497–499.

2. Pappas NJ Jr. Theoretical aspects of enzymes in diagnosis. *Clin Lab Med* 1989;9:595–626.

3. Lehninger AL. *Biochemistry*, 2nd ed. New York: Worth, 1970:183–216.

4. Nemesanszky E, Lott JA. γ-Glutamyltransferase and its isoenzymes: progress and problems. *Clin Chem* 1985;31:797–803.

5. Zimmerman HJ. Function and integrity of the liver. In: Henry JB, ed. *Clinical diagnosis and management*, 17th ed. Philadelphia: Saunders, 1984.

6. Cohen JA, Kaplan MM. The SGOT/SGPT ratio an indicator of alcoholic liver disease. *Dig Dis Sci* 1979;24:835–838.

7. Laposata M. Fatty acid ethyl esters: short-term and long-term serum markers of ethanol intake. *Clin Chem* 1997;43:1527–1534.

8. Oellerich M, Burdelski M, Ringe B, et al. Lignocaine metabolite formation as a measure of pre-transplant liver function. *Lancet* 1989;1:640–642.

9. Morovat A, Trull A, Maguire G. Serum fructose-1,6-bisphosphatase measurements for the early detection of acute hepatocellular damage [Abstract]. *Clin Chem Lab Med* 1999;37(Suppl):S472.

10. Hughes VF, Trull AK, Gimson A, et al. Randomized trial to evaluate the clinical benefits of serum alpha-glutathione S-transferase concentration monitoring after liver transplantation. *Transplantation* 1998;64:1446–1452.

11. Braunwald E. Unstable angina. A classification. *Circulation* 1989;80:410–414.

12. Hearse DJ. Cellular damage during myocardial ischaemia: metabolic changes leading to enzyme leakage. In: Hearse DJ. *Enzymes in cardiology: diagnosis and research*. New York: John Wiley and Sons, 1979.

13. Gruemer HD, Prior T. Carrier detection in Duchenne muscular dystrophy: a review of current issues and approaches. *Clin Chim Acta* 1987;162:1–18.

14. Lott JA, Wolf PL.*Clinical enzymology. A case-oriented approach*. New York: Year Book, 1986.

15. Wu AHB, Schwartz JG. Update on creatine kinase isoenzyme assays. *Diagn Clin Test* 1989;27:16–20.

16. Gilkeson G, Stone MJ, Waterman M, et al. Detection of myoglobin by radioimmunoassay in human sera: its usefulness and limitations as an emergency room screening test for acute myocardial infarction. *Am Heart J* 1978;95:70–77.

17. Wu AHB, Apple FS, Gibler WB, et al. Recommendations for use of cardiac markers in coronary artery diseases. *Clin Chem* 1999;45:1104–1121.

18. Wu AHB. The patient with chest pain. In: Price CP, Hicks JM, eds., *Point-of-care testing*. Washington, DC: AACC Press, 1999:429–448.

19. Report of the Joint International Society and Federation of Cardiology/World Health Organization Task Force on Standardization of Clinical Nomenclature: nomenclature and criteria of diagnosis of ischemic heart disease. *Circulation* 1979;59:607–609.

20. Wu AHB, Feng YJ, Moore R, et al. Characterization of cardiac troponin subunit release into serum following acute myocardial infarction, and comparison of assays for troponin T and I. *Clin Chem* 1998;44:1198–1208.

21. Abbas SA, Glazier JJ, Wu AHB, et al. An analysis of the relative sensitivities of cardiac troponin T and of CK-MB in the detection of myocardial injury during percutaneous transluminal coronary angioplasty. *Clin Cardiol* 1996;19:782–786.

22. Olatidoye AG, Wu AHB, Feng YJ, et al. Prognostic role of troponin T versus I in unstable angina for cardiac events with meta-analysis comparing published studies. *Am J Cardiol* 1998, 81:1405–1411.

23. Hamm CW, Heeschen C, Goldmann B, et al. Benefit of abciximab in patients with refractory unstable angina in relation to serum troponin T levels. *N Engl J Med* 1999;340:1623–1629.

24. Apple FS, Rogers MA, Sherman WM, et al. Comparison of serum creatine kinase and creatine kinase MB activities post marathon race versus post myocardial infarction. *Clin Chim Acta* 1984;138:111–118.

25. Wu AHB, Laios I, Green S, et al. Immunoassays for serum and urine myoglobin: myoglobin clearance assessed as a risk factor for acute renal failure. *Clin Chem*1994;40:796–802.

26. Remaley AT, Wilding P. Macroenzymes: biological characterization, clinical significance, and laboratory detection. *Clin Chem* 1989;35:2261–2270.

27. Borgstrom B. *Lipase*. Amsterdam: Elsevier, 1984.

28. Ranieri E, Ryall RG, Morris CP, et al. Neonatal screening strategy for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis. *BMJ* 1991;302:1237–1240.

29. Wong WCC, Butch AW, Rosenblum JL. The clinical chemistry laboratory and acute pancreatitis. *Clin Chem* 1993;39:234–243.

30. Panteghini M, Pagani F. Clinical evaluation of an algorithm for the interpretation of hyperamylasemia. *Arch Pathol Lab Med* 1991;115:355–358.

31. Kazmierczak SC, Catrou PG, Van Lente F. Diagnostic accuracy of pancreatic enzymes evaluated by use of multivariate data analysis. *Clin Chem* 1993;39:1960–1965.

15

Lipids and Lipoproteins

György Ábel

Michael Laposata

Lipids, also known as fats, include many different compounds that have limited solubility in water. Phospholipids, triglycerides, glycolipids, and cholesterol esters have fatty acids as major building blocks. Free cholesterol is the one major lipid component that does not have a fatty acid as a component. This chapter begins with a description of the structure and metabolism of fatty acids and then the larger fatty acid containing molecules and cholesterol. The structure and metabolism of lipoproteins and a classification of lipid disorders are also included. The chapter concludes with a discussion of the laboratory measurement of various lipids and lipoproteins and the clinical significance of the most frequently encountered abnormalities.

FATTY ACIDS

Part of "15 - Lipids and Lipoproteins"

The fatty acid molecule shown in Fig. 15.1 is one of the fundamental building blocks of lipids. A long-chain fatty acid is shown with the carboxyl end to the right and the methyl end to the left. The carboxyl end of the molecule combines with other molecules when the fatty acid is esterified within more complex lipids. This figure shows palmitic acid, which has 16 carbons and no double bonds, as indicated by the nomenclature 16:0.

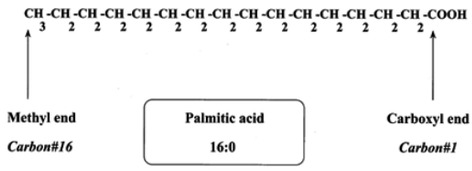


FIGURE 15.1. Molecular structure of a representative fatty acid.

Table 15.1 shows the nomenclature in more detail for several fatty acids. The number to the left of the colon indicates the number of carbon atoms in the fatty acid. The number to the right of the colon indicates the number of double bonds. For example, linoleate Δ18:2 (9,12), also known as linoleic acid, has 18 carbons and two double bonds. The numbers after Δ indicate the locations of the double bonds. In linoleate, there are two double bonds, one between the ninth and tenth carbon and one between the twelfth and thirteenth carbon. As shown in Table 15.1, the five most abundant animal fatty acids are palmitate, stearate, oleate, linoleate, and arachidonate. Many other fatty acids exist in nature.

TABLE 15.1. THE MOST ABUNDANT ANIMAL FATTY ACIDS

Familiar Name	Numerical Nomenclature
Palmitate	16 : 0
Stearate	18 : 0
Oleate	18 : 1 Δ 9
Linoleate	18 : 2 Δ 9, 12
Arachidonate	20 : 4 Δ 5, 8, 11, 14

Fatty acids that have no double bonds, such as palmitate and stearate, are known as saturated fatty acids. Fatty acids like oleate with one double bond are called monounsaturated, and polyunsaturated fatty acids are those with two or more double bonds.

Fatty acids that contain one or more double bonds can be classified as essential or nonessential, and by their omega family. Essential fatty acids are those that must be ingested because humans cannot synthesize them from precursor molecules. Nonessential fatty acids are those that can be synthesized from precursor molecules such as sugars. The designation of the omega family of the fatty acid as (n–x) or (omega–x) is established by counting the number of carbons from the methyl end to the carbon with the nearest double bond. For example, omega-6 fatty acids have the first double-bond six carbons inward from the methyl end. For omega-3 fatty acids, the first double bond from the methyl end is three carbons inward. It should be noted that although fatty acids can be metabolized into other fatty acids, they cannot change families. This means that any modification to the fatty acid does not alter the position of the last double bond.

Fatty acids can also be classified according to the stereochemical configuration of their double bonds as cis or trans. The naturally occurring form of a double bond in a fatty acid is the cis form, with the hydrogen atoms on the same side of the carbon atoms in the double bond (Fig. 15.2). Trans double bonds have the hydrogen atoms on opposite sides of the double bond. Trans fatty acids are present in significant concentration in margarine and have a straight-chain, three-dimensional structure, like saturated fatty acids. This is unlike a fatty acid with a cis double bond because the insertion of a cis double bond produces a bend in the fatty acid molecule resulting in a different three-dimensional

structure. Presumably because of their similar three-dimensional structure, trans fatty acids and saturated fatty acids both elevate serum cholesterol levels.

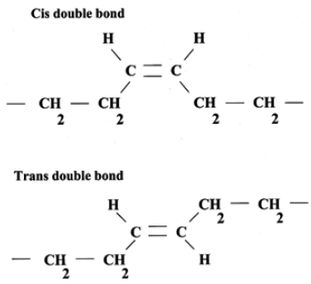


FIGURE 15.2. Stereochemical configuration of cis and trans double bonds. In the cis form, the hydrogen atoms are on the same side of the carbon atoms in the double bond, whereas trans double bonds have the hydrogen atoms on the opposite sides.

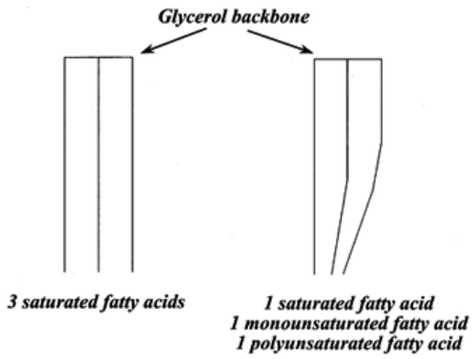


FIGURE 15.3. Structure and composition of triglycerides.

TRIGLYCERIDES

Part of "15 - Lipids and Lipoproteins"

The structure of a triglyceride is shown in Fig. 15.3. A triglyceride is three fatty acids on a glycerol backbone. On the left, a triglyceride molecule is shown with three straight chains of saturated fatty acids. On the right, a triglyceride molecule is shown with one saturated fatty acid in the left position (known as *sn*-1), a monounsaturated fatty acid in the middle (the *sn*-2 position), and a polyunsaturated fatty acid in *sn*-3.

CHOLESTEROL AND CHOLESTEROL ESTERS

Part of "15 - Lipids and Lipoproteins"

Figure 15.4 shows the structure and critical intermediates in the synthesis of cholesterol, another fundamental lipid molecule. Cholesterol contains a sterol nucleus with 27 carbons. It can be synthesized *de novo* from the progressive accumulation of two carbon units and can be ingested by eating dairy products, meat, and some fish. Cholesterol is only present in animals and animal products and cannot be derived from eating plants.

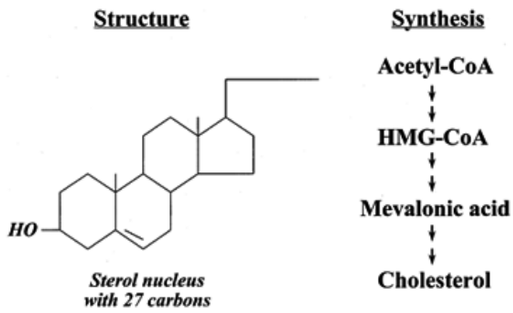


FIGURE 15.4. Molecular structure and biosynthesis of cholesterol.

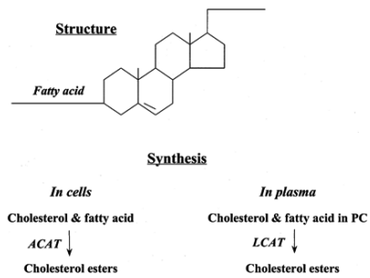


FIGURE 15.5. Structure and biosynthesis of cholesterol esters. ACAT, acyl-CoA cholesterol acyltransferase; LCAT, lecithin cholesterol acyltransferase; PC, phosphatidylcholine

Cholesterol esters are derived from cholesterol by esterification to the third carbon of the cholesterol molecule with a fatty acid (Fig. 15.5). The synthesis of cholesterol in cells is catalyzed by an enzyme, acyl-CoA acyltransferase (ACAT). In the plasma, another enzyme, lecithin cholesterol acyltransferase (LCAT), catalyzes the formation of cholesterol esters from cholesterol and one of the two fatty acids in phosphatidylcholine (PC), also known as lecithin.

PHOSPHOLIPIDS

Part of "15 - Lipids and Lipoproteins"

Phospholipids, like triglycerides, have a glycerol backbone but with only two esterified fatty acids. The third carbon is connected to a polar head group through a phosphate. The most common head groups are choline, ethanolamine, inositol, and serine (Fig. 15.6). There are many others present in much lower amounts. The polar head groups are hydrophilic and therefore water soluble. The water solubility of the head groups and the hydrophobic nature of the apolar fatty acid chains are responsible for the unique membrane-forming feature of phospholipids. It is much more common to have a polyunsaturated or monounsaturated fatty acid in the *sn*-2 position than in the *sn*-1 position.

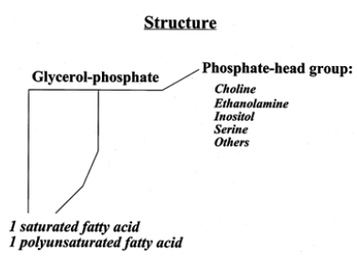


FIGURE 15.6. Structure of phospholipids.

INFLUENCE OF DIETARY FATTY ACID INTAKE ON SERUM CHOLESTEROL CONCENTRATION

Part of "15 - Lipids and Lipoproteins"

There are important clinical issues related to the amount and composition of fatty acids in the diet. Table 15.2 shows the association between serum cholesterol and the ingestion of particular dietary fatty acids. The ingestion of saturated fatty acids, usually contained within triglycerides, with 12 or 14 carbons will elevate the serum cholesterol level. The ingestion of 16:0 fatty acids will also produce an elevation in serum cholesterol. Collectively, the studies suggest that 18:0 fatty acid ingestion is either modestly detrimental or has no effect on the cholesterol level. The ingestion of polyunsaturated fatty acids, such 18:2 or 20:4, generally has a neutral to beneficial effect on the total cholesterol when ingested in modest amounts. Trans fatty acids are detrimental, like saturated fatty acids, and elevate the serum cholesterol concentration.

TABLE 15.2. EFFECT OF THE CHAIN LENGTH AND SATURATION OF DIETARY FATTY ACIDS ON SERUM CHOLESTEROL LEVELS

Fatty Acid	Effect on Serum Cholesterol
12 : 0	Very detrimental
14 : 0	Very detrimental
16 : 0	Detrimental
18 : 0	Modestly detrimental or neutral
20 : 4	Beneficial
Trans fatty acids	Detrimental

Table 15.3 shows the fatty acid composition of a variety of different food fats. Beef tallow, butter, cocoa butter, palm oil, coconut oil, and lard have high concentrations of saturated fatty acid. Coconut oil is particularly dangerous because of its very high concentration of 12:0 and 14:0 fatty acids. Oils with high concentrations of monounsaturated fatty acids include olive oil and canola oil. There are many oils rich in polyunsaturated fatty acids, such as corn oil, cottonseed oil, peanut oil, and soybean oil. Menhaden oil contains fatty acids of the omega-3 family, which have a high number of double bonds. The omega-3 fatty acids have multiple health benefits, which are discussed in a later section. There is some variability between different fish species in the amount of 20:5 n–3 and 22:6 n–3, the major omega-3 fatty acids, which depends in part on where the fish swim (Table 15.4). Fish that swim in cold water have more omega-3 fatty acids than those that swim in warm water.

TABLE 15.4. PRINCIPAL FATTY ACIDS OF SOME OF THE MAJOR FISH OF COMMERCE

	Fatty Acid							
Fish	14:0	16:0 n-7	16:1 n-9	18:1 n-9	20:1 n-9	22:1 n-3	20:5 n-3	22:6
Atlantic menhaden	9	18	11	10	1	T	15	11
Gulf menhaden	9	20	12	11	1	T	13	7
Winterized menhaden	7	15	10	15	3	2	17	10
Pilchard	8	17	9	12	5	3	17	9
Capelin	7	10	10	14	17	14	8	6
Herring	7	16	6	13	13	20	5	6
Anchovy	9	19	9	13	5	2	17	6
Cod Liver	3	13	10	23	T	5	11	12
Sardine	7	17	7	13	8	6	15	10
Mackeral	6	15	5	19	11	8	11	11
Norway pout	6	17	6	19	9	9	11	11
Sandeel	7	17	10	7	13	17	11	7
Sprat	—	17	7	15	10	13	7	11

T, trace.

TABLE 15.3. FATTY ACID COMPOSITION OF COMMON FOOD FAT

Fatty Acid	Menhaden Oil	Soybean Oil	Peanut Oil	Sunflower Oil	Corn Oil	Canola Rapeseed Oil	Butter Oil	Coconut Oil	Olive Oil	Palm Oil	Beef Tallow
4:0	—	—	—	—	—	—	3.2	—	—	—	—
6:0	—	—	—	—	—	—	1.9	0.6	—	—	—
8:0	—	—	—	—	—	—	1.1	7.5	—	—	—
10:0	—	—	—	—	—	—	2.5	6.0	—	—	—
12:0	—	—	—	—	—	—	2.8	44.6	—	—	0.9
14:0	9.0	0.2	0.1	0.1	—	—	10.1	16.8	—	—	3.7
16:0	19.0	10.7	9.5	5.8	11.5	4.8	26.3	8.2	11.0	—	24.9
18:0	3.8	3.9	2.3	4.1	2.2	1.5	12.1	2.8	2.2	0.1	18.9
20:0	—	0.2	1.4	0.3	—	0.6	—	—	0.8	1.0	—
16:1n-7	11.4	0.3	—	—	—	0.5	2.3	5.8	72.5	43.5	4.2
18:1n-9	10.6	22.8	45.6	21.7	26.6	53.2	25.1	—	0.3	4.3	36.0
20:1n-9	1.7	—	1.2	—	—	1.0	—	—	—	0.3	0.3
22:1n-9	0.6	—	—	—	—	0.2	—	1.8	7.9	36.6	—
18:2n-6	2.0	50.8	31.0	66.4	58.7	22.2	2.3	—	0.6	0.1	3.1
18:3n-3	1.0	6.8	—	0.3	0.8	11.0	1.4	—	—	—	0.6
20:5n-3	14.0	—	—	—	—	—	—	—	—	9.1	—
22:6n-3	8.6	—	—	—	—	—	—	—	—	0.2	—

The numbers refer to the percentage of an individual species of fatty acids relative to total fatty acids in the food fat.

There are many epidemiologic data to support the inverse correlation between fish consumption and mortality from coronary heart disease. An increase in fish consumption results in a decrease in the incidence of heart attack and stroke. Early supporting evidence was provided by the finding that Greenland Eskimos consuming more than 400 g of fish per day have a very low death rate from coronary heart disease. It was subsequently shown in epidemiologic studies that Japanese on the island of Okinawa, who consume twice the amount of fish eaten by the Japanese living on the mainland, have a much lower death rate from coronary heart disease, that Japanese from fishing villages who consume 250 g or more of fish per day have significantly less coronary heart disease than Japanese from farming villages who consume no more than 90 g of fish per day, and that Dutch men consuming more than 30 g of fish per day, or on average two fish meals per week, had 50% lower mortality from coronary heart disease than men who did not eat fish.

LIPOPROTEINS

Part of "15 - Lipids and Lipoproteins"

Figure 15.7 shows the structure of a lipoprotein, with a shell composed of one layer of phospholipids with head groups directed outward, cholesterol, and apolipoproteins, and a core with very hydrophobic triglyceride and cholesterol molecules.

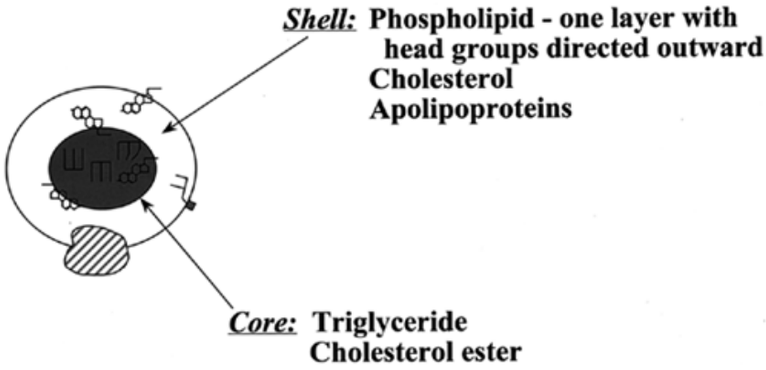


FIGURE 15.7. Schematic structure of a lipoprotein particle.

Lipoproteins are classified into four major groups based on the size of the particle and the content of the core of the lipoprotein (Fig. 15.8). Chylomicrons (CM) are very large lipoproteins formed after ingestion of dietary fat and contain mostly triglyceride in the core. An approximate ratio of triglyceride to cholesterol ester is 20:1. The predominant apolipoprotein in CM is apolipoprotein B. Very low density lipoprotein (VLDL) also has more triglyceride than cholesterol ester in its core, with a ratio of approximately five triglycerides to one cholesterol. There are many apolipoproteins associated with VLDL, with apolipoprotein C and apolipoprotein B predominating. Low-density lipoprotein (LDL) has more cholesterol ester than triglyceride in the core and contains only apolipoprotein B in the shell. High-density lipoprotein (HDL) contains mostly cholesterol ester in the core, and its predominant apolipoproteins are A-I and A-II.

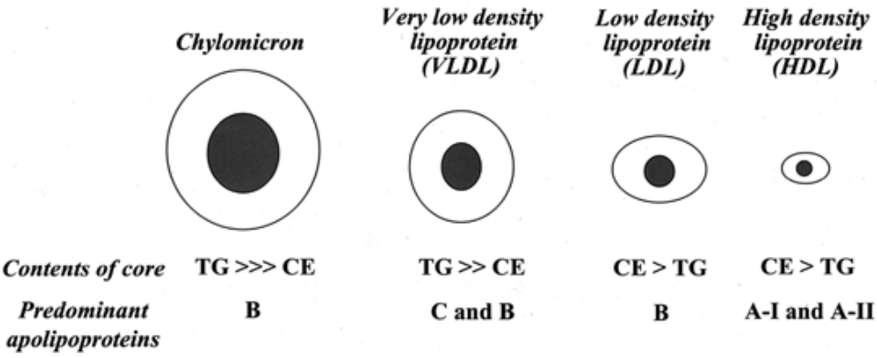


FIGURE 15.8. Comparison of size and composition of the core in the four major classes of lipoproteins. CE, cholesterol ester; TG, triglyceride; Apo, apolipoprotein.

There are multiple subtypes of the apolipoproteins, and each of the lipoprotein classes exists in a continuum of sizes and densities because of differences in the content of the core lipids. In addition, there are many apolipoproteins. Some are related to each other and have the same letter description with different numbering. For example, apolipoproteins A-I to A-IV are primarily present in HDL and CM. Apolipoprotein B-48 is synthesized in the intestines and present in CM, whereas apolipoprotein B-100 is produced by the liver and is present in VLDL, intermediate density lipoprotein (IDL), and LDL. Apoliproteins C-I, C-II, and C-III are synthesized in the liver and a component of CM, VLDL, IDL, and HDL. Apolipoprotein E has three major isoforms noted as apolipoproteins E-2, E-3, and E-4. Apolipoprotein E is produced by the liver and

peripheral tissues, and the isoforms are components of CM, VLDL, IDL and HDL.

Figure 15.9 shows the structure of a unique lipoprotein, lipoprotein (a) [Lp(a)]. This is a LDL-like lipoprotein because the apolipoprotein is apolipoprotein B and the composition of the core resembles that of an LDL particle. It is different from LDL in that apolipoprotein (a) is covalently bound to the apolipoprotein B. The apolipoprotein (a) structure is analogous to plasminogen, a precursor of plasmin, which dissolves clots.

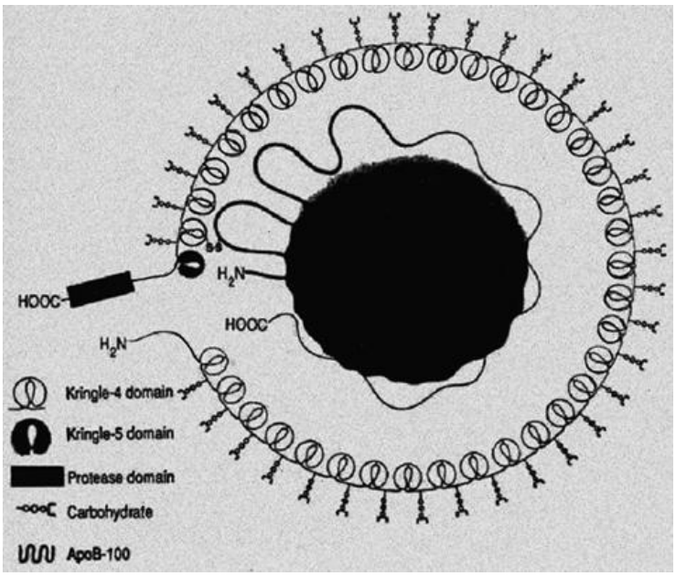


FIGURE 15.9. The unique structure of lipoprotein(a). (From Utermann G. The mysteries of lipoprotein(a). *Science* 1989;246:904–910, with permission.)

It has been demonstrated in multiple studies that an elevated Lip(a) level presents an increased risk for myocardial infarction (MI). The odds ratio for developing a MI for an individual with a Lp(a) level in the top 25% of the values from all individuals and younger than 60 years of age is on the order of two- to fourfold. The Lp(a) concentration is genetically determined and may explain the genetic predisposition to coronary heart disease in some families.

Figure 15.10 shows how LDL is incorporated into cells and the impact of LDL incorporation on cellular metabolism. LDL is shown binding to its receptor. Upon incorporation, the cholesterol in the LDL particle inhibits the synthesis of new cholesterol inside the cell and decreases the synthesis and mobilization of the receptors to the cell surface. This limits further uptake of LDL into the cell and preserves cell function.

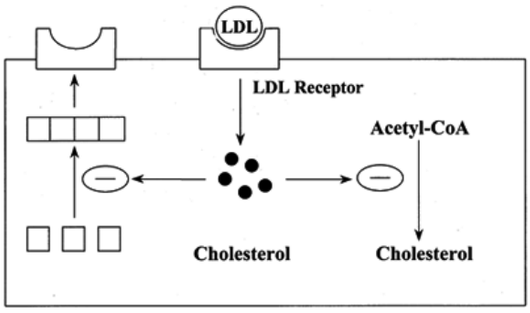


FIGURE 15.10. Low-density lipoprotein (LDL) uptake into cells and its inhibitory effect on cholesterol synthesis and LDL receptor synthesis and recycling.

Figure 15.11 illustrates the consequences of elevated LDL in the plasma, resulting from continued accumulation of LDL and decreased cellular LDL uptake. With a high concentration of LDL in the circulation, some may be oxidized, usually within the blood vessel wall. The oxidized LDL can be ingested by macrophages, which have scavenger receptors rather than LDL receptors. These cells can incorporate large amounts of oxidized LDL and may become “foam cells.”

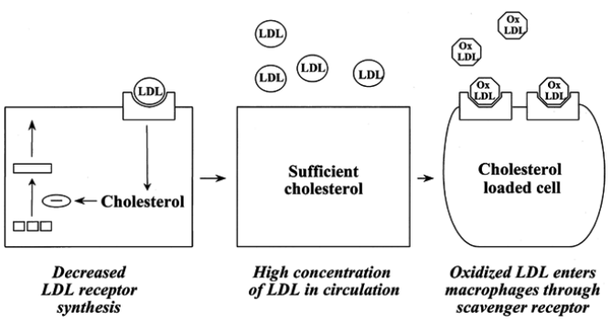


FIGURE 15.11. Pathophysiologic consequences of elevated LDL.

Blood lipid concentrations can be modified by several drugs. When bile acids are secreted into the gut from degradation of cholesterol, bile acid sequestrants bind the acids in the intestine and prevent their reuptake, creating a gradient to promote bile acid secretion into the intestine. The hydroxymethylglutaryl (HMG-CoA) reductase inhibitors (known as the statin drugs) inhibit cholesterol synthesis from precursor molecules by inhibiting the enzyme HMG-CoA reductase that converts HMG-CoA to mevalonate. Nicotinic acid, fibric acid derivatives, and other lipid-lowering drugs work by different mechanisms and alter lipoprotein metabolism in a favorable way.

Figure 15.12 shows what happens to a LDL particle when it is oxidized. At first, it is minimally modified. At this point, it can still bind to the LDL receptor. When it progresses to fully oxidized LDL, binding to the LDL receptor no longer occurs. Native LDL contains apolipoprotein B, a ligand for the LDL receptor. In the

process of oxidation, the polyunsaturated fatty acids in the lipoprotein are oxidized. The oxidized fatty acid breaks into aldehyde fragments, which bind to lysine residues in the apolipoprotein. This biochemical modification prevents the binding of the lipoprotein particle to the LDL receptor. The oxidized LDL is ingested by macrophages by way of the scavenger receptors.

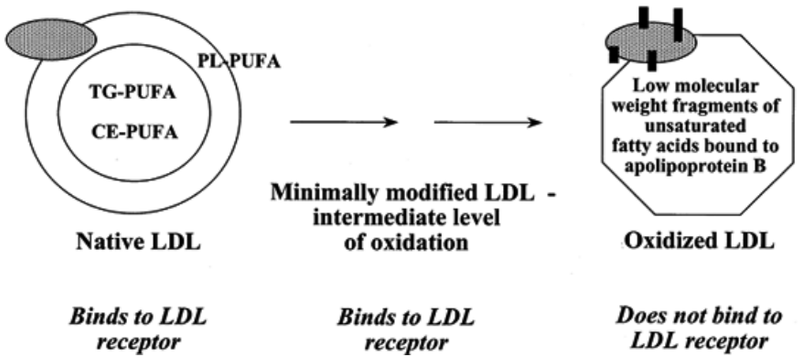


FIGURE 15.12. Effect of minimal and extensive oxidation on the fate of the low-density lipoprotein particle. PL-PUFA, phospholipids containing polyunsaturated fatty acids; TG-PUFA, triglycerides containing polyunsaturated fatty acids.

Figure 15.13 shows excess LDL moving through the circulatory system. It can take a brief detour into the vessel wall, where it can become oxidized. The oxidized LDL then promotes the adhesion of circulating monocytes to the endothelial cells, and then induces them to proliferate and differentiate into macrophages when inside the blood vessel wall. The macrophage then ingests the oxidized LDL and becomes a foam cell. There are a number of cytokines associated with this process. One is platelet-derived growth factor delivered by platelets into the blood vessel wall, which enhances the proliferation of smooth muscle cells. The accumulation of foam cells and smooth muscle cells results in the generation of the atherosclerotic plaque (Fig. 15.14). The fracture of this plaque leads to a thrombotic event, most commonly a MI from ischemia in the coronary arteries or a stroke from plaque rupture in the carotid or cerebral circulation.

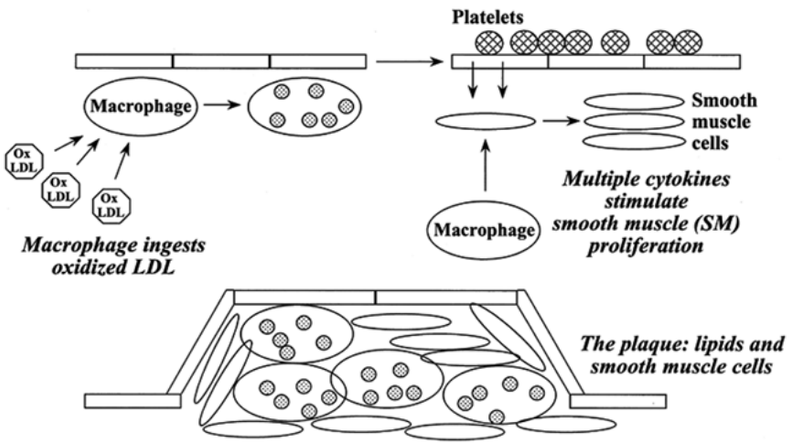


FIGURE 15.14. Mechanism of the formation of the atherosclerotic plaque.

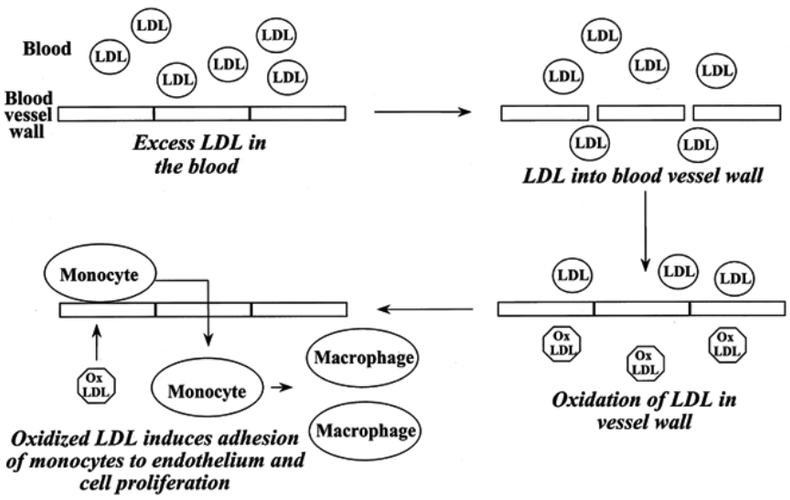


FIGURE 15.13. Events leading to the incorporation of oxidized low-density lipoprotein (LDL) into the blood vessel wall.

FAMILIAL HYPERLIPIDEMIAS

Part of "15 - Lipids and Lipoproteins"

Patients with very high cholesterol levels or other dyslipidemias may have inherited a genetic defect in lipid metabolism. Serum electrophoresis with lipid staining had been widely used for the characterization and classification of hyperlipoproteinemias. Although electrophoresis has been phased out as a diagnostic assay in most laboratories, the classification of hyperlipidemias based on the electrophoretic pattern is still in use. Six different hyperlipoproteinemia patterns have been distinguished based on the electrophoretic pattern. Type I is a severe chylomicronemia secondary to either the absence of lipoprotein lipase of apolipoprotein C-II. Type I hyperlipoproteinemia is indicated by a thick white creamy layer of CM that occupies the upper half of the tube after overnight refrigeration of serum from these patients. Type II hyperlipoproteinemia is characterized by an increase of serum LDL. This abnormality is a result of a genetic defect in lipid or lipoprotein metabolism or secondary to other diseases, such as hypothyroidism, nephrotic syndrome, and diabetes mellitus. Type II is further classified as type IIa and type IIb. In type IIa, only LDL is increased, whereas in type IIb there is an increase in both LDL and VLDL. Refrigerated serum from patients with type IIa lipoproteinemia has a deeper orange color but no significant turbidity. In type IIb, the orange color is combined with moderate turbidity. Type III hyperlipoproteinemia is also called dysbetalipoproteinemia and is characterized by an abnormal accumulation of VLDL that runs in the β region in electrophoresis, unlike the typical VLDL which has pre-β mobility. Many patients with this hyperlipidemia are homozygous for apolipoprotein E-2. The refrigerated serum from type III patients has similar appearance to the serum from type IIb individuals, except that a small creamy layer of CM is also present on the top. The type IV phenotype is characterized by an increase in VLDL. A variety of genetic defects or other diseases are responsible for this disorder. A prominent VLDL (pre-β) band on the electrophoretic pattern is characteristic, and the refrigerated serum is turbid throughout the tube. In type V hyperlipoproteinemia, both VLDL and CM are increased. These patients have defective lipolysis of triglycerides, and usually their lipoprotein lipase level is decreased. Defects in triglyceride metabolism leading to an overproduction of VLDL can occur as a result of excessive alcohol intake, obesity, or diabetes mellitus. The serum pattern shows turbidity throughout the tube with an added layer of CM on the top.

In familial hypercholesterolemia, LDL receptors are defective or deficient. Affected individuals have a cholesterol level on average approximately twice the normal level. Several genetic defects in the synthesis and expression of a LDL receptor have been identified. Patients heterozygous for a defective LDL receptor are prone to premature heart disease with a MI during the fourth or fifth decade. Tendon xanthomas on the dorsum of the hand and the Achilles tendon and corneal arcus are common and important clinical signs of this disorder. The majority of patients with moderate to severe hypercholesterolemia, however, do not have heterozygous abnormality of the LDL receptor. Several genetic factors may be responsible for the condition of these patients. However, the ingestion of too much saturated fat and cholesterol is the most common explanation for the elevation in serum LDL concentration.

LABORATORY MEASUREMENT OF LIPIDS AND LIPOPROTEINS AND THEIR CLINICAL SIGNIFICANCE

Part of "15 - Lipids and Lipoproteins"

In this section, the clinical laboratory aspects of blood lipids are discussed. A more detailed review of the methods is given at the end of this chapter. The total cholesterol value is a quantitation of all the cholesterol and cholesterol esters in all the lipoproteins taken together. HDL cholesterol can be measured in samples from which the LDL and VLDL have been removed by precipitation. The remaining cholesterol and cholesterol esters are in HDL. HDL cholesterol is measured using the same enzymatic reaction as is used for a total cholesterol quantitation. There is an insignificant amount of cholesterol in CM, and it does not influence the total or HDL cholesterol value.

For determination of the calculated LDL cholesterol, the patient must be fasting, because LDL cholesterol is calculated using an equation that includes the triglyceride level, and the serum triglyceride concentration is influenced by recent intake of dietary fat. There are also simple assays for direct LDL measurement that avoid the need to fast. The calculated LDL remains the most popular approach to LDL measurement at present. Unlike LDL cholesterol, the total cholesterol and HDL cholesterol measurements can be made on nonfasting specimens because the dietary intake of lipid is primarily triglyceride rather than cholesterol. Triglycerides are usually ingested in gram amounts, with milligram amounts of ingested cholesterol.

Calculation of LDL cholesterol is performed by subtracting from the total cholesterol the HDL cholesterol level and an approximation of VLDL cholesterol, which is the triglyceride level divided by 5. As noted earlier, in the VLDL particle, there are approximately five triglyceride molecules for every cholesterol molecule (a more detailed description is given later in the LDL cholesteral section of this chapter (p. 319). Triglycerides in excess of 400 mg/dL interfere with the determination of the LDL cholesterol level.

The analytical variability for the typical cholesterol assay in a clinical laboratory is very low, with values on the order of 3% to 5%. Thus, a repeat analysis of the same specimen will give very similar values. Importantly, however, there is a substantial biological variability when a patient is tested on different days. The biological variability ranges from 10% to 20%. Therefore, a person with a cholesterol of approximately 200 mg/dL may, not uncommonly, have values of 180 mg/dL or 220 mg/dL. For this reason, the most accurate assessment of a patient's total cholesterol is established by at least two cholesterol measurements at least 1 week apart.

The measurement of serum or plasma triglycerides is a quantitation of triglyceride in all lipoproteins taken together. This measurement is greatly affected by the ingestion of dietary fat. An accurate measurement of baseline serum triglyceride concentration can only be obtained when the patient has been fasting for 12 to 14 hours.

There are several immunoassays that can be performed for apolipoproteins. There are assays for apolipoprotein B as a measurement of the number of LDL particles; apolipoprotein A concentration reflects HDL levels. These assays have very limited clinical utility. The assay for apolipoprotein (a) as a measure of

Lp(a) is in a state of flux because there is heterogeneity in Lp(a) particles, some of which may be more atherogenic than others, and some assays for Lp(a) measure selected Lp(a) fractions rather than total Lp(a).

The National Cholesterol Education Panel has identified risk factors and developed algorithms for assessment of patients for atherosclerotic vascular disease (Table 15.5 and Table 15.6 and Fig. 15.15, Fig. 15.16 and Fig. 15.17). One risk factor is being a man 45 years or older or a woman 55 years or older in women, or premature menopause not treated with estrogen replacement therapy. Another risk factor is a positive family history of premature coronary heart disease if there is a definite MI or sudden death of the father or other first-degree male relative before the age of 55 years or before the age of 65 of the mother or other female first-degree relative. Cigarette smoking, hypertension with a blood pressure greater than 140/90 mm Hg or treatment with antihypertensives, an HDL cholesterol less than 35 mg/dL, and diabetes mellitus are also risk factors.

TABLE 15.5. RISK FACTORS FOR THE DEVELOPMENT OF ATHEROSCLEROTIC VASCULAR DISEASE

Age (y)
Men ≥45
Women ≥55 or premature menopause without estrogen replacement therapy
Current cigarette smoking
Definite myocardial infarction or sudden death before 55 years of age in father or other male first-degree relative, or before 65 years of age in mother or other female first degree relative
Hypertension
Blood pressure ≥140/90 mm Hg or taking antyhypertensive medication
Low HDL cholesterol
< 35 mg/dl (0.9 mmol/L)
Diabetes mellitus

HDL, high-density lipoprotein.

TABLE 15.6. CLASSIFICATION OF PATIENTS ACCORDING TO THEIR SERUM CHOLESTEROL LEVEL

Total Cholesterol Level	Initial Classification
<200 mg/dl (5.2 mmol/L	Desirable blood cholesterol
200–239 mg/dl (5.2–6.2 mmol/L)	Borderline-high blood cholesterol
≥240 mg/dl (6.2 mmol/L)	High blood cholesterol
<35 mg/dl (0.9 mmol/L)	Low high-density lipoprotein cholesterol

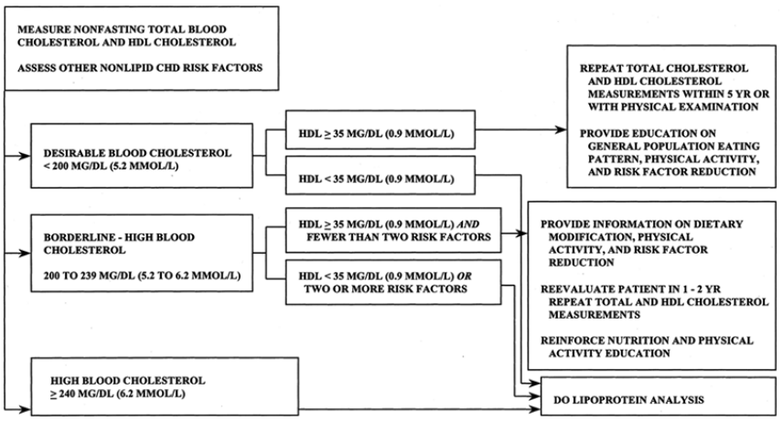


FIGURE 15.15. Algorithm for the initial evaluation of patients who never had a heart attack, stroke, angina, a transient ischemic attack, or peripheral vascular disease. [From Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel (NCEP) on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *JAMA* 1993;269:3015–3023, with permission.]

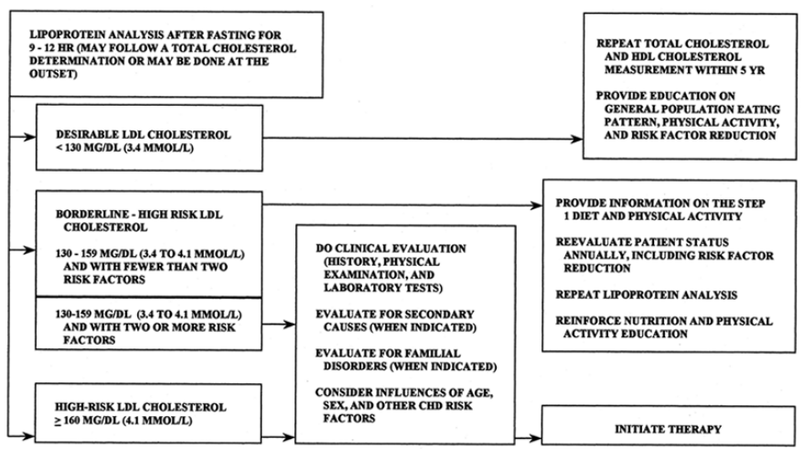


FIGURE 15.16. Algorithm for the assessment of patients without evidence of coronary heart disease. Subsequent classification is based on low-density lipoprotein cholesterol level. [From Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel (NCEP) on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *JAMA* 1993;269:3015–3023, with permission.]

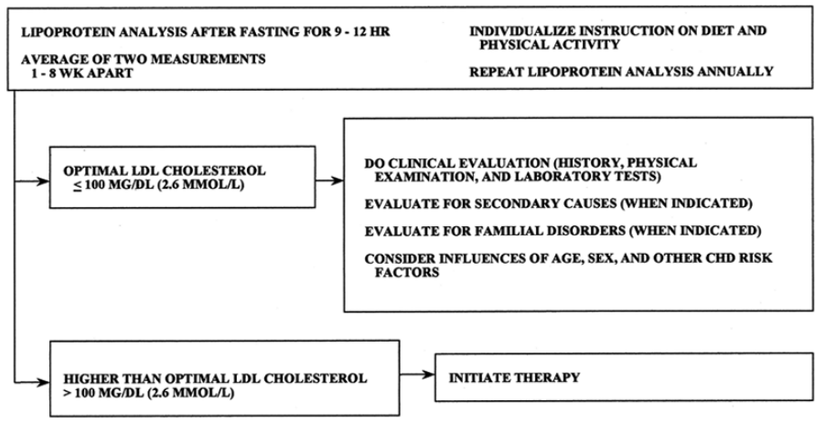


FIGURE 15.17. Algorithm for the laboratory evaluation of patients with existing coronary artery disease. [From Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel (NCEP) on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *JAMA* 1993;269:3015–3023, with permission.]

In the initial classification of the patient relative to total cholesterol level, values for total cholesterol less than 200 mg/dL are considered desirable. It should be noted, however, that somewhat lower values than 200 mg/dL are more desirable because they are associated with a lower risk for atherosclerotic vascular disease. Values between 200 and 239 mg/dL are considered borderline. Values in excess of 240 mg/dL are considered to be high blood cholesterol levels. An HDL cholesterol less than 35 mg/dL is considered low (Table 15.6).

There is one option to eliminate a positive risk factor and that is to have an HDL cholesterol level above 60 mg/dL. Thus, a man older than 45 years of age who has elevated HDL cholesterol, usually as a result of exercise or regular moderate alcohol intake, can eliminate the positive risk factor from his age and have a total risk factor value of zero.

Figure 15.15 shows the algorithm used for assessment of patients who have never had a heart attack, a stroke, angina, a transient ischemic attack, or peripheral arterial vascular disease. These patients are first evaluated with nonfasting total blood cholesterol and HDL cholesterol levels. The LDL cholesterol is not initially measured because it requires fasting and may not be necessary. If the total cholesterol is less than 200 mg/dL and the HDL is greater than 35 mg/dL, the patient is in the most favorable category. Conversely, the patient with a total cholesterol less than 200 mg/dL but an HDL less than 35 mg/dL is further evaluated with an LDL cholesterol. Patients with a borderline value for total cholesterol, an HDL greater than 35 mg/dL, and fewer than two risk factors are educated with regard to the benefits of a dietary modification and physical activity. The patient with a borderline total cholesterol value in the presence of an HDL less than 35 mg/dL must be further studied with an LDL cholesterol value. A blood cholesterol alone greater than 240 mg/dL necessitates an LDL cholesterol determination. Figure 15.16 shows the algorithm after LDL cholesterol quantitation. If the patient has an LDL cholesterol less than 130 mg/dL, general education about dietary fat and physical activity should be provided. If the patient has borderline LDL cholesterol of 130 to 159 mg/dL and fewer than two risk factors, the patient should receive information on diet and physical activity. However, a patient in that same LDL cholesterol range with two or more risk factors merits a thorough evaluation for the cause of the elevated lipid levels, searching for secondary causes or familial disorders to explain the lipid abnormality. A patient with an LDL cholesterol level greater than 160 mg/dL should also be thoroughly evaluated to determine the cause of the hyperlipidemia.

For patients with existing coronary heart disease or other evidence of atherosclerotic vascular disease, LDL cholesterol measurements are obtained in each laboratory evaluation (Fig. 15.17). The goal for LDL cholesterol is much lower in patients with existing disease than for those without clinically evident atherosclerotic disease. The desirable LDL cholesterol level in this group of patients is less than 100 mg/dL. Values higher than 100 mg/dL should prompt an evaluation for the cause of the dyslipidemia and possibly the initiation of lipid-lowering therapy.

Table 15.7 provides a guide for treatment decisions based on LDL cholesterol level and risk factors. For patients without clinically evident atherosclerotic disease who have fewer than two risk factors (group 1), dietary therapy should be initiated at an LDL cholesterol greater than 160 mg/dL. For those without coronary heart disease and with two or more risk factors (group 2), the LDL cholesterol target level for dietary modification is much lower at 130 mg/dL. For those with atherosclerotic vascular disease (group 3), this threshold value is 100 mg/dL for initiation of dietary therapy. Group 1 individuals should receive drug therapy to lower the lipid levels when the LDL cholesterol is greater than or equal to 190 mg/dL, group 2 individuals when the LDL cholesterol is greater than or equal to 160 mg/dL, and group 3 individuals when the LDL cholesterol level is greater than or equal to 130 mg/dL.

TABLE 15.7. GUIDE FOR TREATMENT DECISIONS BASED ON LOW-DENSITY LIPOPROTEIN CHOLESTEROL LEVELS AND RISK FACTORS

Patient Category	Initiation Level for Diet	Low-Density Lipoprotein Goal
Without CHD and with fewer than two risk factors (group 1)	≥160 mg/dl (4.1 mmol/L)	<160 mg/dl (4.1 mmol/L)
Without CHD and with two or more risk factors (group 2)	≥130 mg/dl (3.4 mmol/L)	<130 mg/dl (3.4 mmol/L)
With CHD (group 3)	100 mg/dl (2.6 mmol/L)	>100 mg/dl (2.6 mmol/L)
	For Drug Treatment	
Without CHD and with fewer than two risk factors (group 1)	≥190 mg/dl (4.9 mmol/L)	<160 mg/dl (4.1 mmol/L)
Without CHD and with two or more risk factors (group 2)	≥160 mg/dl (4.1 mmol/L)	<130 mg/dl (3.4 mmol/L)
With CHD (group 3)	≥130 mg/dl (3.4 mmol/L)	>100 mg/dl (2.6 mmol/L)

CHD, Coronary heart disease.

The dietary components of the American Heart Association

Step 1 and Step 2 diets are shown on Table 15.8. The essential features of the diets are their total amount of fat, the percentage of saturated fatty acid, and the amount of cholesterol in milligrams. The Step 1 diet requires that less than 30% of the total calories be represented by fat, that the total cholesterol intake per day is less than 300 mg, and that the total saturated fatty acid content is less than 10% of the total calories. The Step 2 diet is increased in intensity from Step 1 by decreasing the saturated fatty acid percentage to 7% of total calories and decreasing the cholesterol intake to less than 200 mg/day.

TABLE 15.8. DIETARY COMPONENTS OF THE AMERICAN HEART ASSOCIATION STEP 1 AND STEP 2 DIETS

Step 1 diet	
Total fat	<30% of total calories
Saturated fatty acids	<10% of total calories
Polyunsaturated fatty acids	Up to 10% of total calories
Monounsaturated fatty acids	10–15% of total calories
Carbohydrates	50–60% of total calories
Protein	10–20% of total calories
Cholesterol	<300 mg/day
Total calories	To achieve and maintain a desirable weight
Step 2 diet	
Total fat	<30% of total calories
Saturated fatty acids	<7% of total calories
Polyunsaturated fatty acids	Up to 10% of total calories
Monounsaturated fatty acids	10–15% of total calories
Carbohydrates	50–60% of total calories
Protein	10–20% of total calories
Cholesterol	<200 mg/day
Total calories	To achieve and maintain a desirable weight

The following sections of this chapter review the laboratory methods for the measurement of clinically important lipids and lipoproteins in the blood.

MEASUREMENT OF TRIGLYCERIDES

Part of "15 - Lipids and Lipoproteins"

In the reference method, triglycerides are first extracted from the serum with chloroform to remove water-soluble interfering substances, such as glucose and glycerol. The extract is then treated with silicic acid to remove phospholipids, and the triglycerides are hydrolyzed with potassium hydroxide (KOH) to produce glycerol and fatty acids. The glycerol is then oxidized to formaldehyde that forms a colored product with chromotropic acid. The resulting product can be measured at 570 nm.

In the more practical routine methods, triglycerides in plasma or serum are measured enzymatically. Reagents are available that contain all enzymes, buffers, and cofactors needed for the assay.

These reagents are optimized to the analytical system of the manufacturer. In all the different reagent kits, the first step is nearly always a lipase-mediated hydrolysis of triglycerides to glycerol and fatty acids:

$$1\ \text{triglyceride} + 3\ H_2O \xrightarrow{\text{lipase}} 1\ \text{glycerol} + 3\ \text{fatty acid}$$

Glycerol is then phosphorylated by glycerokinase to glycerophosphate, using 1-adenosine 5′-triphosphate. In most methods, glycerophosphate is oxidized by glycerophosphate oxidase to dihydroxyacetone and H_2O_2. The latter is then measured by standard methods, such as the blue color reaction resulting from the peroxidase-mediated condensation of *N*-ethyl-*N*-(2-hydroxy-3-sulfopropyl)-3,5-dimethoxy-4-fluoroanaline and 4-aminoantipyrine. Other methods utilize glycerophosphate dehydrogenase and nicotinamide adenine dinucleotide (NAD) to produce dihydroxyacetone phosphate and NADH (nicotinamide adenine dinucleotide, reduced form). NADH is measured by spectrophotometry at 340 nm or in a reaction catalyzed by diaphorase that produces formazan from a tetrazolium dye using NADH. Formazan can be measured at 500 nm. Enzymatic triglyceride methods do not measure phospholipids or sugars and are linear to a concentration of 700 mg/dL. The methods are very suitable for automation. Free glycerol, however, will contribute to the total amount of triglyceride measured. Glycerol concentrations in fresh serum or plasma generally do not exceed 100 mg/dL. Glycerol levels are increased 50- to 100-fold in hyperglycerolemia, a rare disorder associated with a defective or deficient glycerokinase enzyme. Free glycerol can be measured as described in the assay for triglycerides (see previously), with the omission of the initial lipase step.

MEASUREMENT OF CHOLESTEROL

Part of "15 - Lipids and Lipoproteins"

In the reference method, serum cholesterol esters are hydrolyzed by alcoholic KOH. Cholesterol is then extracted by hexane, and the extract is dried. The dry residue is treated with acetic acid, acetic acid anhydride, and sulfuric acid (Lieberman-Burchard reagent) to develop a color reaction that can be measured at 620 nm. The method has a minimal positive bias compared with the definitive method for cholesterol, which is mass spectrometry. The results can be expressed in either molar or mass concentrations:

Cholesterol concentration (mg/dL) = 38.7 × (mmol/L)

Although the reference method is very accurate, it is time-consuming and not suitable for automation. For routine diagnostic purposes, enzymatic methods have been developed. The first step is an enzymatic hydrolysis of cholesterol esters:

$$\text{Cholesterol ester} + H_2O \xrightarrow{\text{cholesterol ester hydrolase}} \text{cholesterol} + \text{fatty acid}$$

The 3-OH group on cholesterol is then oxidized by cholesterol oxidase to a ketone, with the production of H_2O_2 that can be measured in a peroxidase-catalyzed reaction forming a dye. The method is not highly specific for cholesterol. Beta-hydroxy sterols and plant sterols are also reactive, but their concentration in clinical specimens is negligible compared with that of cholesterol. The commercial reagent kits generally combine all enzymes and other required components, and the formulations vary. Manufacturers usually optimize their methods to one or more automated instruments. When a reagent calibrator-instrument system from a single manufacturer is used, cholesterol measurements should be accurate to within 1% to 3% of the reference values.

HDL CHOLESTEROL

Part of "15 - Lipids and Lipoproteins"

Ultracentrifugation to isolate lipoproteins for subsequent quantitation of lipoprotein components, such as cholesterol, has been the gold standard with which other methods are compared. In density gradient ultracentrifugation, HDL exhibit a density range of 1.063 to 1.21 g/mL. They may contain, however, some Lp(a) that has a slightly overlapping density range at 1.04 to 1.08 g/mL. Lp(a) cholesterol at high concentrations adds to the cholesterol of HDL when isolated by ultracentrifugation. For routine clinical purposes, faster and simpler methods are in common use. HDL cholesterol is usually measured in serum or plasma after the precipitation and removal of apolipoprotein B-100 containing lipoproteins, normally VLDL, IDL, LDL, Lp(a), and CM. A polyanion, reacting with the positively charged groups of lipoproteins and divalent cations that facilitate their reaction have been used for the precipitation reaction. Most of these methods use phosphotungstate, or dextran sulfate in combination with $MgCl_2$. In earlier methods, heparin sulfate and $MnCl_2$ had been widely used, but the residual Mn^{2+} ions were found to interfere with the enzymatic cholesterol methods, resulting in falsely elevated values. Thus, most laboratories avoid precipitants that require manganese. In a variation of the dextran sulfate precipitation method, the precipitating agent is coupled to magnetic particles. The lipoprotein-magnetic precipitating reagent can be rapidly removed without centrifugation, using a magnetic disk. HDL cholesterol can then be measured as described previously. Precipitation in samples with high triglyceride concentrations (>400 mg/dL) may be incomplete and result

in significant error in HDL cholesterol measurement. In these cases, the samples should be subjected to ultracentrifugation and the triglyceride-rich lipoproteins should be removed before precipitation.

An automated homogeneous assay uses sulfated α-cyclodextrin and dextran sulfate and $MgCl_2$ to reduce the availability of cholesterol in LDL, VLDL, and CM for the cholesterol esterase and cholesterol oxidase enzymes. The method eliminates the need for the time-consuming precipitation step. In addition, the cholesterol esterase and cholesterol oxidase enzymes are modified by polyethylene glycol. The modified enzymes exhibit selective catalytic activities toward different lipoprotein fraction, with a preference to react with HDL cholesterol. The resulting H_2O_2 is measured in a peroxidase reaction using 4-aminophenazone to form a purple color dye. Another direct HDL-cholesterol method utilizes antihuman lipoprotein antibodies that bind to LDL, VLDL, and CM. The antibody-antigen complexes formed block enzyme reactions with these lipoproteins, and, therefore, cholesterol esterase and cholesterol oxidase react only with HDL cholesterol. Hydrogen peroxide produced by the enzyme reactions is measured in a peroxidase condensation reaction, yielding a blue color complex. By avoiding the precipitation step and centrifugation, this method is also highly suitable for automated analyzers.

LDL CHOLESTEROL

Part of "15 - Lipids and Lipoproteins"

Methods for the measurement of LDL are based on the assumption that total cholesterol is mainly composed of cholesterol found in VLDL, LDL, and HDL. LDL cholesterol can be calculated or measured directly. To calculate LDL cholesterol, serum or plasma triglyceride, total cholesterol, and HDL cholesterol are determined, and the LDL cholesterol calculated using the equation formulated by Friedewald:

(LDL cholesterol) = (Total cholesterol) - (HDL cholesterol) - (Triglyceride)/5

when concentrations are expressed in milligrams per deciliters. Triglyceride/2.22 should be used if the results are expressed in millimoles per liter. Triglyceride/5 is an estimate of VLDL cholesterol based on the empirically established average ratio of triglyceride to cholesterol in VLDL. The Friedewald equation can be used in most cases to determine LDL cholesterol, but it should not be used for samples with triglyceride concentrations greater than 400 mg/dL. These samples usually also contain CM and CM remnants in which the triglyceride:cholesterol ratios are higher than in VLDL. A different error may occur in patients with type III hyperlipoproteinemia because of the presence of beta-VLDL. Beta-VLDL, which normally does not occur in the blood, contains significantly more cholesterol than normal VLDL and has a triglyceride:cholesterol ratio of only 3:1. Thus, using the Friedewald equation, VLDL cholesterol would be underestimated in these patients. This would also result in an overestimation of the LDL cholesterol.

LDL cholesterol also can be measured directly. The direct assays typically utilize precipitating antibodies specific for apolipoproteins in VLDL, IDL, and HDL. The concentration of LDL cholesterol is measured directly in the supernatant. In a variation of this method, the antibodies are coupled to latex. The antibodies on the beads bind VLDL and HDL, and the complexes are removed by centrifugal filtration. The bead-free filtrate ideally contains only the LDL. The concentration of LDL cholesterol is measured in the solution using an enzymatic method. In other methods, LDL is precipitated with polyvinyl sulfate or heparan sulfate. In these assays, LDL cholesterol in the precipitate is calculated as the difference between the total cholesterol concentration and the cholesterol in the supernatant. Most studies indicate that direct methods are useful and clinically accurate for the measurement of LDL cholesterol concentration, but further validation of some of these assays is needed. An additional limitation of the precipitation methods is that they cannot be used on frozen plasma or serum.

SEPARATION OF LIPOPROTEINS WITH CHOLESTEROL DETERMINATION OF ISOLATED FRACTIONS

Part of "15 - Lipids and Lipoproteins"

Total lipoprotein concentrations can be measured directly using nuclear magnetic resonance spectroscopy. This technique measures the chemical shifts of the methyl and methylene groups of fatty acids in lipoproteins from the VLDL, LDL, and HDL fractions. The results can be converted to lipoprotein cholesterol concentrations using the assumed ratios of cholesterol content in each lipoprotein class. The method is fast but requires expensive instrumentation and considerable user experience.

Ultracentrifugation in vertical rotors using density gradients and measurement of cholesterol in the fractions is another method for the measurement of total lipoprotein concentrations. This method also requires specialized instrumentation and is time-consuming and not suitable for automatization.

Gradient gel electrophoresis has also been used for the fractionation of lipoproteins. Gradient gel electrophoresis provides a better resolution than traditional agarose electrophoresis. The gels can be scanned and the areas under the peaks can be integrated. The cholesterol concentrations can be calculated using the assumed cholesterol content of each lipoprotein fraction. Insufficient resolution of the different fractions has been a major limitation of this method.

These methods have not been validated to the same degree as the ultracentrifugation-polyanion precipitation technique and the Friedewald calculation. Additional studies are required before these newer methods gain wide acceptance.

ISOLATION AND QUANTITATION OF LIPOPROTEINS

Part of "15 - Lipids and Lipoproteins"

Several methods are available to separate lipoproteins for quantitative analysis. HDL can be isolated and cholesterol can be determined using the dextran-sulfate-$MgCl_2$ or the heparin-$MnCl_2$ methods. These methods are similar to the reference methods used for HDL cholesterol measurement (refer to section on HDL cholesterol). Ultracentrifugation can be used to separate CM, VLDL, and beta-VLDL from plasma. After centrifugation

at 105,000*g* for at least 18 hours, a floating white layer composed of triglyceride-rich lipoproteins is separated. A water-clear zone precedes a colored zone near the bottom of the tube, containing lipoproteins with a density greater than 1.06 g/mL. The remaining, denser plasma proteins are located at the bottom of the tube. Lp(a) has a mobility similar to that of VLDL. Using a tube slicer and a syringe equipped with a blunt-end needle, the top and bottom lipoprotein layers can be separated and their respective cholesterol content determined.

Lipoprotein electrophoresis, now considered outdated, had been widely used. The method typically uses agarose as the separation medium. Plasma and the two main ultracentrifuge fractions can be analyzed. CM, the largest, slowest particles, remain at the origin. LDL (beta-lipoprotein) is the fastest fraction; it migrates with the ß-globulins. VLDL (pre-ß-lipoprotein) migrates with the pre-ß-globulins, whereas HDL (α-lipoprotein) with the α-globulins. Lp(a), when present in detectable amounts, migrates with VLDL.

LDL and HDL subclasses can be separated by gradient gel electrophoresis or analyzed by nuclear magnetic resonance. Although still somewhat controversial, these methods allow the identification and quantitation of the more atherogenic, small, dense LDL fraction and the various HDL fractions, which appear to have different antiatherogenic potency.

APOLIPOPROTEINS

Part of "15 - Lipids and Lipoproteins"

Apolipoproteins are measured by various immunoassays, including enzyme-linked immunoassay (ELISA), radioimmunoassay (RIA), immunonephelometry, immunoturbidimetry, and radial immunodiffusion.

Apolipoproteins A-I and B-100 are present in relatively higher concentrations and can be conveniently measured by immunoturbidimetry or immunonephelometry. Most laboratories in the United States use one of these two methods for measuring apolipoproteins A-I and B-100. Radioimmunoassay and radial immunodiffusion have not been extensively used because of the inconvenience of handling radioactive material and the relative imprecision, respectively. Apolipoprotein A-I is smaller and more soluble in water than B; there are fewer technical problems with the measurement of this apoplipoprotein A-I. Apolipoprotein A-I is present as the major protein constituent of HDL. Apolipoprotein A-I is also present, albeit in much smaller amount, in CM. In general, apolipoprotein A-I levels are parallel with the HDL levels. Apolipoprotein B-100 is present in LDL, IDL, VLDL, and Lp(a), the particles of which are considerably different in size and composition. Antibodies that are used for the measurement of apolipoprotein B-100 should equally recognize this apolipoprotein in the different lipoprotein classes. Nonionic detergents, such as Tween-80 or Tween-20, have been used to disrupt the various lipoprotein particles and unmask all antigenic determinants of the apoprotein molecules and make them available for the antibodies. The antibodies should be able to react with their respective epitope, regardless of which class was its source.

Because apolipoproteins C-I and C-II are present at lower concentrations, the more sensitive ELISA and RIA are the preferred methods for their measurement.

Lp(a) is similar to LDL in both lipid composition and the presence of lipoprotein B-100. However, it contains an additional apolipoprotein, apolipoprotein(a), which is covalently linked to apolipoprotein B-100 by a single disulfide bond and likely responsible for the characteristics attributed to the Lp(a) particle. The apolipoprotein(a)–to–B-100 ratio is approximately 1:1. Apolipoprotein(a) features a repetitive structure similar to that of plasminogen; therefore, the development of a reliable immunoassay is very difficult. Thus, there is no accepted reference method for apolipoprotein(a) at present. Monoclonal antibodies have been used for developing enzyme immunoassays. Oxidative modification of Lp(a) and removal of carbohydrate moieties have been attempted to increase specificity. In addition, sandwich assays have also been developed using anti-apolipoprotein(a) and anti-apolipoprotein-B antibodies. Because the expression of epitopes in apolipoprotein B in LDL and Lp(a) are different, the apolipoprotein-B antibodies used in this assay should be carefully selected and characterized.

SUMMARY

Part of "15 - Lipids and Lipoproteins"

This chapter reviews the structure of lipids and lipoproteins, the participation of these entities in the development of atherosclerosis, the use of lipid and lipoprotein test results in the management of patients with and without atherosclerotic vascular disease, and the clinical laboratory methods for the measurement of clinically important lipids and lipoproteins.

DEDICATION

Part of "15 - Lipids and Lipoproteins"

This work is dedicated to the memory of the first author's father, Aladír Ábel, M.D. (1927–1999), former head physician of the Clinical Laboratories, Railway Hospital, Szolnok, Hungary.

SUGGESTED READING

Fatty Acids

Laposata M. Fatty acids: biochemistry to clinical significance. *Am J Clin Pathol* 1995;104:172–179.

Tato F. Trans-fatty acids in the diet: a coronary risk factor. *Eur J Med Res* 1995/96;1:118–122.

Hulshof KFAM, van Erp-Baart MA, Anttolainen M, et al. Intake of fatty acids in Western Europe with emphasis on trans fatty acids: the TRANSFAIR study. *Eur J Clin Nutr* 1999;53:143–157.

Spector AA. Essentiality of fatty acids. *Lipids* 1999;34:S1–S3.

Simopoulos AP. Essential fatty acids in health and chronic disease. *Am J Clin Nutr* 1999;70(Suppl 3):560S–569S.

Pepping J. Omega-3 essential fatty acids. *Am J Health Syst Pharm* 1999;65:719–724.

Triglycerides

Brewer B. Hypertriglyceridemia: changes in the plasma lipoproteins associated with an increased risk of cardiovascular disease. *Am J Cardiol* 1999;83:3F–12F.

Cholesterol and Cholesterol Esters

Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel (NCEP) on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *JAMA* 1993;269:3015–3023.

Ordovas JM, Schaefer EJ. Genes, variation of cholesterol and fat intake and serum lipids. *Curr Opin Lipidol* 1999;10:15–22.

Tall AR. An overview of reverse cholesterol transport. *Eur Heart J* 1990;19:A31–A35.

Lipoproteins, Apolipoproteins, and Dyslipoproteinemias

Ginsberg HN. Lipoprotein physiology. *Endocrinol Metab Clin North Am* 1998;27:503–520.

Olson RE. Discovery of lipoproteins, their role in fat transport and their significance as risk factors. *J Nutr* 1998;128:439S–443S.

Beisiegel U. Lipoprotein metabolism. *Eur Heart J* 1998;19(Suppl A): A20–A23.

Schaefer EJ, Brousseau ME. Diet, lipoproteins, and coronary heart disease. *Endocrinol Metab Clin North Am* 1998;27:711–733.

Hoeg JM. Lipoproteins and atherogenesis. *Endocrinol Metab Clin North Am* 1998;27:569–584.

Davignon J, Genest J. Genetics of lipoprotein disorders. *Endocrinol Metab Clin North Am* 1998;27:521–550.

Packard CJ, Shepherd J. Lipoprotein heterogeneity and apolipoprotein B metabolism. *Arterioscler Thromb Vasc Biol* 1997;17:3542–3556.

Jong MC, Hofker MH, Havekes LM. Role of ApoCs in lipoprotein metabolism: functional differences between ApoC1, ApoC2, and ApoC3. *Arterioscler Thromb Vasc Biol* 1999;19:472–484.

Kamboh MI, Mehdi H. Genetics of apolipoprotein H (β_2-glycoprotein I) and anionic phospholipid binding. *Lupus* 1998;7(Suppl 2):S10–S13.

Utermann G. The mysteries of lipoprotein(a). *Science* 1989;246:904–910.

Marcovina SM, Koschinsky ML. Lipoprotein(a) as a risk factor for coronary artery disease. *Am J Cardiol* 1998;82:57U–66U.

Puckey L, Knight B. Dietary and genetic interactions in the regulation of plasma lipoprotein(a). *Curr Opin Lipidol* 1999;10:35–40.

Scanu AM. Structural and functional polymorphism of lipoprotein(a): biological and clinical implications. *Clin Chem* 1995;41:170–172.

Standardization and clinical measurement of lipoprotein(a). *Clin Chem Lab Med* 1998;36:5–16.

Grundy SM. *Cholesterol and atherosclerosis. Diagnosis and treatment*. Philadelphia: Lippincott, 1990.

Summary of the second report of the National Cholesterol Education Program (NCEP) Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel II). *JAMA* 1993;269:3015–3023.

Gotto AM. *American Heart Association's Cholesterol Education Program. Clinician's manual*. London: Science Press, 1991.

Cardiovascular Risk Factors and Management

Ridker PM. Evaluating novel cardiovascular risk factors: Can we better predict heart attacks? *Ann Intern Med* 1999;130:933–937.

Wood D. European and American recommendations for coronary heart disease prevention. *Eur Heart J* 1998;19(Suppl A):A12–A19.

Cleeman JI. Detection and evaluation of dyslipoproteinemia. *Endocrinol Metab Clin North Am* 1998;27:597–611.

Franklin FA, Dashti N, Franklin CC. Evaluation and management of dyslipoproteinemia in children. *Endocrinol Metab Clin North Am* 1998;27:641–654.

Grundy SM. The role of cholesterol management in coronary disease risk reduction in elderly patients. *Endocrinol Metab Clin North Am* 1998;27:655–675.

Donahoo WT, Kosmiski LA, Eckel RH. Drugs causing dyslipoproteinemia. *Endocrinol Metab Clin North Am* 1998;27:677–697.

16
Endocrine Function and Carbohydrates

Bruce F. Bower

Robert E. Moore

Hormones are a group of diverse biomolecules that are carried by the circulation and interact with specific cells and receptors to produce a precise cellular response. These hormones may be of low molecular weight such as steroids or thyronines or they can be more structurally complex peptides such as thyroid-stimulating hormone (TSH) and growth hormone (GH). Each hormone is under the influence of stimulatory and inhibitory factors that act to provide the exquisite control of biological processes the hallmark of endocrine systems.

Recognizing the extremes in molecular size, it is not difficult to anticipate a similar extreme in physical and chemical characteristics. Because the circulation is essentially a water-based system, the insolubility of organic compounds with their extensive ring structures has been overcome by binding these small molecules to large soluble proteins called binding or transport proteins. This serves several purposes. First, the secreting cells have few limitations in terms of distance from target tissue. Second, this stabilizes the concentration of circulating hormone and is a reserve of inactive hormone that can be readily converted to the active form. The active hormone is usually an unbound or free form that is in equilibrium with the bound form. This relationship between bound and unbound hormone is under biological control. In other words, the concentration of free hormone can be rigorously controlled independent of the concentration of bound hormone.

In the case of peptides, which are generally soluble and have sufficient size to make instantaneous synthesis and secretion improbable, the existence of hormone precursors or prohormones provides the hormone reserve. In this instance, the peptide is synthesized in a form that has little, if any, biological or pharmacologic activity. The precursor hormone is stored and can subsequently undergo enzymatic alteration to be converted to the active peptide. An example of this mechanism is the conversion of proinsulin to insulin. Proinsulin has very little (approximately 10% of insulin) pharmacologic activity, but after a small peptide segment, called C-peptide, has been cleaved, the insulin portion is 100% active. Analogous to the bound-unbound state of small hormones, the active-inactive form of proteins and peptides gives rise to a point of control.

Endocrine control mechanisms take several forms. Feedback control is one of these mechanisms. Depending on the nature of the response of the controlled hormone in relation to the controlling hormone, this process is referred to as either positive or negative feedback. When the secretion of the controlled hormone is stimulated by the increased concentration of a different hormone, the effect is called positive feedback. A common example of this type of control is the observed increase in LH concentration after the increased concentration of estrogens in the late follicular phase of the menstrual cycle. This is an example of positive feedback.

Negative feedback operates in the opposite manner from positive feedback. A hormone under the influence of a stimulatory or releasing hormone will increase in concentration until it is high enough to cause the secretion of that releasing hormone to diminish or stop. Pituitary TSH will stimulate the thyroid gland to secrete thyroid hormone until the concentration of active thyroid hormone reaches physiologic levels. Then the thyroid hormone will inhibit further release of TSH from the pituitary. *In vivo*, this inhibition process can be more complex than a straight linear relationship between the two hormones. Sometimes other hormones do the actual inhibition. In the case of thyrotropin (TSH) secretion, triiodothyronine (T_3) inhibits TSH in the thyrotrope, although the thyroxine (T_4) level in the circulation correlates better with the TSH level. This is the result of an intracellular conversion of T_4 to T_3 and T_3 is the potent inhibitor.

The sequences of events that comprise the feedback system are termed loops. If the hormone acts directly on the cells secreting the hormone either itself or its releasing hormone, then the process is called a short-loop feedback system. However, if the hormone acts on the target tissue and a second hormone is released that returns through the circulation to inhibit the original hormone or its releasing hormone, the process is called long-loop feedback. These feedback loops are illustrated in Fig. 16.1.

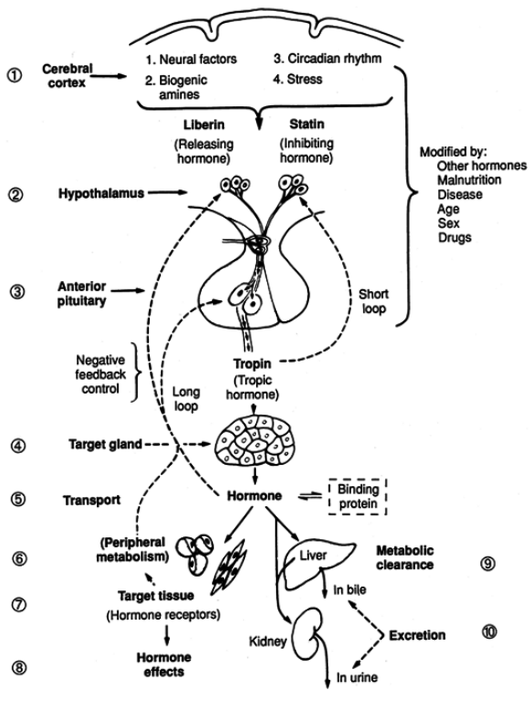


FIGURE 16.1. Stimulation, production, control, and metabolism of hormones. (1) Stimulation of the cerebral cortex resulting in (2) stimulation of the hypothalamus to release tropic hormones from the anterior pituitary (3). The hormone circulates to a target gland (4), where a second hormone can be produced. This hormone is bound to transport or binding proteins (5) and can undergo peripheral metabolism (6) or bind to receptors in target tissues (7), which results in the specific hormone effect (8). Hormones can then be cleared metabolically (9) or excreted from the system (10). If the tropin acts on the hypothalamus, this is an example of the short-loop feedback, whereas if the hormone or peripheral metabolic products circulate to act on the hypothalamus, the control is termed long-loop negative feedback. (From Gornall AG, Luxton AW. Endocrine disorders. In: Gornall AG, ed. *Applied biochemistry of clinical disorders*, 2nd ed. Philadelphia: Lippincott, 1986, with permission.)

A common and diagnostically useful characteristic of hormones is rhythmicity. Hormones are generally secreted on a periodic cycle that is referred to as circadian if the cycle is approximately 24 hours. Cycles shorter are called ultradian, and those that are longer are referred to as infradian. In addition to these cyclic secretions, there are pulsatile releases of hormones triggered by neurologic, environmental, or biochemical events. These secretion patterns can be evaluated and sometimes initiated

by the clinician under carefully controlled conditions to elucidate the underlying pathology. Often the observation that the natural rhythm of secretion has been lost is enough to direct the diagnostic workup.

The preceding description is concerned with normal physiologic events. One abnormal condition is ectopic hormone production. Ectopic production describes a disorder in which a hormone is being produced at a site or by a tissue not normally associated with the production of that hormone. These tissues are generally neoplastic and can produce hormones that can be biologically and structurally identical to, or variants of, normal hormones.

Endocrinology addresses the various clinical disorders that arise from circulating hormone concentration excesses or deficiencies, regulatory system failures, extraneous sources of hormone production, molecular variants, and receptor abnormalities. The clinical laboratory plays a crucial support role in the evaluation of the patient. The laboratory must be aware of the possible approaches to measuring hormones, their metabolites, and receptor function so a clear understanding of the biological processes can be deduced by the clinician. In the following sections, the biochemistry, physiology, laboratory measurement, and clinical interpretation of the more commonly encountered endocrine disorders are reviewed.

RECEPTORS

Part of "16 - Endocrine Function and Carbohydrates"

The circulating concentration of hormones is in the range of 10^{-7} to 10^{-12} M, and of the organs and tissues these hormones come into contact with, only a few are hormonally responsive. The mechanism of hormonal sensitivity and specificity is attributed to receptors. Receptors are cell-associated proteins that are unique in their ability to bind selected molecules, then translate that binding into subsequent cellular processes. These receptors can be located in the cellular membrane where the binding sites are on the external surface of the cell, with the remainder of the protein extending through the membrane to the cytosolic surface. These receptors are referred to as plasma membrane receptors. Plasma membrane receptors move about within the confines of the cellular membrane and interact with molecules that are unable to diffuse through the membrane. Other receptors are intracellular and bind with hormones that are able to diffuse through the plasma membrane. Binding to the intracellular receptor

allows the hormone to be translocated to the nucleus. To regulate cell function, all the elements, hormone, receptors, and the receptor-hormone complex must become activated. Receptors bind other molecules in addition to the specific hormone. The binding of these cross-reacting molecules will result in little or no physiologic response. Molecules that bind and have some activity are referred to as agonists, whereas those that bind and prevent or inhibit hormone activity are referred to as antagonists.

The physical binding of hormones to receptors is dependent on several types of weak chemical forces. These are ionic bonds, hydrogen bonds, van der Waals forces, and hydrophobic-hydrophilic repulsion. Because these forces act over relatively short molecular distances, it is not surprising that there is a stereochemical requirement to maximize the strength of binding. The better the fit between the receptor and its hormone, the greater the affinity of the hormone–receptor complex.

Experimental evidence indicates that cells contain “spare receptors,” i.e., more receptors than are required for maximal cellular response (1). Hormone-receptor binding follows the laws of mass action. Therefore, extra receptors allow the cell to be responsive to lower concentrations of hormone than if the number of receptors were limited. All the cellular receptors are functional and equivalent. At times when hormone concentration is high, these spare receptors bind hormone, but because the cell is already at maximum stimulation, no additional effect is produced. At other times when hormone concentrations are low, small changes in hormone concentration will be effective, because only a small percentage of total available receptors must be occupied to cause maximum stimulation. This is also a point of control. By regulating the number of receptors available, the cell can vary its sensitivity to the circulating concentration of hormone. This phenomenon of reducing receptors on the cell surface when hormone concentrations are high and increasing them when concentrations are low is referred to as downregulation in the first case and upregulation in the latter. The ability of cells to respond to hormone concentrations in this manner may explain the efficacy of pulsatile secretion of some hormones. In pulsatile secretion, hormone is released in short bursts over a period of time. If the receptor population on a cell were fixed, then cellular sensitivity to hormone concentration would be fixed. That is, at some critical concentration the cell would have a maximum response. Higher concentrations would not increase the response and lower concentrations might not evoke a sufficient response. By the cell modulating the receptor concentration, a maximum sensitivity is maintained without a requirement for high hormone levels. One of the unique characteristics of endocrine systems is the graded response. Another process by which the cell can control receptor-binding sites on the membrane surface is to remove the entire hormone–receptor complex. After receptor-hormone binding has taken place, the complex can be removed from the cell membrane and degraded inside the cell with either recycling of the receptor or degradation of the receptor and ligand. This internalization of hormone–receptor complexes renders them inactive and reduces the number of receptors available on the cell wall (2). Not all the mechanisms of hormone–receptor deactivation have been worked out, and there is probably more than one mechanism involved. It is thought that some of the changes in receptor number and function on cell surfaces are owing to the normal turnover of receptor protein. This description of hormone–membrane receptor interactions is generic, and each of the hormone types, e.g., steroids, thyroid hormones, peptides, and catecholamines, alters cellular function in a slightly different manner.

Steroid and thyroid hormones circulate bound to much larger proteins and are in equilibrium with a very low concentration of free hormone. Because of the hydrophobic nature of these hormones, they are able to penetrate the lipid membrane of the cell by simple diffusion. After hormone binding to a cytosolic receptor and translocation to the nucleus, there is a conformational change that permits the activated hormone–receptor complex to bind to a specific region on DNA to initiate new mRNA and ultimately new protein synthesis (3). Figure 16.2 illustrates this process.

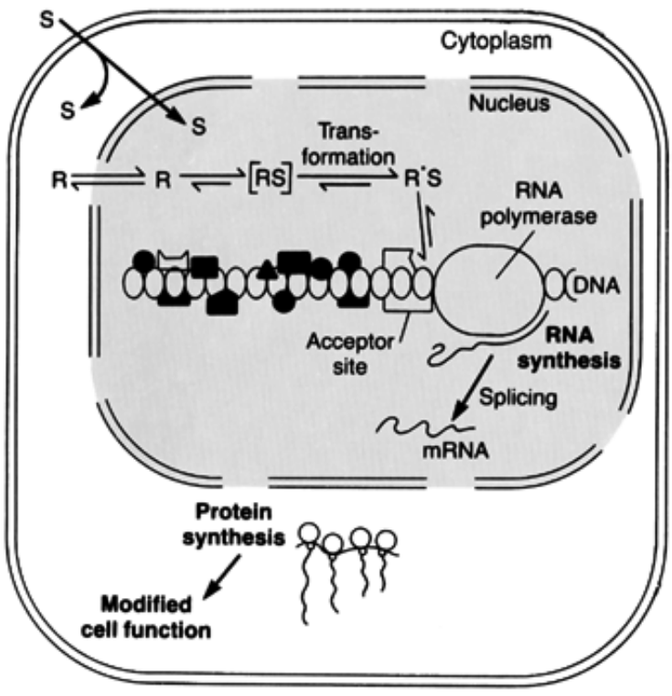


FIGURE 16.2. Steroid hormone initiation of protein synthesis. Steroid diffuses into the cell through diffusion or active transport mechanism and binds with a receptor in the nucleus to form a steroid–receptor complex. This activated complex then activates DNA, which in turn activates the RNA synthesis to eventually lead to protein synthesis and cellular response. S, steroid; R, receptor; RS, steroid-receptor complex. (From Walters MR. Steroid hormone receptors and the nucleus. *Endocrinol Rev* 1985;6:512, with permission.)

Large peptides, amine hormones, and sensory stimuli use a different mechanism for transmembrane signaling (4). There are at least three elements on or in the cell membrane that interact to affect cellular function (3,5). The first is the externally oriented receptor-binding site that binds the peptide or amine hormone (5). The second is a group of heterotrimeric guanine-nucleotide-binding proteins called G proteins. The third element is made up of the numerous effector systems that regulate cellular activity. The process starts with the binding of a hormone to a receptor and forming a receptor–hormone complex. This activated receptor–hormone complex is able to move about in the lipid bilayer of the cellular membrane and couple with a G protein. G proteins act as intramembrane signal transducers that activate cellular effectors. These protein complexes are composed

of three subunits: α, ß, and γ (5). Recent information indicates there are genes for a minimum of 16 α subunits, five ß subunits and 11 γ subunits (6). The possible combinations of heterotrimere may be more than 1,000 (6). In the inactive state the α, subunit binds a molecule of guanosine diphosphate (GDP). Immediately after binding with the activated receptor complex, the G protein exchanges a guanosine triphosphate (GTP) for the GDP and the G protein α subunit; in addition, GTP dissociates from the βγ subunit complex (7). Both the α-GTP complex and ß-γ complex can activate effector proteins (7). An activated receptor is capable of activating multiple G proteins and, in turn, the activated G protein GTP or βγ subunit can activate multiple effectors. This series of events is responsible for the amplification effect of some hormones (8). The effector protein can be an enzyme such as adenylate cyclase or phospholipase C, or it can be an ion channel regulator. This process continues until the GTP on the subunit is dephosphorylated back to GDP and the α subunit combines with the ß-γ subunit generating an inactive G protein. If the effector were adenylate cyclase, the generation of cyclic AMP would halt, and the intracellular level of cyclic AMP would be reduced by cyclic AMP phosphodiesterase. The hormone effect on the cell is terminated. Other cyclic nucleotides may be generated by analogous systems. There are other G proteins that inhibit effector proteins, giving the cell both stimulatory and inhibitory pathways activated by hormone binding to receptors (6,9).

Cyclic AMP has been referred to as a second messenger because it is the intracellular agent responsible for the direct alteration of cellular function (10). However, other intracellular ions and molecules qualify as second messengers, ionic calcium (Ca^{2+}) is an example (11). The intracellular free calcium concentration is very rigidly controlled through a series of transport processes or binding to intracellular structures (11). An important binding protein for intracellular calcium is calmodulin. The calcium–calmodulin complex represents another activated complex that can stimulate intracellular enzymes.

Calcium ion can be released in the cell through the action of a membrane-bound phosphodiesterase (phospholipase C), which generates the second messengers inositol 1,4,5-triphosphate (IP_3) and diacylglycerol (DAG) (4,11). IP_3 can liberate calcium from bound stores within the cell. DAG and the released calcium can activate protein kinase C, which will ultimately phosphorylate proteins. These phosphorylated proteins, generally enzymes, can either activate or inhibit further cellular processes.

Larger molecules such as insulin and epidermal growth factor (EGF) bind to the extracellular domain of a transmembrane spanning receptor, which is linked through a transmembrane sequence to an intracellular enzymatic domain often possessing tyrosine and serine kinase activity (12). Autophosphorylation of the enzymatic site leads to serine and tyrosine phosphorylation of a variety of intracellular proteins, including enzymes, which results in intracellular response.

There is an additional small group of hormones that appear to affect phosphorylation through kinases without the aid of a second messenger (10). Insulin and EGF activate receptor-localized tyrosine kinases that phosphorylate tyrosine residues on a variety of intracellular enzymes after hormone-receptor binding on the cell surface. This receptor complex spans the cell membrane, with the receptor-binding site presented to the external environment and the tyrosine kinase on the inside of the cell.

The awareness of specific tissue receptors is clinically relevant. Hormone abnormalities can arise from failures in hormone production and from failure caused by receptor abnormalities.

HYPOTHALAMIC-PITUITARY AXIS

Part of "16 - Endocrine Function and Carbohydrates"

The hypothalamic-pituitary axis can be thought of as an integrated neuroendocrine transducer in which central neural input, neuropeptide release, and both positive and negative long-loop systemic hormonal feedback regulate the production and secretion of no fewer than six anterior pituitary hormones (13). The control of anterior pituitary function by hypothalamic peptide–releasing hormones is mediated via the hypophyseal portal vascular plexus, which links the hypothalamus and the hypothalamic-releasing peptides with their corresponding anterior pituitary hormones. Table 16.1 lists these releasing peptides and their associated pituitary hormones. These central stimulating effects are modified by feedback loops from the peripheral endocrine glands in which a product hormone is secreted—adrenal gland: cortisol; thyroid: T_4 and T_3; ovary and testis: estrogen, testosterone, and the gonadal peptide inhibin; and GH: somatomedin C/insulinlike growth factor 1 (SM-C/IGF-1).

TABLE 16.1. RELEASING PEPTIDES AND HORMONES

Hypothalamic Releasing Peptides	Associated Hormone
Growth hormone–releasing hormone	Growth hormone
Somatostatin	Growth hormone
Thyrotropin-releasing hormone	Thyrotropin
Gonadotropin-releasing hormone	Follicle stimulating hormone and luteinizing hormone
Corticotropin-releasing hormone	Corticotropin

Growth hormone releasing hormone and somatostatin peptides regulate pituitary growth hormone secretion; thyrotropin releasing hormone regulates pituitary thyrotropin secretion; gonadotropin (luteinizing hormone) releasing hormone regulates both pituitary follicle-stimulating hormone and luteinizing hormone; corticotropin-releasing hormone regulates corticotropin.

An essential feature of the hypothalamic-pituitary axis is episodic secretion of hypothalamic hormones with resultant episodic secretion of pituitary hormones. Pulsatile secretion is defined by pulse, frequency, pulse amplitude, and resultant quantity of the hormone secreted. Episodic secretion complicates endocrine analysis by introducing temporal variations in the serum concentration of pituitary hormone but appears necessary to avoid downregulation of the hormone receptors through which these peptides exert their biological influence. An additional complexity is the marked diurnal variations in pituitary peptide secretion, particularly of corticotropin (ACTH) and GH, in which sleep-related nocturnal secretion predominates with resultant high early-morning values and much lower late-afternoon and evening values. Documentation of the loss of this physiologic diurnal variation may be useful clinically and requires appropriate morning and evening reference intervals for hormones such as ACTH and cortisol.

ANTERIOR PITUITARY

Part of "16 - Endocrine Function and Carbohydrates"

The pituitary gland (hypophysis) weighs approximately 0.5 g and is located in the sella turcica, a depression of the sphenoid bone. The anterior pituitary or adenohypophysis develops from ectodermal cells of the central nervous system and neuroectodermal cells from the neural crest, to eventually comprise 75% to 80% of the total pituitary mass. The anterior pituitary is further subdivided into a pars tuberalis and a pars distalis, the latter containing the cells that synthesize and secrete all the anterior pituitary hormones. The major anatomical features of the pituitary are shown in Fig. 16.3.

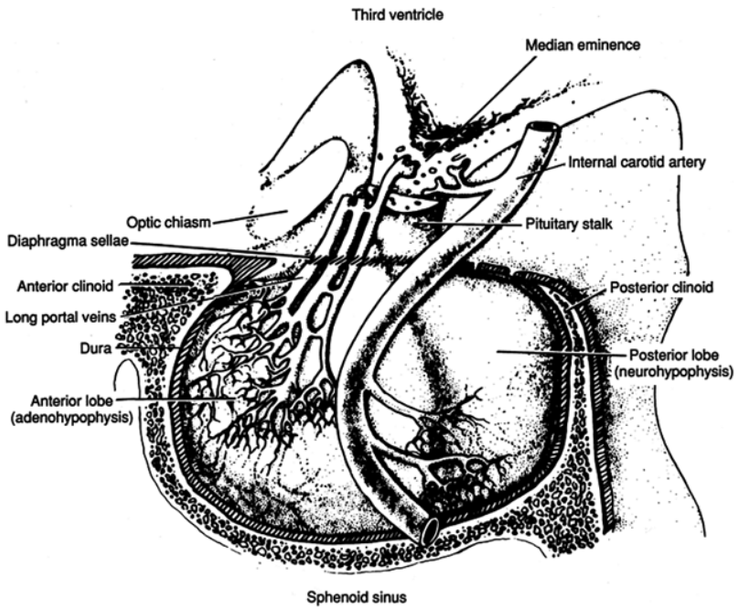


FIGURE 16.3. Major anatomical features of the pituitary gland. (From Frohman LA. *Diseases of the anterior pituitary. Endocrinology and metabolism*. New York: McGraw-Hill, 1981:151–231, with permission.)

The hypothalamus is connected to the pituitary through the pituitary stalk. Blood and hormones from the hypothalamus communicate with the anterior pituitary through the portal venous plexus. The venous circulation of the brain and the pituitary is illustrated in Fig. 16.4.

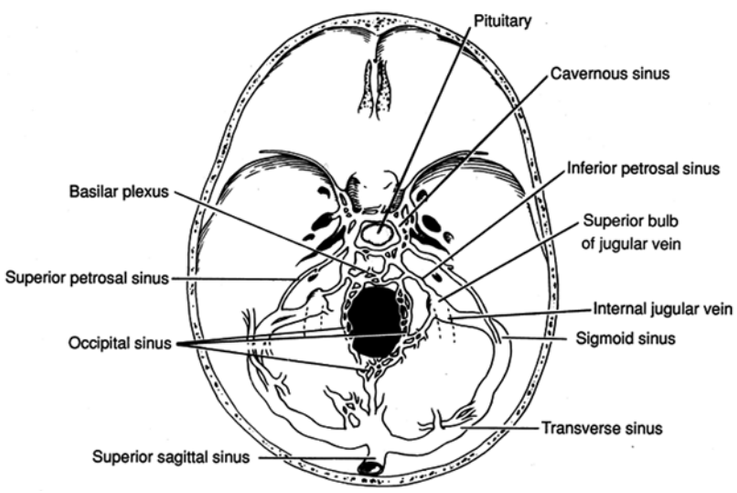


FIGURE 16.4. Venous system of the pituitary gland communicating with the systemic circulation. (From Findling JW, Aron DC, Tyrrell JB, et al. Selective venous sampling for ACTH in Cushing's syndrome: differentiation between Cushing's disease and the ectopic ACTH syndrome. *Ann Intern Med* 1981;94:647–652, with permission.)

Growth Hormone

GH, a single-chain peptide composed of 191 amino acids, is the most abundant of the anterior pituitary hormones. It is released in a pulsatile manner from the somatotropic cells and is under the direct control of two other hormones, growth hormone–releasing hormone (GHRH), which stimulates GH secretion, and somatostatin (SS), which inhibits secretion (14). GHRH and SS are controlled and secreted into the hypophyseal portal circulation by the neuropeptide cells of the hypothalamus.

The association of GH with skeletal growth has been recognized for some time. However, GH is not a true growth factor because its effect on growth is mediated through the somatomedins, a group of peptides synthesized by liver and cartilage.

SM-C, also referred to as IGF-1, is one of these somatomedins and is a true cellular growth factor acting on skeletal components to produce linear growth.

There is a general rhythmicity to the secretion of GH, with the greatest secretion taking place during deep sleep. However, this circadian character has superimposed on it evoked responses to physiologic and biochemical stresses.

Growth Hormone–Releasing Hormone

GHRH is a small peptide, present in three molecular forms having 37, 40, and 44 amino acids. All three of these molecular forms have a biologically active segment in the first 29 amino acids of the sequence. The two larger forms of GHRH are more potent than the smaller form in the human. It has been suggested that the different forms arise from a posttranslational enzyme cleavage. The active GHRH reaches the pituitary by way of the portal circulation.

Somatostatin

Somatostatin was actually observed before GHRH when an extract of hypothalamus was noted to inhibit the release of GH. Occurring as a 28-amino acid precursor that is enzymatically cleaved to the biologically active 14-amino acid form, somatostatin has an inhibitory effect on several hormones in addition to GH, including insulin, glucagon, foregut peptides, and thyrotropin. GH inhibition is considered to take place at the pituitary by the binding of somatostatin to membrane receptors, which causes both a decrease in cyclic AMP concentration and a reduction in cyclic AMP effect. Somatostatin also inhibits human TSH at the pituitary level, as well as several other hormones throughout the body.

GH Abnormalities

GH Excess

Excess GH is the cause for two pathologic conditions, each of which is associated with a particular age group. In children, excess GH leads to giantism, whereas in adults it produces acromegaly. The acromegalic has a distinctive physical appearance associated with acral growth of flat bones and large, broad features such as hands, feet, and nose. Additional facial characteristics are impressive supraorbital ridges and prognathism. The skin of these individuals is thick and oily, and physical examination will reveal enlarged organs. Laboratory findings supporting the diagnosis include an elevated and nonsuppressible serum GH and an elevated serum SM-C/IGF-1. The most convenient method to demonstrate GH nonsuppressibility is glucose administration. This is usually accomplished by giving an oral glucose load of 1.75 g/kg, to as much as 100 g, and then measuring GH levels at 30-minute intervals for as long as 2 hours. Table 16.2 lists the more common stimulation and suppression tests and their expected reference intervals.

TABLE 16.2. GROWTH HORMONE REFERENCE INTERVALS

Condition	Preparation	Collection Time	Results	Increased Levels	Decreased Levels
Normal	Fasting	Early AM		Giantism, acromegaly, ectopic secretion, stress, exercise, prolonged fasting	Growth hormone deficiency Hypopituitarism Adrenal cortical hyperfunction
Child			1–10 µg/L		
Adult					
Female			< 10 µg/L		
Male			<2 µg/L		
L-Dopa	500 mg adult	0, 30, 60, 90, 120	10 µg/L or		No response if glucose 120 mg/dl; poor response in hypopituitarism
stimulation	10 mg/kg child	180 min	5 µg/L above baseline		
Arginine	Child 0.5 g/kg	0, 30, 60 min	Fasting <5 µg/L		Hypopituitarism
stimulation	Adult 30 g/kg 30 min		Peak 10 µg/L at 30–60 min		no response
Insulin	0.1 U/kg adult	0, 15, 30, 45, 60	10 µg/L or		Poor response in hypopituitarism, hypothyroidism
stimulation	0.05 U/kg child IV	90 min	5 µg/L above baseline Glucose level ≤50% of baseline		
Exercise (vigorous)	20 min		5 µg/L		
Glucose suppression	75 g or 1.75 g/kg after fast	0, 30, 60, 90, 120 min	Decrease to <5 µg/L	None to poor suppression in acromegaly, giantism, and ectopic secretion	

GH Deficiency

GH deficiency is primarily a problem of children before they have reached normal stature. Children with short stature with decreased growth velocity who are not malnourished or genetically predisposed to short stature and who have no underlying systemic illness should be suspected for GH deficiency. Serum SM-C/IGF-1 provides a useful screening technique. Insulin-induced hypoglycemia (insulin tolerance test), arginine administration (arginine tolerance test) with blood samples drawn over a 60- to 90-minute period, and stimulation with α agonists such as dopamine and clonidine, will provide sufficient information to confirm the diagnosis. The GH should peak at levels greater than 10 mg/L in normally responsive individuals.

A literature consensus of reference values for provocative testing is summarized in Table 16.2. Although normal reference intervals are listed for fasting specimens, it is generally agreed that single random specimens for GH are neither cost-effective nor diagnostically informative.

Prolactin

Prolactin (hPrL) is a 198-amino acid protein with an approximate molecular weight of 22 kd. hPrL can occur as polymers, with the dimeric form being approximately 40 to 50 kd and a much larger form with a molecular weight in excess of 100 kd. The dimeric form is usually referred to as big hPrL, whereas the much larger aggregate is referred to as big big hPrL. When the polymeric forms are present in the circulation, they generally

constitute less than 2% of the total hPrL. In addition to being present in low concentration, these polymeric forms are significantly less active biologically and pharmacologically and do not seem to represent serious pathologic conditions.

Monomeric hPrL is secreted from the anterior pituitary and enters the general circulation. Inhibitory control of hPrL is accomplished by the neurotransmitter dopamine. Dopamine secretion is initiated by stimulation of the neurons in the median eminence in response to high concentrations of hPrL in either the general circulation or the hypophyseal portal circulation. Although dopamine is capable of suppressing the secretion of hPrL, it does not appear that dopamine concentrations are sufficient to be the only control mechanism, and it is speculated there may be other inhibitory factors. In any event, hPrL appears to be the only anterior pituitary hormone without a long-loop feedback control.

hPrL secretion is stimulated by thyrotropin-releasing hormone (TRH) and perhaps some other hPrL-releasing factor. In addition to the biochemical stimulation, hPrL secretion can be initiated by suckling, whereby stimulation of the nipple through the sucking process causes a neuroreflex of the spinal cord to the neurosecretory cells of the hypothalamus. It is thought that TRH stimulates hPrL secretion at the pituitary level, whereas drugs such as levodopa and insulin (as well as stress) stimulate at the level of the hypothalamus. Secretion of hPrL is episodic, and concentrations are generally increased during the sleep period, with the highest concentrations occurring just at the time of awakening.

Lactation

Breast development through puberty requires the presence of estrogen and progesterone in conjunction with GH, adrenal steroids, and hPrL. Although estrogen and progesterone are the most important steroids for development, all these seem to be necessary but not sufficient conditions for breast development. Development is only partial until pregnancy, at which time ducts and acini mature, but the process of lactation is inhibited. Through pregnancy, placental lactogen, estrogen, progesterone, and hPrL increase dramatically, and the breast development is completed. The roles of hPrL and placental lactogen during this time are unclear.

After parturition, estrogen and progesterone concentrations decrease and hPrL concentration remains high, stimulating the synthesis of enzymes as well as initiating milk secretion. If at weaning, hPrL concentrations decrease and milk secretion will slowly taper off and cease. It is evident from this process that hPrL is the primary hormone of lactation. Suckling, in addition to stimulating hPrL secretion, stimulates an increase in the concentration of oxytocin. Oxytocin causes constriction of the tissue around the acini, forcing the milk into the primary collection ducts. During this feeding cycle, hPrL concentration increases, with peak levels being attained between 30 and 60 minutes. The ultimate concentration is dependent on the length and intensity of the suckling taking place. This elevated concentration of hPrL then causes the synthesis of milk components that will be available for the next feeding.

Gonadotropin-releasing hormone (GnRH), inhibited by the lactation process, is responsible for depressed levels of LH and follicle-stimulating hormone (FSH), and this is probably the reason there is a contraceptive effect associated with lactation. Nonlactating women will resume normal menses with normal levels of gonadotropins two to three times sooner than women who continue to lactate after delivery.

Hyperprolactinemia

There are multiple etiologies of hyperprolactinemia, with hPrL-secreting adenomas being among the most common pituitary neoplasms (15). However, drugs such as estrogens, some anti-hypertensives, phenothiazines, tricyclic antidepressants, metoclo-pramide,

and methyldopa can all cause an increase in hPrL concentration. Hypothyroidism, renal insufficiency, and the natural states of pregnancy and lactation are also reasons for increased hPrL secretion. In cases of prolactinoma, GnRH secretion is inhibited by the elevated hPrL, which leads to inhibition of ovulation and amenorrhea. In women, the most common prolactinoma is the microadenoma measuring less than 1 cm in diameter. Microadenomas are generally slow growing and can be monitored with periodic measurement of the hPrL level, provided the patient is willing to accept the consequences of elevated hPrL such as galactorrhea and amenorrhea. In the case of macroadenoma (i.e., lesions in excess of 1 cm in diameter), the lesion is generally less common, more aggressive, associated with much higher levels of hPrL and may produce neuroanatomical symptoms such as headache and visual impairment.

Laboratory Results

Commercial immunoassays for hPrL are available that use chemiluminescent, enzymatic, fluorescent, or radioactive labels with only the radioactive methods falling into disfavor. Therefore, the choice of methodology should be the one that is the most familiar and compatible to the laboratory work flow. The patient should fast for 12 hours and should not be taking phenothiazines, estrogens, monoamine oxidase, tricyclic antidepressants, antihypertensives, or methyldopa. In addition, hypothyroidism should be ruled out and if that cannot be done with certainty by history and physical examination, then a concomitant sample for TSH analysis should be drawn. The preferred specimen is serum, which should be drawn into an appropriate container and placed immediately in the cold to clot. Blood specimens should be drawn before physical examination of the breast as breast manipulation can elevate serum hPrL concentrations. After clotting is complete, the specimen should be centrifuged in a refrigerated centrifuge, the serum separated from the cells, and the serum frozen if it is not to be processed immediately. If the specimen is to be analyzed within 4 to 6 hours, refrigeration until time of assay is acceptable. Storage for longer periods requires freezing. Reference intervals for normal as well as some selected clinical situations are given in Table 16.3.

TABLE 16.3. PROLACTIN REFERENCE INTERVALS

Condition	Patient Preparation	Specimen Collection	Results (µg/L)
Normal	Fasting 12 h	Serum early AM collection	
Male		Chilled	9.5–16.1
Female		Process in cold	12.4–18.1
Pregnancy	Same	Same	
1st trimester			<80
2nd trimester			<160
3rd trimester			<400

GONADOTROPINS

Part of "16 - Endocrine Function and Carbohydrates"

FSH and LH

LH and FSH are synthesized and released from gonadotrophs located in the anterior pituitary in response to the stimulatory effect GnRH [sometimes called LH-releasing hormone (LHRH)]. Biologically active GnRH is a decapeptide that is derived from a much larger, biologically less active propeptide of approximately 70 amino acids. The active GnRH is secreted by neurons in the median eminence, where it enters the capillaries and is carried to the anterior pituitary through the hypophysial portal system. GnRH is secreted in a pulsatile fashion and seems to be sensitive to the circulating level of estrogen in women and testosterone in men. It has been demonstrated that GnRH is synthesized early in the development of the fetus and stimulates the secretion of LH and FSH until peak concentrations are reached at approximately 20 weeks of gestation. Then from the 20th week to birth, LH and FSH, along with testosterone, decline in concentration to the levels normally associated with the newborn.

LH and FSH are glycoproteins composed of two subunits referred to as the α and ß subunits. This polypeptide character is shared with two other hormones, human chorionic gonadotropin (hCG) and TSH. The α subunit of these four hormones is identical, and experiments have shown that dissociation of the α and ß subunits and recombination of the α subunit of any one of the four hormones to a ß subunit of another hormone, will produce the hormonal activity associated with the hormone of the original ß subunit. The α subunit is a necessary component for hormone binding to receptor sites, but specificity and hormone effect are conferred by the ß subunit. In addition, these four hormones show considerable homology in the structure of the ß subunit. This was a serious problem in the early immunoassays, in which antibodies could not distinguish between hCG and LH and even had some cross-reactivity with FSH and TSH. Today, with the advent of monoclonal antibodies and a better knowledge of the structure of these ß subunits, it is possible to construct assays that are very specific and very sensitive for the individual hormones. Table 16.4 lists one set of reference

intervals for men and women. Other reference values may be found, depending on the assay and the calibrators used.

TABLE 16.4. GONADOTROPIN (LUTEINIZING HORMONE/FOLLICLE-STIMULATING HORMONE) REFERENCE INTERVALS

Condition	Patient Preparation	Collection Time	LH Results (IU/L)	FSH Results (IU/L)
Normal				
Male Female	Serum (refrigerate)	Random specimen acceptable	1–8	1–7
Follicular phase	Same	Same	1–12	1–10
Ovulatory peak			15–105	6–26
Luteal phase			1–12	1–10
Postmenopausal			15–70	30–120

LH, luteinizing hormone; FSH, follicle-stimulating hormone.

The glycoprotein character of these hormones is the result of several carbohydrate residues, of which sialic acid seems to be the most important. The sialic acid content is variable and appears to be influenced by the endocrine environment at the time of synthesis. Sialic acid residues are protective of the gonadotropin, and as the sialic acid content increases, there is an extension in the half-life and biological activity of the particular gonadotropin.

Once LH and FSH have been secreted from the gonadotrophs, they are carried through the general circulation to the gonads. In the female, FSH stimulates follicle development and receptors for LH binding, whereas LH stimulates follicular production of estradiol, ovulation, and formation of the corpus luteum. In men, FSH stimulates Sertoli cell development and LH induces testosterone secretion. It is these hormones from the gonads that can feed back through the long-loop system and control the GnRH and the gonadotropin secretion.

One of the first symptoms of pituitary insufficiency can be gonadal failure. The most direct approach to ruling out primary gonadal insufficiency from pituitary insufficiency is a direct measurement of the pituitary gonadotropins. If the pituitary hormones are depressed, that is highly suggestive of pituitary insufficiency. If, conversely, the circulating levels of LH and FSH are high, pituitary deficiency can be ruled out. In Klinefelter's syndrome, a disorder of sexual differentiation, it is common to measure elevated pituitary gonadotropins and depressed testosterone. In some individuals, this picture can be altered by the finding of a normal testosterone level. This problem arises because the assay measures total circulating testosterone, which includes the very large bound fraction. The biologically active fraction, or free testosterone, is low and can only be determined by an assay that measures free testosterone in the presence of bound testosterone.

Laboratory Methods

In the measurement of LH and FSH, as in many other polypeptides, there is no special patient preparation required. That is, fasting is not an absolute requirement because most immunoassays are tolerant of mild lipemia. Fresh serum specimens are preferred because they are the least troublesome to the laboratory, but plasma is certainly an acceptable specimen. As mentioned previously, current assays are able to measure specifically any of the glycoproteins in the presence of the others, and therefore cross-reactivity and sensitivity are no longer measurement problems. In making measurements of specimens derived from

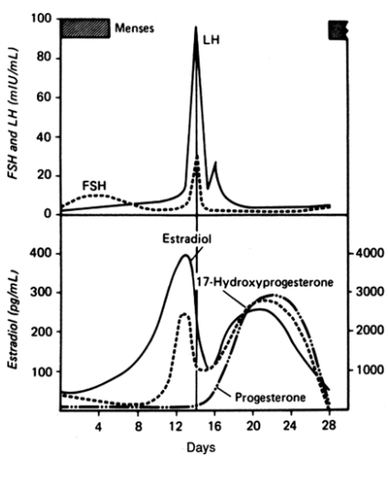


FIGURE 16.5. The cyclic secretion of gonadotropins and sex steroids during the normal 28-day menstrual cycle. Note that the estradiol and 17-hydroxyprogesterone peak precedes the mid-cycle surge of the gonadotropins luteinizing hormone and follicle-stimulating hormone. After the mid-cycle gonadotropin peak, there is a second peak of 17-hydroxyprogesterone and progesterone, which decreases to baseline levels when the cycle begins again. (From Ode WD, Moyer DL. *Physiology of reproduction*. St. Louis: Mosby, 1971, with permission.)

female patients, the most critical piece of information, with the exception of the actual level, is the menstrual status of the woman. As Fig. 16.5 demonstrates, the LH and FSH concentrations vary considerably throughout the menstrual cycle. It is imperative that the physician have access to accurate menstrual information when interpreting gonadotropin results.

POSTERIOR PITUITARY

Part of "16 - Endocrine Function and Carbohydrates"

Physiology

Antidiuretic hormone [(ADH), vasopressin] is secreted from hypothalamic neurons of the supraoptic nuclei and transported to the posterior pituitary by axoplasmic flow. From these storage sites, it is released into the systemic circulation in response to osmotic and volume stimuli and circulates to the kidney. At the renal tubule level, the effect of ADH is to increase water permeability of the distal tubule collecting duct, permitting water resorption resulting in concentration of the final urine. The osmoreceptors responsible for ADH release are located in the anterior hypothalamus near the supraoptic nuclei and are responsive to small changes in extracellular tonicity.

A physiologic set point for ADH secretion exists at an approximate plasma osmolality of 280 mOsm/kg. Below this set point ADH secretion is suppressed. ADH secretion rises rapidly above this set point, with maximal ADH levels achieved at extracellular tonicities of approximately 300 mOsm/kg, at which point thirst and water seeking are also stimulated in an integrated fashion. The threshold and sensitivity for ADH secretion vary among individuals and are increased by hypovolemia and hypercalcemia, as well as by many drugs, including lithium and carbamazepine. The threshold is lowered by hypervolemia, glucocorticoids, alcohol, and opiates.

ADH secretion is also altered by isosmotic changes in extracellular volume. Decreases in volume amplify the ADH release to any given osmotic stimulus, and increased volume suppresses ADH secretion in a similar manner. Nausea, pain, and many drugs such as narcotics are also potent nonosmotic stimulators of ADH secretion.

Clinical Disorders

Diabetes Insipidus

Defects in ADH secretion (central) or renal tubular action (nephrogenic) result in chronic polyuria, polydipsia, and thirst. In addition, functional suppression of ADH occurs with chronic psychogenic water drinking, producing a similar clinical picture (16). Central diabetes insipidus can result from a variety of hypothalamic etiologies, including trauma, various neoplasms, and granulomatous and infectious disorders. A heritable familial autosomal dominant central disorder is also described. In a significant percentage of patients with central diabetes insipidus, no specific etiology can be determined and the patients are labeled idiopathic.

Nephrogenic diabetes insipidus can occur as a result of renal parenchymal disorders such as amyloidosis, polycystic renal disease, and sickle cell disease, as well as after relief of urinary tract obstruction. Drugs such as lithium, demeclocycline, and methoxyflurane can also produce nephrogenic diabetes insipidus. A familial X-linked recessive disorder is also recognized (17).

The diagnosis of diabetes insipidus and the differentiation between central and nephrogenic etiologies is achieved by a structured dehydration test. In short, the procedure requires the determination of baseline values for body weight, serum and urine osmolality, and serum electrolytes. Fluid is withheld, and body weight and urine osmolality are monitored to assess the adequacy of the dehydration stimulus. The end point is achieved when there has been a 3% to 5% decrease in body weight. In patients with central diabetes insipidus despite an adequate dehydration stimulus, there is a failure to achieve urine osmolalities greater than the serum osmolality in the presence of low plasma ADH levels. Confirmation of the diagnosis is achieved by the administration of an ADH analog, such as desamino-D-arginine vasopressin (DDAVP), which produces a urine concentration increase of 50% over the preadministration value. This demonstration of the renal tubule to respond to ADH confirms the failure to secret ADH as the etiology of the diabetes insipidus. In patients with nephrogenic diabetes insipidus, there is an inability to concentrate urine to an osmolality greater than serum osmolality and there is an insufficient response after administration of an ADH analog. Patients with excessive water drinking (primary polydipsia) have an acquired defect in urinary concentration attributable to a physiologic resistance to ADH owing to the high volume of water intake and the reduced medullary tonicity. Consequently, urine concentration during dehydration may be submaximal but will exceed 300 mOsm/kg with no further increase in urine osmolality after exogenous ADH or DDAVP administration. More prolonged fluid restriction or therapeutic ADH administration restores full urine-concentrating ability.

Primary Hypodipsia

In some patients with or without coexistent central diabetes insipidus, defects in thirst perception owing to destruction of hypothalamic osmoreceptors are also present. These defects lead to an inappropriate thirst response to dehydration and at times striking asymptomatic hypernatremia and hyperosmolality. Other causes of hypernatremia, including inadequate fluid intake secondary to coma or physical restraints preventing access to fluids, should be excluded. Plasma ADH levels are usually appropriate for the degree of osmolality unless diabetes insipidus is also present. Table 16.5 lists some common plasma and urine laboratory results for polyuria.

TABLE 16.5. URINE OSMOLALITY FOR DIABETES INSIPIDUS OF DIFFERENT ORIGINS AND POLYDIPSIA

	Neurogenic Diabetes Insipidus	Nephrogenic Diabetes Insipidus	Psychogenic Polydipsia
Random plasma osmolality	↑	↑	↓
Random urine osmolality	↓	↓	↓
Urine osmolality during mild water deprivation	No change	No change	↑
Urine osmolality during nicotine or hypertonic saline	No change	No change	↑
Urine osmolality after vasopressin intravenously	↑	No change	↑
Plasma vasopressin	Low	Normal or high	Low

From Greenspan PS. *Basic endocrinology*, 3rd ed. Norwalk, CT: Appleton & Lange, 1991, with permission.

Syndrome of Inappropriate Secretion of ADH

The continued secretion of ADH despite appropriate serum osmolality and continued unrestricted water intake, results in dilutional hypoosmolality and hyponatremia, inappropriately increased urinary osmolality, and increased urinary sodium excretion (18). Many etiologies may be responsible. Ectopic production of ADH by nonendocrine carcinoma, several acute and chronic pulmonary conditions, and a variety of intracranial disorders

are the most common etiologies. The diagnosis is established clinically by documenting hyponatremia and hypoosmolality in the setting of normal extracellular volume. Spurious hyponatremia secondary to hyperglycemia, dysproteinemia, and hyperlipemia should be excluded, as well as hypothyroidism and hypoadrenalism, which can impair renal free water excretion. (18).

THYROID

Part of "16 - Endocrine Function and Carbohydrates"

Thyroid Physiology

Iodide is concentrated by the thyroid gland, organified into tyrosine units on thyroglobulin, and stored in colloid-containing thyroid follicles until secretion. The metabolism of iodine is shown in Fig. 16.6. Synthesis and secretion of T_4 and T_3 is controlled by pituitary TSH secretion through classic negative feedback regulation in combination with the stimulatory effects of hypothalamic TRH. Serum T_3 arises both from direct thyroid gland secretion (15%) and from peripheral monodeiodination of the outer tyrosine ring of T_4 (85%). The synthesis and control of thyroid hormones is diagrammed in Fig. 16.7. A biologically inactive reverse T_3 (RT_3) also arises from peripheral monodeiodination of the inner tyrosine ring of T_4.

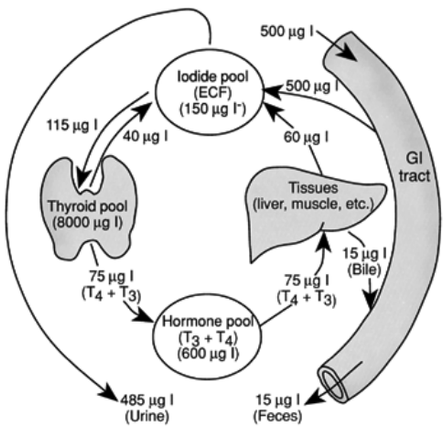


FIGURE 16.6. Daily iodine adsorption, synthesis and excretion in a healthy subject. (From Greenspan FS, Rapoport B. Thyroid gland. In: Greenspan FS, ed. *Basic and clinical endocrinology*, 3rd ed. Norwalk, CT: Appleton & Lange, 1991:188–246, with permission.)

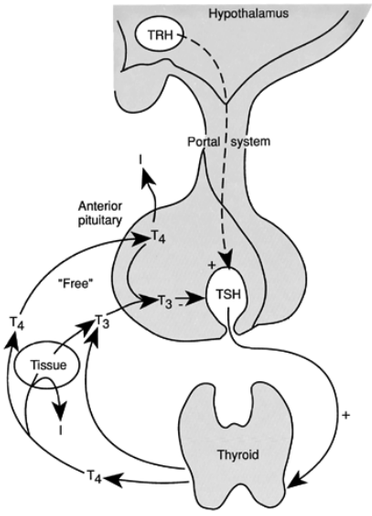


FIGURE 16.7. The hypothalamic-hypophyseal-thyroid axis in the production of thyroid hormones. Pathways marked with a | indicate inhibition, whereas those with a + indicate stimulation. I indicates deiodination of the parent thyroxine molecule to produce triiodothyronine. (From Greenspan FS, Rapoport B. Thyroid gland. In: Greenspan FS, ed. *Basic and clinical endocrinology*, 3rd ed. Norwalk, CT: Appleton & Lange, 1991:188–246, with permission.)

Thyroid Transport

Circulating T_3 and T_4 are bound to a variety of thyroid transport proteins, including thyroid-binding globulin (TBG), thyroid binding prealbumin, and albumin. Only approximately 0.03% of T_4 and 0.3% of T_3 are “free” and hence biologically active. It is the excess or deficiency of the free T_4 fractions that determines hyperthyroidism and hypothyroidism. Because many drugs and disease states affect the concentrations of the thyroid transport proteins, it is important that thyroid function studies distinguish between elevations or depressions of the bound or total concentration, and changes in the free fraction of T_4 and T_3 (19).

The most common abnormalities in thyroid hormone studies are related to the widespread use of oral estrogens for contraception and postmenopausal estrogen replacement therapy. The estrogen-related increase in the hepatic production of TBG produces an elevation of serum total T_4, with preservation of normal free T_4 levels. A less common cause for abnormal total T_4 levels is a hereditary X-linked familial trait that occurs in approximately one in 5,000 newborns. This genetic defect causes an increase in TBG levels with the resultant elevation of total T_4 levels.

Clinical Disorders

Hyperthyroidism

Independent of the etiology, hyperthyroidism is characterized by the presence of elevated levels of thyroid hormone. The diagnosis of hyperthyroidism is established by thyroid function studies including biochemical, radioactive iodine uptake, and thyroid scans. In the extreme, the condition is easily diagnosed, but there is a spectrum of presentations that may be more problematic. The most common etiology of hyperthyroidism is Graves' disease,

or diffuse toxic goiter. Graves' disease is considered to have a major autoimmune component. Graves' disease patients have circulating antibodies that were originally called long-acting thyroid stimulators but are now referred to as either thyroid-stimulating antibodies or thyroid-stimulating immunoglobulins. The mode of action of these antibodies is to bind to the TSH receptor and stimulate the thyroid, causing the secretion of thyroid hormones. This stimulation of the thyroid is not sensitive to feedback mechanisms and therefore does not respond to therapeutic measures directed at TSH or TRH control. Other etiologies of hyperthyroidism include toxic solitary or multinodular thyroid disorders, inflammatory disorders of the thyroid gland such as subacute thyroiditis, hyperthyroidism secondary to placenta chorionic thyroid stimulators, excess exogenous thyroid administration, and, rarely, TSH production from a pituitary adenoma. With the exception of the latter, circulating free T_4 and free T_3 levels, as determined directly or calculated by indexing techniques, are elevated and serum TSH levels, as measured by sensitive immunoassay methods, are suppressed. When reviewing the laboratory results, an occasional patient will have only an elevated T_3 (T_3 toxicosis), in support of the diagnosis (19).

To prove TSH suppression in patients with hyperthyroidism, provocative testing with TRH can be undertaken (TRH, 250 to 500 mg i.v., serum TSH at 0 and 30 minutes). Normal patients will respond with a two- to fivefold rise in serum TSH concentration, whereas patients with hyperthyroidism will not exhibit an increase in TSH levels. This suppression is the result of a competent feedback loop responding to the elevated T_4 and T_3. With the introduction of new, sensitive TSH assays capable of measuring serum TSH levels in the range of 0.01 μU/mL, TRH testing has become less frequent. Hyperthyroidism has the potential of manifesting itself in one or more organ systems without obvious thyroid dysfunction. Therefore, evaluation of thyroid status is a prudent action when evaluating patients with such common and diverse complaints as weight loss, sexual dysfunction, cardiac disease, or complaints of generalized weakness.

Hypothyroidism

The diagnosis of primary hypothyroidism is simplified by documentation of an elevated serum TSH. If the negative feedback loop were intact, the thyroid gland would be secreting thyroid hormones in response to the elevated levels of TSH. In patients with central or secondary hypothyroidism (5%), serum TSH levels are inappropriately normal, indicating failure at the pituitary-hypothalamic level to increase TSH concentration in response to decreased circulating thyroid hormone levels. Some hypothyroid patients (a small percentage) will have serum total and free T_4 levels within normal reference intervals and will be recognized as having primary hypothyroidism on the basis of an isolated serum TSH elevation or an exaggerated serum TSH response after TRH administration. Occasional spurious serum TSH elevations can arise from interference in immunoassays by circulating autoantibodies to TSH antimurine antibodies used in the TSH assays. TRH administration results in an excess increase in serum TSH in patients with indeterminate or borderline elevations of basal serum TSH levels.

Neonatal screening for congenital hypothyroidism has become nearly universal. Initial screening is performed using serum T_4 determinations by methods adapted to specimens collected and transported on filter paper. These specimens are then analyzed for TSH and an elevated result is strong evidence for primary hypothyroidism The incidence of neonatal primary hypothyroidism is approximately 1:5,000. The disorder is sporadic and usually nonfamilial, requiring universal screening of newborns to permit early diagnosis and prompt institution of thyroid replacement therapy. False-positive results (low T_4 result) may occur with heritable TBG deficiency or decreased serum TBG concentrations related to prematurity.

The Euthyroid Sick State

A unique problem arises in the thyroid evaluation of critically ill hospitalized patients. A complex series of events, including decreased serum thyroid binding protein concentration, circulating inhibitors of thyroid binding, decreased peripheral T_4 deiodination to T_3, qualitative abnormalities in serum TSH glycosylation,

and quantitative abnormalities in TSH secretion, combine to produce distortions of conventional thyroid function tests in these patients. These effects are further complicated by the common use of drugs such as glucocorticoids, beta-blockers, and DOPA (which directly suppress pituitary TSH secretion) in these critically ill patients. To resolve the true thyroid status often requires direct measurement of free T_4 rather than indirect estimates of free thyroid hormone levels. An elevation of serum TSH of more than 15 μU/mL is satisfactory evidence of coexistent primary hypothyroidism in these critically ill patients. Current evidence suggests that at least some of these changes are the result of a metabolic adaptation to catabolic systemic illness and that thyroid replacement therapy is unwarranted and unhelpful.

Nodular Thyroid Disease

Laboratory thyroid function studies play a limited role in the diagnosis and management of patients with nodular and multi-nodular thyroid disease. Multinodular thyroid glands may produce hyperthyroidism if sufficient non-TSH-regulated thyroid function arises from autonomous nodules. Conversely, multi-nodular thyroid disease can be a secondary result of chronic TSH hypersecretion in patients with underlying hypothyroidism. In either case, the functional diagnosis must be separated from any anatomical etiology.

Patients with carcinoma of the thyroid present with isolated thyroid nodules and are usually euthyroid. The initial diagnosis and subsequent monitoring of therapy for papillary and follicular thyroid carcinoma may be supplemented by measurement of TSH, serum thyroglobulin and serum calcitonin levels. The latter may be enhanced by stimulatory tests employing pentagastrin and intravenous calcium administration.

Thyroid Therapy

Thyroid function studies are employed in the initial diagnosis of thyroid dysfunction and in the subsequent monitoring of the patient's response to therapy. It is essential to understand the effect of therapy on thyroid function and laboratory thyroid function testing. Increasingly, third-generation serum TSH assays are becoming the monitor of antithyroid therapy of hyperthyroidism, thyroid replacement therapy in primary hypothyroidism, and thyroid suppression in patients with benign and malignant thyroid neoplastic disorders (20).

Laboratory Diagnosis

Until recently the basis for all laboratory methods used in a thyroid evaluation was immunochemical. Modern automated chemistry analyzers are now capable of performing adequate total T_4 and T_3 resin uptake assays. The exquisite sensitivity and specificity of current antibodies have made the other thyroid testing a routine event in most clinical laboratories. Total T_3, free T_4, free T_3, TSH, TBG, and even some of the byproducts such as RT_3 can be measured accurately and precisely with only a fraction of the effort required several years ago. Because of current technology, laboratories are able to choose procedures for thyroid function testing that will fit in with the existing work flow and other techniques employed in the clinical laboratory.

Special patient preparation is not required for biochemical thyroid studies. If possible, a serum specimen obtained after an overnight fast is preferred. However, if time or circumstance does not permit the collection of a fasting specimen, a random specimen may be used. Serum should be separated from the cells and refrigerated until the time of assay or frozen for analyses that will be performed longer than 8 hours after the specimen collection. There is an extensive list of drugs that interfere with thyroid physiology and/or specimen analysis, and therefore the laboratory should periodically refer to one of the publications that deals with drug interference in common laboratory tests. Some of the more commonly occurring drugs that are known to affect testing are prostaglandin, estrogen, oral contraceptives, androgens, aspirin, corticosteroids, lithium, some antibiotics, phenothiazines, dexamethasone, propranolol, and certainly many others (21).

As is the case with all laboratory procedures, each laboratory should establish their own reference intervals for the population they serve. As a guide, serum T_4 is between 4.5 and 12 mg/dL, T_3 resin uptake between 20% and 35%, serum total T_3 between 100 and 220 ng/dL, and serum TSH between 0.5 and 5.0 IU/L. Because of the diversity of these reference intervals, Table 16.6 lists thyroid function studies in terms of increase, decrease, or normal. This is a common technique used to show the relationship among the thyroid studies without having to be concerned about the specific reference intervals of a method or geographic location.

TABLE 16.6. THYROID FUNCTION STUDIES

	Serum					
Clinical Disorder	**T_4**	**T_3U**	**fT_4I**	**T_3**	**fT_3I**	**TSH**
Elevated T_4						
Hyperthyroidism	↑	↓	↑	N/ ↑	↑	↓
Increased TBG						
Estrogens	↑	↓	N	↑	N	N
Pregnancy	↑	↓	N	↑	N	N
Drugs						
1-T Rx	N/ ↑	N	N/ ↑	N	N	↓
Glucocorticoids	N/ ↑	N	N/ ↑	↓	↓	N
Propranolol	N/ ↑	N	N/ ↑	N/ ↓	N/ ↓	N
Amiodorone	↑	↑	↑	↓	↓	N
Thyroid resistance	↑	↑	↑	↑	↑	N/ ↑
T_3 Toxicosis	N	N	N	↑	↑	↓
Depressed T_4						
Hypothyroidism						
Primary	↓	↓	↓	N	N	↑ ↑
Secondary	↓	↓	↓	N	N	N/ ↓
Decreased TBG						
Androgens	↓	↑	N	↓	N	N
Hepatic failure	↓	↑	N	↓	N/ ↓	N
Nephrosis	↓	↑	N	↓	N/ ↓	N
Drugs						
T_3	↓	N	↓	N/ ↑ [a]	↓	N
Phenytoin	↓	↑	N/ ↓	↓	N	N
Thyroid carcinoma	N	N	N	N	N	N
	↑ = increased ↓ = decreased N = normal					

[a] Time-dependent value.
T_4, thyroxine; TBG, thyroid-binding globulin; Rx, prescriptions; T_3, triiodothyroxine.

The measurement of free T_4 or the estimation of free T_4 by indexing procedures is now widely practiced and has complemented total T_4 measurements. In the past, free T_4 measurements were done by equilibrium dialysis, which was most commonly done in research laboratories or large clinical reference laboratories. Early attempts to approximate the free T_4 concentration by using calculated indices, although cost-effective, proved to be limited in some clinical applications. Proprietary analog procedures to estimate free T_4 were an improvement over the calculated indices but were still problematic when measuring free T_4 in samples from sick euthyroid patients. If the laboratory is thoroughly familiar with the limitations of the analog procedures, they can prove useful in free T_4 assessment.

There is growing enthusiasm for measuring serum TSH (by a third-generation ultrasensitive TSH assay) as an initial step in thyroid evaluation followed by a free T_4 in those cases that require further investigation. This approach will replace the total T_4 and T_3 resin uptake that is currently done by a large number of laboratories, as the costs of the former become more competitive. The serum ultrasensitive TSH is also employed to monitor thyroid hormone replacement and suppression therapy in patients receiving thyroid therapy. Serum TSH should be maintained within or slightly below the normal reference interval for thyroid replacement therapy or in the case of thyroid suppression, in a range that, although below the reference interval, remains easily measurable.

Triiodothyronine

Total T_3 measurements are influenced by many factors, including protein concentration. Free T_3 measurement fulfills a similar function

to free T_4 because it is a direct measurement of the physiologically active hormone and is not affected by protein concentration.

T_3 Uptake

This is the historical way of measuring thyroid hormone binding to serum proteins. Although there are many approaches to this assay, all T_3 uptake assays attempt to estimate the relative saturation of all the serum proteins capable of binding T_4. This result, when multiplied by the total T_4 result, produces a value referred to as free T_4 index (FTI). Before there were direct measurements for free T_4, this approach gave a reasonable estimate of free T_4. Today there are several methods available for the quantitation of free T_4, so the FTI is less important.

Thyroid-Stimulating Hormone (Thyrotropin)

Many commercial immunoassays are available that are in the ultrasensitive or supersensitive category, meaning that circulating TSH at levels of 0.01 to 0.03 U/L can be accurately measured (22). It is this generation of TSH assay that is being considered as one of the first assays for thyroid screening.

Thyroid-Binding Globulin

Assays for TBG are available that are sufficiently reliable and simple so that most laboratories can perform them on a routine basis. It should be remembered that the specificity of immunoassays results in a quantitation of the TBG protein and not an assessment of the available thyroid-hormone binding sites. Therefore, the serum TBG assay is not a functional replacement for T_3 uptake or for measurement of the serum free T_4 concentration.

Evocative Testing

Stimulation by the intravenous administration of TRH amplifies the information available from measurement of steady-state thyroid and TSH levels. TRH (250 to 500 μg) is administered intravenously over approximately 1 minute. Serum TSH levels are obtained at time zero and at 30 or 60 minutes. In normal subjects, an increase of from 5 to 20 mIU/mL over baseline is obtained. In patients with hyperthyroidism, the low serum TSH does not increase after TRH administration. In patients with primary hypothyroidism, an excessive rise in serum TSH occurs at 30 and 60 minutes. In hypothyroidism secondary to pituitary or hypothalamic disease, serum TSH increase is either absent, suboptimal, or delayed. The wide availability of the new ultrasensitive serum TSH measurement has diminished the need for dynamic TRH testing.

Serum Thyroglobulin

Serum thyroglobulin determinations are an important aspect of monitoring for residual or recurrent differentiated papillary and

follicular thyroid carcinoma, either at the time of original diagnosis or after surgical or radioisotope therapy. Elevated serum thyroglobulin indicates residual carcinoma even in the absence of clinical or thyroid scan evidence of residual disease. Maximum sensitivity of serum thyroglobulin measurements is achieved after either withdrawal of thyroid suppression or administration of exogenous recombinant TSH administration to elevate serum thyroglobulin levels. Elevated serum thyroglobulin levels are also seen in inflammatory thyroid disorders such as subacute thyroiditis.

Serum Calcitonin

Serum calcitonin determinations are employed to screen for medullary thyroid carcinoma. Medullary thyroid carcinoma may be sporadic or associated with multiple endocrine neoplasia syndrome type 2 (MEN 2). MEN 2A is a multiple endocrine disorder including medullary thyroid carcinoma, hyperparathyroidism, and pheochromocytoma. MEN 2B is a similar MEN disorder but with the addition of clinical mucosal neuromas. MEN1 is an independent disorder of pituitary adenoma, hyperparathyroidism, and pancreatic or foregut neuroendocrine tumors including gastrinomas, insulinomas, and other neuroendocrine neoplasms. MEN 2 and MEN 2A are caused by mutations in the *ret* oncogene located on chromosome 10. Molecular techniques for DNA screening of the *ret* oncogene have been useful in diagnosis and screening of the *ret* oncogene in families suspected of having this autosomal dominant disorder.

ADRENAL CORTEX

Part of "16 - Endocrine Function and Carbohydrates"

Physiology of Adrenal Biosynthesis and Regulation

The adrenal cortex is composed of three distinct anatomical layers, each associated with specific steroid hormone synthesis and secretion. The glomerulosa is the most superficial and is responsible for mineralocorticoid biosynthesis and resultant renal sodium:potassium homeostasis. The glomerulosa is controlled principally by the renin–angiotensin system.

The fasciculata is the second layer of the adrenal cortex, is the thickest and most prominent on examination, and is responsible for glucocorticoid production, including cortisol. The fasciculata is under the regulation of pituitary ACTH in a long-loop negative feedback from peripheral cortisol. Stimulation is accomplished through the release of ACTH-releasing hormone from the central hypothalamus. ACTH is the product of the anterior pituitary processing a large precursor peptide, pro-opiomelanocorticotropin, which also contains melanocyte stimulating hormone (MSH) and endorphin sequences. Cortisol circulates in both the bound and free form. The bound portion (90%) is associated with a specific transport protein, cortisol-binding globulin (CBG), and serum albumin. The 10% of cortisol that circulates as free cortisol is the biologically active fraction.

The reticularis is the innermost zone of the adrenal cortex and is responsible for androgen and estrogen production. Regulation is achieved through pituitary ACTH levels but only in response to cortisol feedback from the fasciculata, an important interrelationship in congenital adrenal hyperplasia syndromes related to adrenal enzyme defects.

Another important regulator of adrenal steroid synthesis and secretion is the renin–angiotensin system that selectively regulates adrenal mineralocorticoid production. This system is strongly influenced by adrenergic factors and by sodium and potassium concentrations.

Clinical Disorders

Cushing Syndrome

Cushing's syndrome is the clinical disorder that results from sustained glucocorticoid excess. Clinical features include central weight gain, weakness, and facial plethora as well as hirsutism and menstrual irregularity in female patients. When Cushing's syndrome is suspected, screening laboratory studies are indicated to establish glucocorticoid excess followed by additional studies to define the etiology of the disorder and to plan specific therapy.

The most common etiology of Cushing's syndrome is pituitary ACTH–dependent, bilateral adrenal hyperplasia (Cushing's disease). Other etiologies include primary tumors of either adrenal gland, bilateral macronodular adrenal hyperplasia, or pigmented micronodular hyperplasia. Finally, some nonendocrine neoplasms of the lung and pancreas in addition to carcinoid tumors of the bronchi and thymus can produce ectopic nonpituitary ACTH syndromes (23).

Elevated urinary excretion of cortisol and cortisol metabolites is reflected in 17-hydroxysteroid, and 17-ketosteroid determinations. Urinary creatinine should always be included to ensure adequate collections. Elevated serum cortisol and loss of normal physiologic diurnal variation of serum cortisol are complementary or alternative methods of establishing quantitative glucocorticoid excess. Obesity alone may increase urinary cortisol metabolite excretion and estrogens, either endogenous secondary to pregnancy or exogenous secondary to anovulatory or estrogen replacement therapy, may increase CBG and result in false-positive serum cortisol elevations. In ACTH-dependent Cushing's syndrome secondary to either pituitary ACTH–dependent Cushing's syndrome or ectopic ACTH production from a paraendocrine neoplasm, plasma ACTH concentrations will be either frankly elevated or inappropriately normal, i.e., nonsuppressed.

Dexamethasone, a potent glucocorticoid derivative, may be used to document qualitative glucocorticoid abnormality by demonstrating lack of physiologic adrenal suppressibility. Normal patients or those with uncomplicated obesity will demonstrate a decrease of 50%, and often 90% or more, in urinary cortisol excretion when treated with dexamethasone (0.5 mg orally every 6 hours for 48 hours). A single dose of dexamethasone, 1 mg orally at 11 p.m., will suppress serum cortisol at 8 a.m. the following morning. Patients with any of the Cushing's syndrome etiologies will not demonstrate physiologic serum or urinary suppression. False positives may arise from estrogen administration, which increases serum cortisol binding (CBG), or the administration of drugs such as phenobarbital and analeptics, which accelerate metabolic clearance of dexamethasone.

Patients with bilateral adrenal hyperplasia suppress urinary cortisol excretion with larger doses of dexamethasone (2 mg orally every 6 hours, for a subsequent 48 hours) (24). Those with autonomous primary adrenal disorders such as adrenal adenoma or carcinoma or those with autonomous ACTH secretion arising from paraendocrine neoplasms will also fail to suppress with the larger dose of dexamethasone (25).

Alternatively, exogenous ACTH or endogenous ACTH stimulated through inhibition of cortisol synthesis by pharmacologic agents such as metyrapone may also provide useful supplementary information (Table 16.7).

TABLE 16.7. CUSHING'S SYNDROME

	Normal	Cushing's Syndrome			
Screening Studies					
Urinary free cortisol	<80 µg/24 h	>100 µg/24 h			
Overnight dexamethasone suppression study	<5µg/dl	>5 µg/dl			
	Normal	**Pituitary Dependent**	**Adrenal Adenoma**	**Adrenal Carcinoma**	**Ectopid ACTH**
Definitive Studies					
Baseline					
Urinary					
17-OHCS (mg/24 h)	2-8	>100%	>100%	>100%	>>100%
17-Ketosteroid	5-15			>>100%	
Cortisol (µg/24 h)	20-80	>100%	>100%	>100%	>100%
Serum					
Cortisol 8 AM (µg/dl)	6-26				
Cortisol 4-8 PM	4-14				
Dehydroepiandrosterone sulfate (µg/dl)	40-360			>>100%	
Plasma corticotropin pg/mL	20-80	>50	<20	<20	>>50
Suppression					
Dexamethasone (0.5 mg every 6 h × 48 h)	>50%	>50%	>50%	>50%	>50%
Dexamethasone (2.0 mg every 6 h × 48 h)	>90%	>90%	>50%	>50%	>50%
Stimulation					
ACTH stimulation (0.5 mg i.v./8 h)	>50%	>100%	Often >50%	10%	Variable
Metapyrone stimulation: 750 mg every 6 h × 24 h	>50%	>50%	<50%	<50%	<50%
Corticotropin-releasing hormone (100 µg i.v.)					
Serum cortisol		>20%	<20%	<20%	<20%
Plasma corticotropin		>50%	<20%	<20%	<20%

Pituitary localization as well as lateralization can be accomplished by measurement of plasma ACTH obtained by bilateral petrosal sinus catheterization (26). This information is particularly helpful when transsphenoidal pituitary surgery for correction of ACTH-dependent Cushing's disease is anticipated.

Adrenal Insufficiency

Adrenal insufficiency (Addison's disease) can arise from either primary adrenal cortical failure or because of secondary pituitary ACTH deficiency. The disorder can present as both a chronic and an acute syndrome.

The most common etiology of chronic adrenal insufficiency is autoimmune adrenalitis. Other etiologies include granulomatous disorders such as tuberculosis, sarcoidosis, adrenal leukodystrophy, and metastatic adrenal involvement. Autoimmune adrenalitis may occur as an isolated phenomenon or as a component of polyglandular autoimmune syndromes. Polyglandular autoimmune syndrome type I includes patients with adrenal insufficiency, hypoparathyroidism, and mucocutaneous moniliasis. Polyglandular autoimmune syndrome type II includes patients with adrenal insufficiency, diabetes mellitus, and thyroid dysfunction. Fatigue, weight loss, nausea, and hyperpigmentation are clinical presentations in patients with chronic adrenal insufficiency. The latter is related to the overproduction of pituitary ACTH and MSH sequences, which are derived from pituitary processing of pro-opiomelanocortin. Acute adrenal insufficiency is most often seen in the intensive care setting in stressed and anticoagulated patients. Hypotension and shock dominate the clinical picture of acute adrenal insufficiency.

Secondary adrenal insufficiency arises as a result of infiltrative or destructive disorders of the hypothalamic-pituitary axis or from functional suppression by exogenous glucocorticoid administration.

Congenital adrenal hyperplasia syndromes are the result of heritable defects in the enzymes required for cortisol biosynthesis. The resultant disorders are characterized both by cortisol deficiency and often by androgen excess resulting from ACTH stimulation secondary to cortisol deficiency. The most common disorders involve deficiencies of the 21-hydroxylase enzyme involved in cortisol and mineralocorticoid synthesis. In the newborn, syndromes of glucocorticoid and mineralocorticoid deficiency as well as ambiguous external genitalia in female patients are common. In later childhood and adolescent stages, disorders of hirsutism, virilization, and menstrual irregularity are seen.

Isolated Mineralocorticoid Deficiency

A syndrome of isolated mineralocorticoid deficiency is seen in older patients who present with recurrent hyperkalemia but normal cortisol and ACTH secretion. This disorder is particularly

common in patients with diabetes mellitus and underlying kidney diseases such as interstitial nephritis. Evaluation requires exclusion of primary adrenocortical insufficiency by serum cortisol and ACTH testing followed by documentation of defective plasma renin and serum aldosterone response to sodium restriction and upright posture.

Adrenal Virilization and Congenital Adrenal Hyperplasia

Virilization secondary to adrenal disorders occurs in patients with congenital adrenal hyperplasia syndromes and as a component of Cushing's syndrome secondary to adrenocortical androgen excess. Rarely, adrenal cortical adenoma may secrete androgens alone, producing isolated adrenal androgen excess. Elevated serum dehydroepiandrosterone (DHEA) sulfate, an excellent marker of adrenal androgen excess, is the most consistent finding, as are selective elevations of urinary 17-ketosteroids.

The syndromes of congenital adrenal hyperplasia are a group of inborn disorders of adrenal cortisol biosynthesis that have defects in cortisol biosynthesis and often androgen excess as a consequence of increased ACTH secretion. The most common enzyme defect is at the 21-hydroxylase step (27). Affected female newborns present with ambiguous genitalia as a result of intrauterine exposure to excessive adrenal androgens. Affected male newborns have normal external genitalia and share with their female counterparts variable renal sodium loss as a result of mineralocorticoid deficiency. Affected individuals not properly diagnosed at birth may present in later childhood with sexual precocity in males and virilization in females as a result of ongoing adrenal androgen production. The diagnosis of congenital adrenal hyperplasia is established by measurement of serum 17-hydroxyprogesterone, which is the precursor for the 21-hydroxylase enzyme. Recently, a more subtle defect in 21-hydroxylase has been recognized in girls at puberty with resultant hirsutism and menstrual irregularity. Diagnosis requires measurement of 17-hydroxyprogesterone before and 60 minutes after ACTH (Cortrosyn) administration. Other relatively common adrenal enzymatic disorders occur at the 11-hydroxylase and 3ß-dehydrogenase steps. The former is associated with increased desoxycorticosterone synthesis, resulting in hypertension rather than sodium loss, and the latter by profound cortisol, mineralocorticoid, and androgen deficiency. The laboratory diagnosis is established by documentation of increased 11-desoxycortisol in the 11-hydroxylase disorder and by the ratio of pregnenolone:progesterone, 17-hydroxypregnenolone:17-hydroxyprogesterone, or DHEA:androstenedione in the 3ß-dehydrogenase deficiency. Finally, it should be noted that all steroid hormones exert their biological actions through cellular receptors and that clinical disorders that mimic adrenal biosynthetic abnormalities may arise as a result of defects in glucocorticoid, mineralocorticoid, or androgen receptors.

Adrenal Hypertension

Hypertension resulting from primary adrenal mineralocorticoid excess can be caused either from mineralocorticoid-secreting solitary adrenal adenomas (aldosterone-secreting adenomas) or from bilateral adrenal adenomatous hyperplasia. Hypokalemia and elevated urinary potassium excretion are related to increased mineralocorticoid biosynthesis, and plasma renin is suppressed by the resultant extracellular volume expansion. Failure of serum aldosterone to suppress after intravenous saline infusion and of plasma renin to increase after sodium restriction, upright posture, and diuretic administration are additional characteristics (28). The specific etiology can then be distinguished by the use of adrenal computed tomography (CT) scanning, ^{131}I-iodo-cholesterol adrenal scanning, or differential adrenal vein catheterization studies. Rarely, other mineralocorticoid-secreting adrenal adenomas may produce a similar clinical picture.

Secondary aldosterone hypersecretion as a result of increased renin secretion is associated with renal artery stenosis, accelerated hypertension of any etiology, and, in some patients, oral estrogen administration, particularly anovulatory agents, as a result of estrogen-mediated hepatic angiotensinogen secretion.

Adrenal Incidentaloma

With the widespread use of CT and magnetic resonance imaging for abdominal imaging, a substantial number of adrenal incidentalomas have been recognized. It is suggested that as many as 5% of the population may have such incidental adrenal nodules. These represent adenomas and other focal adrenal gland enlargements in patients without previously recognized adrenal abnormality. Laboratory evaluation should include a serum cortisol, before and after overnight dexamethasone administration, serum electrolytes, serum DHEA sulfate, and urinary vanilmandelic acid (VMA) and metanephrines to exclude an occult functional adrenal lesion. If the nodules are less than 5 cm and nonfunctional, no intervention is necessary and imaging should be repeated in 6 months. Nodules greater than 5 cm should be removed surgically to exclude possible occult adrenal carcinoma.

THE ENDOCRINE PANCREAS AND GASTROINTESTINAL PEPTIDES

Part of "16 - Endocrine Function and Carbohydrates"

The islet cells of the pancreas and the endocrine cells of the foregut constitute an integrated endocrine system functionally related to alimentation and metabolic fuel integration. The islets of Langerhans of the pancreas contain alpha, beta, and delta cells respectively secreting glucagon, insulin, and somatostatin. The close proximity of these cells permits both paracrine, (i.e., local interactions at a cell-to-cell level) and more traditional endocrine actions. The common embryologic origin of these endocrine cells from histochemically distinct amine uptake and decarboxylation (APUD) cells of the embryologic neuroectoderm results in the presence of cells in the adult foregut that secrete gastrin, motilin, vasoactive intestinal peptide (VIP), and somatostatin and in peptide-secreting cells that are distributed from the hypothalamic-pituitary axis, through the thyroid and bronchial epithelium, to the distal small bowel. These foregut peptides are involved in the integration of intestinal secretion and motility. All these sites may become clinically involved in endocrine neoplastic syndromes.

Diabetes Mellitus

The most common endocrine disorder of the endocrine pancreas is type I insulin-dependent diabetes mellitus (IDDM). In this disorder, there is progressive loss of pancreatic beta-cell function and without treatment the ultimate development of hyperglycemia and ketoacidosis. Antibodies to islet cells are present in a high percentage of cases at the time of clinical onset, and there is a relationship between IDDM and the HLA histocompatibility loci DR3 and DR4 on the sixth chromosome.

IDDM accounts for approximately 10% of patients with diabetes. Ninety percent of patients with diabetes have type II diabetes, noninsulin-dependent diabetes, which is often associated with obesity and is present particularly in patients older than 40 years of age. In this disorder, insulin resistance complicates a quantitatively less severe defect in insulin secretion. Islet cell antibodies are negative and there is no relationship to HLA abnormalities.

A final group of younger patients with hyperglycemia may have maturity onset diabetes of youth. This disorder in some kindreds is related to a glucokinase deficiency and is often associated with an autosomal dominant family history, acanthosis nigricans, and central obesity.

An additional category of impaired glucose tolerance (IGT) is reserved for patients with abnormalities of glucose tolerance testing but without fasting hyperglycemia or other overt evidence of diabetes mellitus. Diabetes can also occur secondary to other endocrine disorders such as acromegaly or Cushing's disease, in which elevations of counterregulatory hormones develop. IGT may initially present during pregnancy in patients with subclinical insulin deficiency as a result of the secretion of placental hormones that antagonize insulin action.

Diagnosis of Diabetes Mellitus

The diagnosis of diabetes mellitus is usually straightforward when fasting hyperglycemia is present. A fasting blood sugar greater than 125 mg/dL (5.0 mmol/L) on two occasions is diagnostic of diabetes mellitus.

There has been substantial revision of the diagnostic criteria for diabetes mellitus prompted by the presence of significant macrovascular and microvascular complications at the time of initial diagnosis using previous criteria. Normal fasting blood sugar is now less than 110 mg/dL, diabetes is more than 126 mg/dL, and a new category, impaired fasting glucose, analogous with IGT with fasting blood sugar of 110 to 126 mg/dL.

Greater sensitivity but decreased specificity is achieved by the use of oral glucose tolerance testing, which is unnecessary if fasting hypoglycemia is documented (29). The patient should have eaten at least 150 g of carbohydrate for 3 days preceding the test. Glucose (75 g) is administered to nonpregnant adults at time zero (1.75 g/kg in children) and serum samples are obtained at zero, 1, and 2 hours. For screening during pregnancy, 50 g of glucose is administered without respect to fasting between weeks 24 and 28 of pregnancy. If blood sugar exceeds 140 mg/dL 60 minutes after the administration of glucose, a formal glucose tolerance test is performed with 100 g of glucose and with serum samples obtained at zero, 1, 2, and 3 hours. For the evaluation of patients with suspected reactive hypoglycemia, 75 to 100 g of glucose is administered and blood sugars obtained at zero, 1, 2, 3, 4, and 5 hours and additionally with spontaneous symptoms. In nonpregnant adults blood sugars greater than 200 mg/dL at either 1 or 2 hours is diagnostic of diabetes. In pregnant women, gestational diabetes is diagnosed if two of the following are achieved: fasting blood sugar is greater than 105 mg/dL, 1-hour blood sugar is 160 mg/dL, 2-hour blood sugar is 165 mg/dL, or 3-hour blood sugar is 145 mg/dL. A summary of the criteria for making the diagnosis of diabetes mellitus is given in Table 16.8.

TABLE 16.8. CRITERIA FOR THE DIAGNOSIS OF DIABETES MELLITUS

	FBS (mg/dl)	1 h (mg/dl)	2 h (mg/dl)	3 h (mg/dl)
Adult				
Diabetes mellitus				
Fasting	>126			
Glucose tolerance test (75 g glucose or 1.75 g glucose/kg body weight)	>126	>200	>200	—
Impaired glucose tolerance (75 g glucose)	>110	140–200		
Gestational diabetes mellitus				
Screening (50 g glucose)		>140		
Definitive (100 g glucose)	>105	>190	>165	>145
Pediatric				
Diabetes mellitus				
Fasting	>140			
Glucose tolerance test (1.75 g glucose/kg body weight)	>140	>200	>200	

Management of Diabetes Mellitus

The contemporary management of diabetes mellitus relies heavily on the use of frequent capillary blood sugars. This process has become extremely convenient for patient and physician with the widespread use of portable glucometers. Several times a day, a patient can obtain his or her own blood specimen obtained with disposable lancets, place the drop of blood on a reagent strip, and read the glucose level on the glucose meters. A log is kept and reviewed with the physician at the next visit. This process will enable the patient to monitor his or her glucose as often as necessary with the

objective of maintaining much tighter glucose control. Although these devices make a significant contribution to diabetic management, it is important to understand the limitations of the individual devices. Some are sensitive to hematocrit and others are limited in the lower or upper range of glucose concentrations.

Urinary glucose monitoring is not sufficiently precise for optimal diabetic control, although the measurement of urinary ketones may be helpful during episodes of acute illness and potential ketoacidosis. These blood sugar results are supplemented by the measurement of one of several glycolated proteins, including glycosylated hemoglobins, hemoglobin A_1c, glycosylated albumin, and fucosamine. These proteins undergo nonenzymatic glycation as a function of integrated blood sugar levels (30). The concentrations of these glycated proteins are a function of the degree of hyperglycemia and the half-life of the involved cells or proteins. Glycosylated hemoglobin, as an example, gives an integrated estimate of blood sugar over the approximately 3-month life span of the erythrocyte. Glycosylated proteins such as hemoglobin A_1c should be monitored every 3 to 6 months in stable patients and every 6 to 12 weeks at the onset of therapy. For unstable diabetics requiring changes in therapy and in pregnancy complicated by diabetes, the monitoring period should be 6 to 12 weeks until the patient is stabilized. A target range of hemoglobin A_1c less than 7.0% is sought to minimize the development and progression of microvascular and possibly macrovascular complications.

Hypoglycemia

Hypoglycemia, either documented or clinically suspected, can be divided into disorders in which symptoms occur when fasting and are often associated with decreased levels of consciousness or mentation (i.e., fasting hypoglycemia) and disorders that typically occur within several hours after eating and are typically associated with sympathomimetic symptoms such as anxiety, palpitations, and anxiety (31).

Fasting hypoglycemia can arise from disorders of insulin secretion, from autoimmune mechanisms involving either the insulin molecule or the insulin receptor, from defects in counterregulatory hormone secretion such as cortisol and growth hormone, and from exogenous means such as insulin or sulfonylurea administration. Measurement of blood sugar and serum insulin levels after an overnight fast and during spontaneous symptoms is essential. With blood sugars below 50 mg/dL, serum insulin should be undetectable and the ratio of serum insulin (uU/ml) to blood glucose (mg/dL) less than 0.3. In addition, measurement of the serum insulin C-peptide or serum proinsulin may be useful in identifying endogenous versus exogenous insulin, because C-peptide and proinsulin are not present in exogenous insulin preparations. Finally, antibodies to insulin may be helpful in identifying both exogenous insulin administration and spontaneous insulin antibodies with partial agonist actions.

Patients with postprandial symptoms may have reactive hypoglycemia. The symptoms are dominated by shakiness, anxiety, fatigue, and hunger rather than loss of consciousness. Postgastrectomy patients and patients with early type II diabetes mellitus with delayed but ultimately normal or even elevated serum insulin levels may present in this manner.

The criteria for the diagnosis of reactive hypoglycemia are problematic, and blood sugars obtained at the time of spontaneous symptoms are more appropriate and preferred when this diagnosis is being considered. The diagnosis of reactive hypoglycemia is frequently made uncritically by patients and physicians alike. Many subjective symptoms that may or may not be related to eating or are associated with food relief are attributed to low blood sugar. This tendency is compounded by the excessive and uncritical use of the glucose tolerance test to establish this diagnosis.

The mean blood glucose nadir in normal subjects after glucose administration is 64 mg/dL, and 10% of normal subjects have blood sugars below 47 mg/dL. Failure to recognize these variances has led to the overdiagnosis of hypoglycemia by glucose tolerance testing and has led many authorities to recommend that glucose tolerance tests not be employed in the evaluation of patients with suspected hypoglycemia.

Islet Cell–Foregut Peptides

Neoplasms of the cells of the pancreatic islet and the neuroectodermal cells of the foregut arise infrequently but may produce distinctive and characteristic endocrine syndromes (32). Benign and malignant tumors of the beta cells of the pancreas produce the insulinoma syndrome, characterized by fasting hypoglycemia and inappropriately elevated serum insulin levels, which is reviewed under fasting hypoglycemia.

Gastrin

Tumors of the delta cells of the islets or of gastrin-secreting cells of the foregut produce the gastrinoma syndrome (Zollinger-Ellison syndrome), characterized by recurrent duodenal and distal small bowel peptic ulcer disease and diarrhea secondary to increased gastric acid secretion and resultant small bowel pH disruption. These tumors are often malignant and may be multifocal. The diagnosis is based on the documentation of elevated fasting serum gastrin levels by immunochemical methods (more than 100 pg/mL) in the setting of increased gastric acid secretion or by the paroxysmal increase (more than 200 pg/mL) in serum gastrin after intravenous secretin administration (2 U/kg i.v.) when basal serum gastrin levels are not clearly elevated.

Vasoactive Peptide

A second syndrome associated with watery diarrhea and resultant hypokalemia as a result of increased diarrheal potassium loss is the watery diarrhea–achlorhydria–hypokalemia syndrome or Verner-Morrison syndrome. Gastric acid secretion is not elevated in this disorder, which results from the excessive secretion of vasoactive intestinal peptide (VIP) from pancreatic islet cell tumors. Serum VIP levels by immunoassay are elevated (more than 50 pg/mL) in the majority of patients with this disorder.

Glucagon

A syndrome characterized by hyperglycemia, anemia, weight loss, and a migratory necrotizing dermatitis skin rash is related to

glucagon-secreting neoplasms of the islet cells of the pancreas. These tumors are often malignant and extensive before diagnosis is established. Serum glucagon levels are elevated (more than 100 pg/mL). The etiology of the rash, which is seen in approximately two thirds of the patients, is unknown.

Somatostatin

Several patients with diabetes, diarrhea, and gallstones secondary to decreased gallbladder contractility have been described secondary to somatostatin-secreting islet cell pancreatic tumors. These symptoms are consistent with the known actions of somatostatin on insulin secretion and gastrointestinal and biliary tract motility. Serum somatostatin levels by immunoassay are elevated.

Carcinoid Syndrome

The syndrome characterized by flushing, diarrhea, asthma, and right-sided endocardial valvular thickening and dysfunction is produced by carcinoid tumors of the chromaffin cells of the distal intestine and bronchial epithelium. These symptoms are related to episodic serotonin and bradykinin secretion. Diagnosis is usually confirmed by documentation of elevated urinary excretion of 5-hydroxyindoleacetic acid (a metabolite of serotonin) or by increased serum serotonin levels.

Other Gastrointestinal Peptide Syndromes

Many other gastrointestinal peptides are known, including pancreatic polypeptide (PP), gastric inhibitory peptide (GIP, bombesin), secretin, and cholecystokinin. No recognized endocrine syndromes are associated with these peptides. Elevation of serum levels by radioimmunoassay may serve as markers in other APUD endocrine syndromes. Serum ß-hCG, α-fetoprotein, and chromogranin-A may also be useful as serum markers of pancreatic-gastrointestinal pancreatic syndromes.

HYPERCALCEMIA/HYPOCALCEMIA METABOLIC BONE DISEASE

Part of "16 - Endocrine Function and Carbohydrates"

Physiology

Serum Calcium Regulation

Maintenance of serum ionized calcium is essential to a variety of critical biological functions, including nerve transmission, muscle contractility, and exocrine secretion. Serum calcium regulation is achieved by the interactions of parathormone, vitamin D and its activated metabolites, and calcitonin. These hormones act at the skeletal, gastrointestinal, and renal levels to maintain a stable ionized serum calcium concentration in the setting of highly variable dietary calcium intake. The major hormones responsible for calcium homeostasis and their modes of action are listed in Table 16.9.

TABLE 16.9. CALCIUM REGULATION BY HORMONE AND SITE OF ACTION

	Bone	Kidney	Intestine
Parathyroid hormone	Increases resorption of calcium and phosphate	Increases reabsorption of calcium, decreases reabsorption of phosphate; increases conversion of 25-OHD_3 to $1,25(OH)_2D_3$; decreases reabsorption of bicarbonate	No direct effects
Calcitonin	Decreases resorption of calcium and phosphate questionable effect on vitamin D metabolism	Decreases reabsorption of calcium and phosphate:	No direct effects
Vitamin D	Maintains Ca^{2+} transport system	Decreases reabsorption of calcium	Increases absorption of calcium and phosphate

From Greenspan FS. *Basic endocrinology*, 3rd ed. Norwalk, CT: Appleton & Lange, 1991, with permission.

Dietary calcium intake is approximately 1,000 mg daily. Eighty percent of dietary calcium is initially absorbed, although 60% is subsequently secreted into the gastrointestinal tract as calcium-containing exocrine pancreatic secretion and sulcus entericus, with a resultant net absorption of dietary calcium of approximately 20%. Calcium is also available from the skeletal system, which serves as a reservoir of calcium through the integrated processes of new bone formation and resorption of existing bone, as part of continuous bone remodeling. Approximately 10,000 mg of ionized calcium is filtered at the glomerulus of the kidney, although efficient proximal and distal tubule resorption limits urinary calcium excretion to only 100 to 300 mg daily.

A fall in serum ionized calcium results in a rise in serum parathormone as a result of the direct negative feedback regulation of parathormone by serum ionized calcium. The rise in serum parathormone simultaneously increases bone resorption of calcium, increases renal calcium tubular resorption, and activates the 1-hydroxylation of 25-hydroxyvitamin D to its active metabolite 1,25-dihydroxyvitamin D, which in turn increases gastrointestinal calcium absorption and amplifies parathormone-mediated bone resorption. These actions serve to return the serum calcium to a normal range. Spontaneous increases in serum calcium have the opposite effects with resultant decrease in serum calcium. Calcitonin appears to have little influence on steady-state serum calcium regulation but moderates the rise in serum calcium during unsteady-state serum calcium changes. Through these hormonal, dietary, and physical activity relationships, both serum calcium and bone mineral remodeling are maintained.

Clinical Disorders

Hypercalcemia

Clinical hypercalcemia results from the pathologic disruption of serum calcium regulation by several clinical disorders (33).

Primary hyperparathyroidism is the most common etiology of hypercalcemia and is characterized by a rise in serum ionized

calcium as a result of increased parathormone secretion, most commonly from a parathyroid adenoma (90%) involving one of the four parathyroid glands, although occasionally as a result of multiple adenomas or hyperplasia involving all four parathyroid glands. The resultant hypercalcemia can produce symptoms of lassitude and increased fatigue. Increased bone resorption can produce demineralization and fractures, and increased urinary calcium secretion can lead to renal calculi. With newer and widespread monitoring of serum calcium by biochemical profile determinations, many older patients are found to have relatively asymptomatic hypercalcemia and hyperparathyroidism. In all these instances, serum parathormone is elevated when determined by appropriate immunoassay techniques (34). This is in contradistinction to other causes of hypercalcemia in which serum parathormone levels are suppressed by nonparathyroid-dependent hypercalcemia. Hyperparathyroidism may occur alone or as a component of MEN 1 and 2, in which case parathyroid hyperplasia involving all four parathyroid glands is seen. In MEN 1 hyperparathyroidism is associated with pancreatic islet cell neoplasms and pituitary adenomas. In MEN 2, hyperparathyroidism is associated with medullary thyroid carcinoma and pheochromocytoma, often bilateral. Both disorders are inherited as autosomal dominant disorders.

Hypercalcemia is one of the most common metabolic complications of malignancy, and several unique mechanisms are recognized. Hematologic malignancies such as multiple myeloma are associated with the production of osteolytic cytokines such as interleukin-1 and tumor necrosis factor. In addition, some lymphomas are associated with increased 1,25-dihydroxyvitamin D levels as a result of tumor-related extrarenal 1-hydroxylase enzyme systems. Some epithelial tumors produce a parathyroid hormone related peptide, or humoral hypercalcemia of malignancy peptide, which is similar in amino acid composition to parathormone in the amino acid terminus of the molecule, although highly distinct in its complete structure (35). Hypercalcemia in these patients is generally severe and associated with decreased serum parathormone levels and with both hypercalcuria and hypochloremia as a result of suppression of parathormone tubular actions on calcium and hydrogen ion resorption.

Sarcoidosis and other granulomatous disorders such as tuberculosis, coccidiomycosis, and berylliosis may be associated with extrarenal 1-hydroxylase enzyme activity, which results in hypercalcemia from increased dietary calcium absorption.

Some drugs may elevate serum calcium, including vitamins A and D, thiazide diuretics, and lithium. The mechanisms are quite distinct. Vitamin D excess leads to increased circulating vitamin D levels, thiazides decrease urinary calcium excretion, and lithium appears to alter transmembrane calcium transport.

The syndrome of familial hypocalcuric hypercalcemia is associated with mild hypercalcemia, normal to slightly elevated serum parathormone concentrations, and low urinary calcium excretion (urinary calcium clearance:urinary creatinine clearance less than 0.01). The disorder is an autosomal dominant abnormality in which an inactivating mutation of a transmembrane calcium transporter is present that is similar to that in patients with lithium-associated hypercalcemia. Care should be exercised to avoid unnecessary parathyroid surgical exploration in these patients.

Hypocalcemia

Hypocalcemia results from failure of the parathormone–vitamin D system to maintain stable ionized calcium levels. It is important to recall that approximately 50% of serum calcium is bound to serum albumin. As a result, a decrease in serum total calcium can arise from decreased serum albumin concentration without any change in biologically active ionized serum calcium. Ionized serum calcium needs to be either measured directly or estimated by adding 0.8 mg/dL of serum calcium to the observed serum total calcium for every 1.0 g/dL decrease in serum albumin concentration.

Hypoparathyroidism results from failure of appropriate parathormone secretion or action. Hypoparathyroidism may follow thyroid or parathyroid surgery in which parathyroid glands are compromised or may arise as idiopathic hypoparathyroidism as a result of immunologic parathyroid dysfunction. The latter may arise as a component of one of two multiple endocrine immunodeficiency syndromes: type I with idiopathic hypoparathyroidism, adrenal insufficiency, and mucocutaneous moniliasis or type II with idiopathic hypoparathyroidism, immunologic thyroid dysfunction (such as Hashimoto's thyroiditis or Graves' disease), and IDDM.

Pseudohypoparathyroidism is an uncommon disorder in which parathormone secretion is normal, but parathormone action is deficient as a result of blunted parathormone receptor activity. Serum parathormone levels are elevated, and the patients do not respond to exogenous parathormone administration by either a rise in serum calcium or urinary cyclic AMP excretion. In some patients, this defect is associated with a defective guanidine-nucleotide coupling protein, which serves to integrate the external parathormone binding receptor with the internally located adenyl cyclase catalytic subunit.

Because magnesium is necessary for both parathormone secretion and action, magnesium depletion is associated with hypocalcemia and decreased serum parathormone levels. Replacement of magnesium corrects both the hypocalcemia and the impaired parathormone secretion, which is not responsive to calcium replacement alone.

Demineralization

The participation of the skeletal system in the maintenance of serum calcium homeostasis gives rise to several significant clinical syndromes in which loss of skeletal mineral content may progress to the point of demineralization and fracture.

Osteoporosis

The most common demineralization disorder is osteoporosis. Osteoporosis is responsible for more than 1.5 million fractures annually in the United States, with annual costs estimated to be more than 15 billion dollars and with a substantial increase in mortality, particularly in relation to hip fractures (36). Most patients have primary osteoporosis related to age, gender, race and menopausal status in which a gradual loss of bone mineral occurs as a result of decreased bone formation relative to bone resorption. The residual bone is qualitatively normal although quantitatively reduced in mass. Additional contributory factors include

limited dietary calcium intake, excessive alcohol or tobacco use, and family history. Osteoporosis can arise secondary to hyperthyroidism, hyperparathyroidism, Cushing's disease (particularly related to steroid use), myeloproliferative disorders, and malabsorptive disorders. As a result of the relatively low levels of negative calcium balance, laboratory parameters are usually quite unremarkable. Contributory secondary factors can be excluded by laboratory means. Bone mineral density studies including dual x-ray absorptiometry (DEXA) and quantitative CT scanning can define bone mineral loss before fracture. Markers of bone formation such as bone-specific alkaline phosphatase and osteocalcin, as well as markers of bone resorption such as urinary N-telopeptide and deoxypyridinoline, are helpful in defining high bone turnover states. These markers may be additionally helpful in monitoring early treatment response and may complement serial DEXA evaluation of bone mineral mass.

Osteomalacia

Defective bone mineralization despite adequate bone matrix formation is seen in osteomalacia. The resultant bone is qualitatively abnormal, with wide osteoid seams of unmineralized bone collagen owing to inadequate bone mineralization. Malabsorption and chronic acidotic disorders are common contributing abnormalities. Drugs such as phenytoin and barbiturates may be factors as a result of hepatic enzyme induction and increased clearance of vitamin D. Inborn errors of vitamin D metabolism and action, such as vitamin D–dependent rickets and vitamin D–resistance rickets, also occur. Although serum calcium levels are well maintained despite the lack of adequate calcium for mineralization, decreased serum phosphorus levels as a result of secondary hyperparathyroidism, increased alkaline phosphatase activity, and low urinary calcium excretion are commonly seen. Definitive diagnosis may require bone biopsy.

Paget's Disease

Paget's disease, a common disorder of older patients is characterized by increased osteoclastic bone resorption, chaotic new bone formation, and elevated serum alkaline phosphatase levels. The latter may lead to the initial diagnosis of Paget's disease in asymptomatic patients without fracture or other skeletal abnormality.

Azotemic osteodystrophy is the complex metabolic bone disease that is commonly seen in patients with chronic renal insufficiency. As dialysis extends life in these patients, the combination of chronic acidosis, secondary hyperparathyroidism as a result of phosphate retention, vitamin D abnormalities as a result of decreased renal 1-hydroxylation of vitamin D, and skeletal aluminum accumulation all lead to a progressively disabling metabolic bone disease.

Metabolic Stone Disease

Renal calculi can arise from increased solute excretion (calcium, uric acid, cysteine), from decreased urinary volume or pH (uric acid, calcium oxalate, calcium phosphate), as a result of loss of biological inhibitors of urinary crystallization (e.g., citrate, pyrophosphate, magnesium), or secondary to underlying obstruction and infection. Evaluation requires exclusion of hypercalcemic disorders; quantitation of urinary calcium, oxalate, and uric acid excretion; and exclusion of underlying obstruction or infection (37).

REPRODUCTIVE ENDOCRINOLOGY

Part of "16 - Endocrine Function and Carbohydrates"

Physiology

Pituitary gonadotropin secretion in both men and women is initiated by the secretion of hypothalamic GnRH into the hypothalamic-hypophyseal portal vasculature in pulsatile fashion at 60- to 90-minute intervals (38). The resulting stimulation of pituitary FSH and LH secretion is further modulated by both positive and negative feedback by gonadal steroids and nonsteroidal peptides such as the inhibins.

In men, LH binds to receptors on the interstitial Leydig cell of the testis where it stimulates the production of testosterone and estradiol. FSH binds to receptors on the Sertoli cell of the seminiferous tubule, leading to the production of inhibin and an androgen-binding protein produces the high local concentrations of testosterone needed for spermatogenesis. Inhibin feeds back through the systemic circulation to downregulate pituitary FSH secretion, whereas estradiol and testosterone feed back to downregulate pituitary LH secretion.

In women, similar but much more dynamic hypothalamic-pituitary-ovarian relationships are present. FSH binds to the granulosa cells of the developing follicle where it stimulates oocyte maturation, estradiol secretion, and inhibin production. Estradiol and inhibin downregulate FSH secretion by feedback inhibition. LH stimulates estrogen and androgen secretion from the interstitial cells of the ovary. These steroid hormones feed back initially to negatively inhibit LH secretion. At a critical threshold of estradiol, however, positive feedback—probably from a different set of hypothalamic neurons—results, with a resultant estrogen-stimulated peak in LH and FSH secretion, which is necessary for ovulation. Thereafter, the site of ovulation undergoes luteinization and becomes morphologically the corpus luteum, the site of progesterone secretion during the second half or luteal phase of the cycle. Under the influence of estrogen, the endometrial lining of the uterus undergoes proliferative changes. These proliferative changes are further modified by progesterone-dependent glycogen deposition during the luteal phase of the cycle to prepare the endometrium for implantation should fertilization occur. If fertilization does not occur, these levels decline and withdrawal of estradiol and progesterone brings about loss of the endometrium as a menstrual cycle preliminary to initiation of the subsequent menstrual cycle.

Fetal testicular androgen secretion is necessary for intrauterine male external genital differentiation. Female external genital differentiation occurs passively in the absence of directive testosterone secretion. Pituitary gonadotropin secretion is low during pediatric development in both sexes. At puberty, an increase in serum FSH and LH secretion occurs as a result of increased hypothalamic GnRH secretion, initially leading to secondary sexual maturation in both sexes and ultimately to reproductive function.

Clinical Disorders

Male

Hypogonadism

Delayed or absent sexual maturation in males can occur as a result of lack of hypothalamic-pituitary gonadotropin secretion, failure of gonadal response to pituitary gonadotropins, or failure of androgen action at the target cell receptor or postreceptor level.

When the defect is at the pituitary-hypothalamic level, low serum FSH and LH levels as well as low testosterone levels will be present. With organic lesions such as pituitary adenomas, craniopharyngiomas, or infiltrative or granulomatous disorders, other pituitary functions such as GH, TSH, and ACTH function may be impaired. The association of isolated hypogonadotropic hypogonadism with anosmia is known as Kallman's syndrome.

Not all abnormalities are organic. A common clinical problem is functional delay, in which ultimately normal physiologic sexual maturation may be delayed into late adolescence or early adulthood.

When the defect is at the testicular level, testosterone levels are low and pituitary gonadotropin levels are elevated as a result of lack of physiologic feedback regulation of both testosterone and inhibin. In addition to traumatic testicular injury, Klinefelter's syndrome, a relatively common chromosomal disorder associated with an additional X chromosome (i.e., 47,XXY or 47,XXY/46,XY versus 46,XY in normal males), may be present.

Sexual Dysfunction

Sexual dysfunction, including decreased libido and failure of erection and ejaculation, often prompts pituitary-testicular endocrine evaluation. Although endocrine factors, including low testosterone states, may be present, other nonendocrine etiologies, including vascular and neurogenic disorders, must be considered. Disorders resulting in serum hPrL elevation may also be present.

Gynecomastia

Gynecomastia is the presence of clinically significant subareolar breast tissue in male subjects. The disorder can arise from testosterone deficiency, inhibition of testosterone action by drugs such as spironolactone or cimetidine, and neoplastic disorders with either ectopic chorionic gonadotropin or estrogen secretion. Pubertal gynecomastia is a common disorder in male adolescents associated with the initiation of physiologic pubertal gonadotropin secretion, which is usually self-limited and not associated with sustained androgen abnormality.

Female

Amenorrhea

Primary amenorrhea, the absence of menses by age 16, may result from endocrine, chromosomal, anatomical, or functional disorders. With endocrine abnormalities, there is usually an associated lack of appropriate physiologic sexual maturation. Turner's syndrome, with a 45/XO karyotype including mosaic variations in which only a portion of cells express a loss of the X chromosome, is usually associated with short stature. Elevated serum FSH levels implicate primary ovarian failure. Decreased serum gonadotropins indicate either organic or functional hypothalamic-pituitary dysfunction, the latter often related to low body weight and vigorous physical activity. Normal gonadotropin levels generally indicate loss of the physiologic cycle necessary for ovulation and menses. In patients with amenorrhea, pregnancy should always be excluded by appropriate hCG determinations.

Secondary amenorrhea is the cessation of periods in a previously menstruating patient. Pregnancy must again be excluded by beta hCG determination. The history of having had periods documents an intact uterus and functional endometrium in the absence of gynecologic intervention. hPrL excess, as a result of either psychotropic medications such as phenothiazines or pituitary microprolactinomas and macroprolactinomas, requires exclusion by serum hPrL measurement. Premature ovarian failure is associated with elevated serum gonadotropin levels. In patients with polycystic ovary syndrome, variable clinical and laboratory evidence of androgen excess is present.

Hirsutism and Virilization

Excessive hair growth in androgen-responsive areas is a common clinical problem (39). Many drugs such as danazol, minoxidil, phenytoin, glucocorticoids, and cyclosporine may produce hirsutism. If other signs of androgen expression in addition to hirsutism are present, such as acne, amenorrhea, temporal hair recession, and masculinization, systemic androgen excess or virilization is present. Laboratory evaluation of hirsutism should include measurement of serum testosterone to document biochemical androgen excess and serum DHEA sulfate, an adrenal androgen, to distinguish adrenal from ovarian etiology. Additional laboratory studies may include measurement of serum free testosterone to distinguish free biologically active versus bound testosterone, androstenedione, cortisol, and pituitary serum FSH and LH levels. In patients with adrenal and ovarian neoplasms, the elevated serum androgen levels also qualitatively fail to suppress after dexamethasone administration.

Patients with latent-onset 21-hydroxylase congenital adrenal hyperplasia may have elevated serum 17-hydroxyprogesterone levels, although often only after Cortrosyn stimulation testing (0.25 mg i.v., serum 17-hydroxyprogesterone at 0 and 60 minutes). Such tests should not be performed during the luteal phase of the menstrual cycle, when 17-hydroxyprogesterone secretion by both the ovarian corpus luteum and the adrenal gland may be present. Androgen suppression with dexamethasone may also be demonstrated.

Menopause

Physiologic cessation of menses occurs in the mid to late 40s in a majority of patients secondary to ovarian senescence. This change is monitored most appropriately by elevation of serum FSH levels as a result of decreased estrogen and inhibin negative feedback suppression. Premature ovarian failure may occur at any age. The disorder may occur either as an isolated defect or as a component of a polyglandular autoimmune syndrome and the diagnosis documented by premature serum FSH elevation.

Laboratory Diagnosis

Figure 16.8 illustrates a diagnostic workup to determine the various causes of amenorrhea.

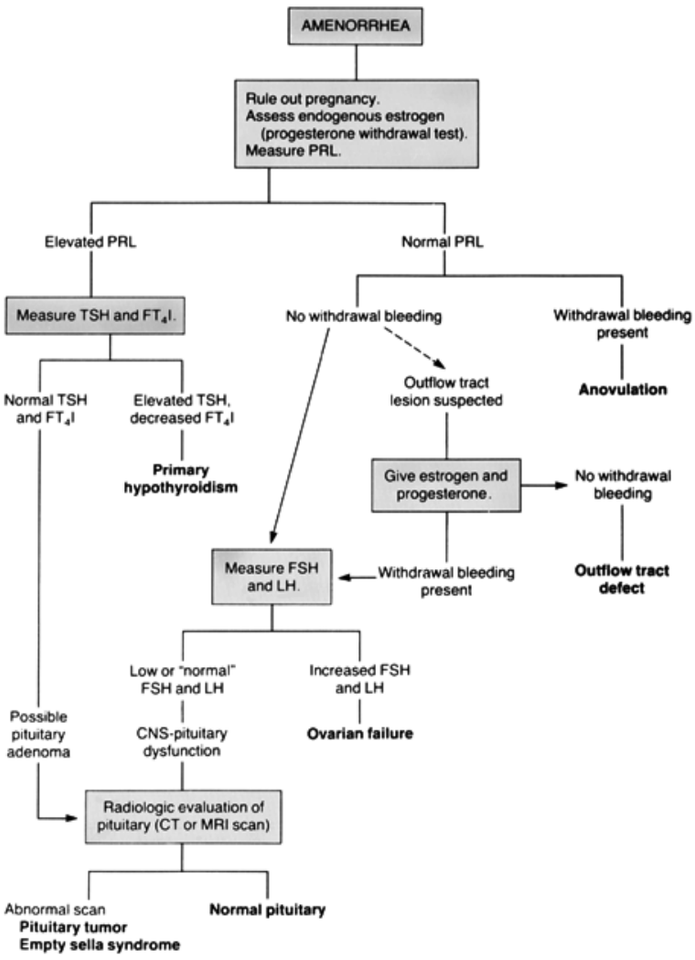


FIGURE 16.8. Potential decision tree for evaluation of amenorrhea. (From Goldfien A, Monroe SE. Ovaries. In: Greenspan FS, ed. *Basic and clinical endocrinology*. Norwalk CT: Appleton & Lange, 1991:442–490, with permission.)

ADRENAL MEDULLA

Part of "16 - Endocrine Function and Carbohydrates"

The adrenal medulla is composed of chromaffin cells, which produce and secrete the vasoactive catecholamines epinephrine and norepinephrine. These cells share a common embryologic origin with the cells of the sympathetic nervous system and with chromaffin cells of extrasympathetic distribution ranging from the carotid bodies to the organ of Zuckerkandl at the bifurcation of the aorta. The symptoms attributable to catecholamine excess are usually related predominantly to their cardiovascular effects.

Clinical Disorders

Pheochromocytomas are tumors arising from the adrenal medulla or sympathetic ganglia that produce hypertension that may be either labile or sustained and is often associated with episodic symptoms of anxiety, headache, tremor, and weight loss. Such tumors may present sporadically or as a component of MEN types IIa and IIb in association with medullary carcinoma of the thyroid and hyperparathyroidism. The diagnosis is often first considered as a result of labile hypertension noted at the time of general anesthesia and surgery.

A high index of suspicion is warranted in the presence of labile hypertension, particularly in younger patients with episodic symptoms including anxiety, tremulousness, and weight loss.

Laboratory Diagnosis

The diagnosis of pheochromocytoma is based on the demonstration of increased concentrations of the catecholamines and their metabolites supplemented by anatomical localization employing CT scanning (40). Because of the labile and variable nature of plasma catecholamines, more integrated urinary collections for measurement of urinary catecholamines and their metabolites, metanephrines and vanillylmandelic acid, are preferred. All collections should include urinary creatinine to ensure adequacy of collection and to permit expression of catecholamines and metabolites per 1,000 mg of creatinine excretion. Such collections may encompass brief periods of spontaneous symptomatology or 24-hour periods with or without spontaneous symptoms. Attention must be given to the exclusion of multiple potential interfering medications. Initial anatomical localization is attempted with abdominal CT imaging. Arteriography or nuclear medicine localization by adrenal imaging with metaiodobenzylguanidine is reserved for those patients who fail to localize with CT scanning or in whom extraadrenal pheochromocytoma is suspected.

REFERENCES

1. Thompson EB. Comment: single receptors, dual second messengers. *Mol Endocrinol* 1992;6:551–556.

2. Williams JA. Mechanisms of hormone secretion and action. In: Greenspan FS, ed. *Basic and clinical endocrinology*, 3rd ed. Norwalk, Connecticut: Appleton & Lange 1991.

3. King AC, Cuatrecasas P. Peptide hormone-induced receptor mobility, aggregation, and internalization. *N Engl J Med* 1981;305:77.

4. Findlay J, Eliopoulos E. Three dimensional modeling of G protein-linked receptors. *Trends Pharmacol Sci* 1990;11:492–499.

5. Gilman AG. G proteins: transducers of receptor-generated signals. *Annu Rev Biochem* 1987;56:615–649.

6. Hildebrandt JD. Role of subunit diversity in signaling by heterotrimeric G proteins. *Biochem Pharmacol* 1997;54:325–339.

7. Wieland T, Chen CK. Regulators of G-protein signalling: a novel protein family involved in timely deactivation and desensitization of signalling via heterotrimeric G proteins. *Naunyn Schmiedebergs Arch Pharmacol* 1999;360:14-26.

8. Simon MI, Strathman MP, Gautam N. Diversity of G proteins in signal transduction. *Science* 1991;252:802–808.

9. Spiegel AM, Simonds WF, Jones TL, et al. Antibodies as probes of G-protein receptor effector coupling and of G-protein membrane attachment. *Biochem Soc Symp*1990;56:61–69.

10. Rink TJ. Receptor-mediated calcium entry. *FEBS Lett* 1990;268:381–385.

11. Dreyer M, Rudiger HW. Genetic defects of human receptor function. *Trends Pharmacol Sci* 1988;9:98–102.

12. Rarick HM, Artemyev NO, Hamm HE. A site on rod G protein alpha subunit that mediates effector activation. *Science* 1992;256:1031–1033.

13. Page MD, Webster J, Dieguez C, et al. A clinical update on hypothalamic-pituitary control. *Acta Neurochir Suppl Wien* 1990;47:48–57.

14. Dieguez C, Page MD, Peters JR, et al. Growth hormone and its modulation. *J R Coll Physicians Lond* 1988;22:84–91.

15. Abboud CF, Laws ER Jr. Diagnosis of pituitary tumors. *Endocrinol Metab Clin North Am* 1988;17:241-280.

16. Goldman MB, Luchins DJ, Robertson GL. Mechanisms of altered water metabolism in psychotic patients with polydipia and hyponatremia. *N Engl J Med* 1988;318:397–403.

17. Merendino JJ Jr, Speigel AM, Crawford JD, et al. Brief report: a mutation in the vasopressin V-2 receptor gene in a kindred with X-linked diabetes insipidus. *N Engl J Med* 1993;328:1538–1541.

18. Vokes TJ, Robertson GL. Disorders of antidiuretic hormone. *Endocrinol Metab Clin North Am* 1988;17:281–299.

19. Bayer MF. Effective laboratory evaluation of thyroid status. *Med Clin North Am* 1991;75:1–26.

20. Wartofsky L. Use of the sensitive TSH assay to determine optimal thyroid hormone therapy and avoid osteoporosis. *Annu Rev Med* 1991;42:341–345.

21. Wenzel KW. Pharmacologic interference with *in vitro* tests of thyroid function. *Metabolism* 1981;30:717–728.

22. Ross DS, Ardisson LJ, Meskell MJ. Measurement of thyrotropin in clinical and subclinical hyperthyroidism using a new chemiluminescent assay. *J Clin Endocrinol Metab* 1989;69:648–684.

23. Aaron DC, Findling JW, Tyrrell JB. Cushing's disease. *Endocrinol Metab Clin North Am* 1987;16:705–730.

24. Flack MR, Oldfield EH, Cutler GB Jr., et al. Urine free cortisol in the high dose dexamethasone suppression test for the differential diagnosis of Cushing's syndrome. *Ann Intern Med* 1992;116:211–217.

25. Carpenter PC. Diagnostic evaluation of Cushing's syndrome. *Endocrinol Metab Clin North Am* 1988;17:445–473.

26. Oldfield EH, Doppman JL, Nieman LK, et al. Petrosal venous sampling with and without corticotropin-releasing hormone for the differential diagnosis of Cushing's syndrome. *N Engl J Med* 1991;325:897–905.

27. New MI, White PC, Paug S, et al. The adrenal hyperplasias. In: Scriver CR, Beaudet AL, Sly WS, et al., eds. *The metabolic basis of inherited disease*, 6th ed. New York: McGraw-Hill, 1989:1881–1917.

28. Young WF Jr, Klee GG. Primary aldosteronism. Diagnostic evaluation. *Endocrinol Metab Clin North Am* 1988;17:367–395.

29. Office guide to diagnosis and classification of diabetes mellitus and other categories of glucose intolerance. *Diabetes Care* 1993;16(Suppl 2):4.

30. Brownlee M, Cerami A, Vlassara H. Advanced glycosylation end products in tissue and the biochemical basis of diabetic complications. *N Engl J Med* 1988;318:1315–1321.

31. Service JH. Hypoglycemia. *Endocrinol Metab Clin North Am* 1988;17:601–616.

32. Vinik AI, Moattari AR. Treatment of endocrine tumors of the pancreas. *Endocrinol Metab Clin North Am* 1989;18:483-518.

33. Bilezikian JP. Hypercalcemia. *Dis Mon* 1988;34:737–799.

34. Marcus R. Laboratory diagnosis of primary hyperparathyroidism. *Endocrinol Metab Clin North Am* 1989;18:647–658.

35. Burtis WJ, Wu TL, Insogna KL, et al. Humoral hypercalcemia of malignancy. *Ann Intern Med* 1988;108:454-457.

36. Riggs RL, Melton LJ. The prevention and treatment of osteoporosis. *N Engl J Med* 1992;327:620–626.

37. Coe FL, Parks JH, Asplin JR. The pathogenesis and treatment of kidney stones. *N Engl J Med* 1992;327:1141–1152.

38. Marshall JC, Kelch RP. Gonadotropin-releasing hormone: role of pulsatile secretion in the regulation of reproduction. *N Engl J Med* 1986;315:1459–1468.

39. Rittmaster RS, Loriaux DL. Hirsutism. *Ann Intern Med* 1987;106:95–107.

40. Sheps SG, Jiang N, Klee GG. Diagnostic evaluation of pheochromocytoma. *Endocrinol Metab Clin North Am* 1988;17:397–414.

17

Electrolytes and Acid-Base Balance

Elizabeth Lee-Lewandrowski

Robert W. Burnett

Kent Lewandrowski

The essential concepts necessary to understand the mechanisms underlying acid-base, fluid, and electrolyte disorders are described here. For more details, an excellent discussion of this topic by Rose (1) is recommended.

- SODIUM, POTASSIUM, AND CHLORIDE
- INTERPRETATION OF ELECTROLYTE DISORDERS
- pH, PCO2, AND BICARBONATE
- INTERPRETATION OF ACID-BASE DISORDERS

SODIUM, POTASSIUM, AND CHLORIDE

Part of "17 - Electrolytes and Acid-Base Balance"

Normal Physiology and Homeostasis

Measurements of electrolyte concentrations are among the most commonly performed tests in the clinical laboratory, both because there is a wide variety of disorders that may cause electrolyte abnormalities and because the administration of intravenous and parenteral fluids necessitates periodic reassessment of serum electrolytes and osmolality.

In a healthy person, homeostatic mechanisms maintain the water and electrolyte composition of the body within relatively narrow limits. The total body water constitutes 60% of the lean body mass and is distributed as intracellular fluid (ICF) and extracellular fluid (ECF). The distribution of water in a lean 70-kg man is shown in Table 17.1. Note that by convention extracellular fluid includes both blood plasma and the fluid inside blood cells.

TABLE 17.1. Water Distribution in a 70-kg Person

Component	Volume (*liters*)
Total body water (60% of lean body mass)	42
Intracellular fluid (66% of total body water)	28
Extracellular fluid (34% of total body water)	14
Interstitial fluid	10.5
Vascular compartment (including blood cells)	3.5

TABLE 17.2. Electrolyte Composition of Plasma

Cations	
Sodium	135–145 mmol/L
Potassium	3.2–5.0 mmol/L
Calcium (total)	2.3–2.6 mmol/L
Magnesium	0.7–1.1 mmol/L
Trace elements	˜ 1 mmol/L
Anions	
Chloride	95–105 mmol/L
Bicarbonate	24–32 mmol/L
Phosphate	0.8–1.5 mmol/L
Other (proteins, anions of organic acids, sulfate)	˜ 21 mmol/L

It is important to understand the difference between the ECF and the effective circulating volume (ECV). The ECV is the portion of the ECF that is actively perfusing tissues. Normally the ECF and the ECV are similar, but in some pathologic states (such as patients with heart failure, ascites, or edema), the effective circulating volume may be decreased despite an increase in the total ECF. These patients behave as if they are volume depleted despite an increase in the ECV. From a physiologic perspective, the ECV, not the ECF, is regulated by the volume receptors. A fall in the ECV results in sodium and water retention.

The ECF and ICF are in osmotic equilibrium, but the electrolyte composition of the two compartments differs. Table 17.2 shows the composition of the major electrolytes in plasma. Sodium is the major cation in the ECF, whereas potassium ion predominates in the cells. Sodium salts are thus the major component of the plasma osmolality.

The regulation of the plasma sodium, osmolality, and the effective circulating volume are complex and closely interrelated. The plasma sodium and osmolality are regulated by water balance, whereas volume is maintained by regulating sodium (2).

The plasma sodium and osmolality are regulated by water intake and renal water excretion. Renal water excretion is regulated by pituitary antidiuretic hormone (ADH), which acts on the renal collecting tubules. A change in the plasma osmolality is detected by hypothalamic osmoreceptors. An increase in the plasma osmolality stimulates thirst and ADH secretion. ADH enhances renal water conservation, resulting in a concentrated urine. Collectively these adaptations restore the plasma osmolality to normal. ADH secretion may also be stimulated by a decrease in the effective circulating volume, which is detected by baroreceptors in the aorta and carotid artery (3). Conversely, a fall in osmolality produces water excretion and a dilute urine.

Sodium and water balance is regulated by the kidney. Several variables influence renal sodium excretion, including the plasma sodium concentration, aldosterone, atrial natriuretic peptide (ANP), as well as other minor factors (2).

Volume regulation is monitored by receptors in the carotid sinuses, the aortic arch, and the juxtaglomerular arterioles. The juxtaglomerular sensors regulate the renin–angiotensin–aldosterone system. Aldosterone acts on the renal collecting tubules to increase sodium resorption and potassium secretion. Sodium resorption leads to an expansion of the ECF volume. The extrarenal sensors regulate ANP, which promotes sodium excretion by the kidney. In addition to aldosterone and ANP, other factors may also influence volume regulation. For example, a fall in the ECV results in stimulation of the sympathetic nervous system with vasoconstriction and an increase in cardiac output and an increase in aldosterone. An expansion of the ECV produces the opposite effect.

Potassium salts are the major intracellular ions, whereas only 2% of the total body potassium is extracellular. The normal plasma potassium is 3.2 to 5.0 mmol/L, whereas the normal intracellular concentration is approximately 150 mmol/L. Redistribution of potassium between cells and the ECF occurs in several

situations. For example, if potassium is lost through the kidney for some reason, the plasma level does not drop as much as might be expected because of the cellular potassium that is available for redistribution. Potassium balance is dependent on intake and renal excretion. A small amount of potassium is also excreted by the gastrointestinal tract, but this fraction may increase markedly in some hyperkalemic states.

Potassium is filtered by the kidney and resorbed in the proximal tubule. However, the distal renal tubule is the major site for regulation of renal potassium excretion. Aldosterone regulates potassium excretion in the distal tubule (aldosterone promotes potassium excretion).

Laboratory Measurement

Flame Photometry

In this method for sodium and potassium, a sample is diluted and introduced into a high-temperature, air-propane flame. In the flame, electrons in sodium and potassium ions are excited and emit light as they return to their ground state. The wavelength of emitted light is characteristic of the particular ionic species, resulting in a high specificity for this method.

Flame photometry is a robust method in that the result is not generally affected by the matrix of the specimen. Thus, aqueous calibration materials may be used, and flame photometry is suitable for analysis of sodium and potassium in serum, plasma, urine, or other body fluids. Most flame photometers use an internal standard such as a cesium or lithium salt, which is added to the diluent to compensate for dilution imprecision and for signal fluctuations owing to changes in the aspiration rate of the sample. Instruments for clinical use are generally dual channel, so that sodium and potassium may be measured simultaneously. Because of the excellent accuracy and precision of flame photometry, this technique is the basis for the reference method for sodium and potassium in serum.

Ion-Selective Electrodes

An ion-selective electrode (ISE) is a device that develops a small voltage when in contact with a solution containing a particular ion. The critical element in the electrode is a thin membrane of a material chosen for its ability to bind one ion species preferentially over other ions in the sample.

Sometimes these membranes are made of glass, for example, the well-known pH electrode and some sodium electrodes. More often, the membrane contains a large organic molecule, called an ionophore, which has different affinities for different ions.

To measure the voltage at the ISE, a second (reference) electrode must be used to complete a circuit. The voltage at the reference electrode must be held constant. Consequently, the reference electrode is isolated from the sample (or calibrator) solution by a “salt bridge” or “liquid junction.” A small voltage also exists at the interface between the salt bridge and the sample solution that is dependent on the composition of both the sample and the bridge solution. This voltage is called the liquid junction potential, *Ej*, and is one of the largest sources of error in ISE methods.

The difference between *Ej* with the unknown solution and *Ej* with the calibration solution is called the residual liquid junction potential. This quantity must be zero for the total cell voltage to reflect accurately the activity of the ion of interest. In other words, the voltage at the ISE itself should be the only voltage that varies when the test solution is changed. Although this is never completely true, the contributions of the bridge solution and the test solution to the liquid junction potential are not equal, and consequently the composition of the bridge solution can be chosen to minimize the variation in *Ej* with different test solutions.

A saturated, or nearly saturated, solution of potassium chloride is one popular choice of bridge solution for general-purpose ISE methods. However, electrode systems designed specifically for measurements on blood or plasma may use a different bridge solution to minimize the residual liquid junction potential in the particular test solution matrix.

Another common source of error in ISE methods is imperfect selectivity of the electrode for the ion of interest. For example, some ISEs intended for chloride measurements also respond, to a lesser but measurable extent, to bicarbonate ions in the sample. The matrix of the samples to be measured determines how selective the ISE must be. For example, for measurements in plasma, which has a sodium concentration approximately 30 times higher than the potassium concentration, a sodium electrode that responds slightly to potassium might be acceptable, whereas a potassium electrode with a significant sodium response could not be tolerated. ISEs are available for determination of sodium, potassium, and chloride as well as several other ions in body fluids.

Coulometry

One of the most accurate methods for serum chloride uses the principle of coulometry. This method measures the number of coulombs (which is proportional to the number of electrons) transferred in an electrochemical reaction. An apparatus called a chloride titrator has been designed for the determination of chloride in serum. A sample is added to an acid diluent in a cell containing a silver anode and a reference electrode, and a constant

current is imposed that generates silver ions at a constant rate. Insoluble silver chloride is formed until all the chloride ion in the sample has been used. Excess silver ions are then generated and detected, either potentiometrically or amperometrically, using a third electrode, and the titration is stopped. Because a constant current is used for the duration of the titration, the number of silver ions generated (which equals the number of chloride ions in the sample) is directly proportional to the duration of the titration, which is usually on the order of 1 or 2 minutes.

This method is both accurate (in the absence of other halides) and precise and is the accepted reference method for serum chloride. Possible interferences include iodide, bromide, and any other anion that will react with the silver ion. Coulometry is also a robust method with respect to the sample matrix and is thus suitable for chloride analysis in serum, plasma, urine, and other body fluids.

Spectrophotometry

Sodium and Potassium

A spectrophotometric method for sodium and potassium measurement has been developed based on the affinity of these ions for particular large synthetic, cyclic organic molecules. These molecules, called macrocyclic ionophores, include crown ethers, cryptanols, spherands, and related structures (4). Like the ionophores used in ISEs, these molecules are designed to selectively complex ions such as sodium and potassium. The difference is that in this case the formation of the complex results in a change in color, which is measured spectrophotometrically. The color change can be accomplished either by extracting the cation of interest along with an anionic dye into a nonaqueous phase containing the ionophore or by using an ionophore that includes a chromophore and is water soluble. In the latter case, binding of the cation by the ionophore results in a spectral shift in the chromophore. The change in absorbance is proportional to the sodium or potassium concentration in the sample.

These methods are suitable for use with very small sample volumes, and because the methods are based on spectrophotometry, they can be readily integrated into multitest analyzers without requiring an additional instrument module for ISEs or flame photometry.

Chloride

Chloride in serum and other body fluids can be measured spectrophotometrically using a reagent containing mercuric thiocyanate and ferric ion. Chloride displaces thiocyanate to form mercuric chloride, and free thiocyanate then combines with ferric ion to form a reddish complex with an absorbance maximum of approximately 500 nm. The sensitivity of the method can be increased by adding mercuric nitrate to the reagent. The free mercuric ions combine with the first available chloride ions so that only excess chloride is responsible for the formation of the colored ferric thiocyanate, as shown below.

$Hg(NO_3)_2 + Cl^- \rightarrow HgCl_2 + 2NO_3^-$

$Hg(SCN)_2 + Cl^-$ (excess) $\rightarrow HgCl_2 + 2SCN^-$

$3SCN^- + Fe^{3+} \rightarrow Fe(SCN)_3$ (reddish color)

This method was once the most widely used method for measuring chloride levels, but it is being replaced by the adaptation of ISE technology in clinical laboratory instruments.

Ion Activity and Ion Concentration

When a sodium ISE is brought into contact with a sample, the electrode signal reflects the activity, not the concentration, of sodium ions in the sample. The activity of an ion may be thought of as an effective concentration that takes into account the influence of all other ionic species present in the solution. The activity is always less than the concentration, and the two quantities are related by the equation

$a = \gamma c$

where a is activity, c is concentration, and γ is called the activity coefficient. The activity coefficient of an ion depends primarily on the total ionic strength of the solution, which for normal plasma is approximately 160 mmol/L.

Activity measurements are important because biochemical reactions involving ions proceed at rates that are dependent on activity and are only indirectly related to concentration. However, because the ionic composition of plasma is closely regulated, the activity coefficients of ions in solution stay within narrow limits, and there is nearly a constant relationship between the activity and the concentration of the ion. Thus, the distinction between activity and concentration is of more theoretical than practical importance in the clinical laboratory.

Plasma Volume Versus Water Volume

The distinction between the total volume of a plasma (or serum) sample and the volume of water contained in the sample is of much more importance than that between activity and concentration because the water fraction of plasma can vary greatly in different patient samples. Normally, plasma contains 93% water by volume. Most of the remaining volume results from dissolved proteins. If the protein concentration is markedly abnormal, the water fraction of plasma will change correspondingly. Triglycerides, either in solution or as chylomicrons, can also significantly reduce the water space when present at high concentrations.

Homeostatic mechanisms for electrolytes respond to the activity of ions in the *water space* of plasma. Hyperproteinemia or hyperlipidemia may result in a marked decrease in the water fraction, but the activity of sodium ions in the water space will not change, and thus the osmoreceptors controlling sodium balance will not be affected. The sodium concentration as measured by an ISE in an undiluted sample will likewise be unaffected. In contrast, a flame photometer or an ISE measurement on a diluted sample will reflect the amount of sodium in a fixed volume of sample. The measured result will be lower because of the volume occupied by the increased protein or lipid resulting in a misleading sodium value referred to as pseudohyponatremia.

ISE methods that use undiluted samples are sometimes called direct methods, whereas those that require dilution of a sample are called indirect. It has been recommended that

methods using undiluted samples be calibrated to agree with methods using diluted samples for specimens of normal protein and lipid composition. This avoids having two different (and overlapping) sets of reference intervals for sodium, potassium, and chloride. Unfortunately, a single method of calibration has not been widely adopted, resulting in some variability of results from different instruments even on samples of normal composition.

Specimen Selection and Specimen Handling

Although sodium, potassium, and chloride ions are inherently stable, there is a large difference between the concentrations of these ions in plasma and in blood cells, especially for potassium. When using plasma samples, care must be taken to avoid hemolysis and to promptly separate blood cells after the specimen is obtained. If visible hemolysis is present in a plasma sample for potassium analysis, either another specimen should be obtained or a note of the presence of hemolysis should be included with the laboratory report.

Serum concentrations of potassium are normally approximately 0.3 mmol/L higher than in plasma because of potassium released from platelets and leukocytes during the clotting process. This difference can be much greater if the platelet or leukocyte count is very high. Many instruments that use ISEs in undiluted samples are capable of using whole blood as the sample. This is advantageous in situations in which short turnaround time is important but has the disadvantage that hemolysis cannot be detected.

Osmolality

The major contributors to the plasma osmolality are sodium salts, glucose, and urea. Plasma osmolality can be measured directly by osmometry (normally 275 to 290 mOsm/kg) or estimated by summing the concentrations of the major solutes (sodium, chloride, bicarbonate, urea, glucose) as shown in the following formula. Because sodium is the major cation, the sum of the anions and cations can be approximated by multiplying the sodium concentration by 2. Thus, only sodium, urea, and glucose need to be measured.

$$\text{Calculated osmolality (mOsm/kg)} \cong 2C_{\text{sodium}}\,(\text{mmol/L}) + \frac{C_{\text{glucose}}\,(\text{mg/dL})}{18} + \frac{C_{\text{urea N}}\,(\text{mg/dL})}{2.8}$$

A more complete discussion of formulas for calculated osmolality may be found in reference 5. The factors 18 and 2.8 are used to convert milligrams per deciliter to millimoles per liter and would not be required for glucose and urea concentrations reported in Sl units.

Unlike sodium and glucose, urea freely crosses cell membranes and does not influence the distribution of water in the body (3). Therefore, the effective plasma osmolality can be calculated by dropping the term for urea and may be more useful than the measured osmolality for evaluating disorders of fluid and electrolyte balance in patients with renal failure and an elevated plasma urea.

INTERPRETATION OF ELECTROLYTE DISORDERS

Part of "17 - Electrolytes and Acid-Base Balance"

Hyponatremia

Hyponatremia is defined as a plasma sodium less than 135 mmol/L. Usually hyponatremia is associated with hypoosmolality. Hyponatremia results in water movement into cells, producing cellular swelling with cerebral edema and metabolic encephalopathy. The severity of the these symptoms depends on the degree of hyponatremia and the rate of change of the plasma sodium. Acute hyponatremia less than 125 mmol/L may produce severe symptoms, whereas an equivalent level of chronic hyponatremia may be asymptomatic.

An approach to the differential diagnosis of hyponatremia is shown in Fig. 17.1. In principle, hyponatremia may result from loss of sodium or an increase in body water. The kidney will normally excrete excess water to prevent hyponatremia. Consequently, most patients with hyponatremia exhibit abnormal water excretion as occurs in patients with effective volume depletion, renal failure, inappropriate ADH secretion, and other causes as shown in Fig. 17.1 (6). Three categories of hyponatremia may be distinguished based on the plasma osmolality.

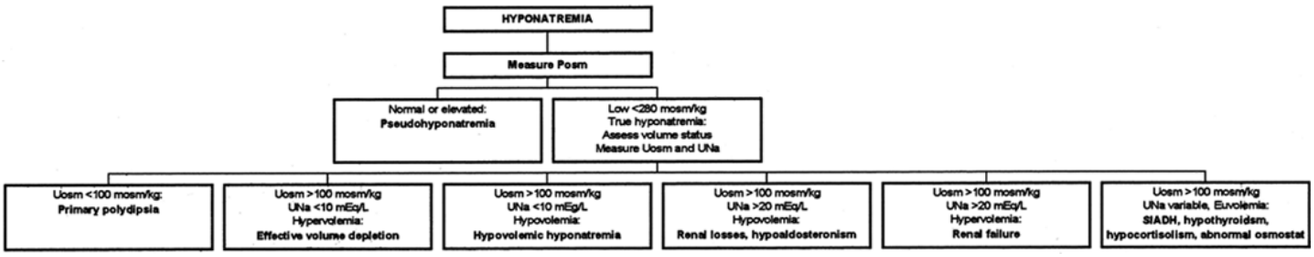


FIGURE 17.1. Selective differential diagnosis and laboratory approach to hyponatremia. Posm: Plasma osmolality; Uosm: Urine osmolality; UNa: Urine sodium; SIADH: Syndrome of inappropriate antidiuretic hormone release.

Hyponatremia with a Low Plasma Osmolality

Effective circulating volume depletion stimulates thirst and secretion of ADH. The result is hypoosmolality and hyponatremia. As described previously, some of these patients have hypervolemia (heart failure, ascites, edema) despite a low effective circulating volume. Other cases exhibit hypovolemia (vomiting, diarrhea, skin losses).

Severe renal failure may result in water retention. Sodium losses may produce hyponatremia as occurs in hypoaldosteronism and salt-losing syndromes (polycystic kidney disease, tubulointerstitial diseases). Hypothyroidism and hypocortisolism produce effective volume depletion, resulting in water retention by the kidney.

The syndrome of inappropriate secretion of ADH (SIADH) is a common cause of hyponatremia especially in hospitalized patients. SIADH occurs in a variety of conditions, including ectopic ADH production by tumors (e.g., bronchogenic carcinoma especially of the small cell type), various pulmonary disorders (e.g., tuberculosis, pneumonia), central nervous system (CNS) diseases (e.g., brain tumors, CNS infections, trauma, hemorrhage, or infarction), a variety of drugs, and miscellaneous conditions (3). The diagnosis of SIADH is based on clinical criteria including hyponatremia, hypoosmolality with normovolemia, a urine osmolality greater than 100 mOsm/kg with an inappropriate urine sodium concentration, normal renal, adrenal, and thyroid function in a patient not taking diuretics.

Rarely, hyponatremia is observed in a condition called primary polydipsia or in patients with abnormal osmoregulator function. Hyponatremia owing to deficient oral salt intake is also quite rare.

Hyponatremia with an Elevated Plasma Osmolality

Hyponatremia with an elevated plasma osmolality may be observed in patients with hyperglycemia or after administration of

mannitol, both of which draw water out of cells, resulting in hyponatremia. Hyperosmolality owing to an elevated plasma urea is not associated with hyponatremia because urea does not alter the distribution of water in the body

Pseudohyponatremia

A condition called pseudohyponatremia may occur with serum or plasma samples containing large amounts of lipids (hyperlipidemia) or proteins (e.g., paraproteinemia). This condition is a laboratory artifact, which was discussed previously, and has no clinical significance other than calling attention to the underlying cause of the hyperlipidemia or hyperproteinemia. The osmolality is unaffected because osmometers measure only the solutes in the plasma water.

Laboratory Differential Diagnosis

A selective differential diagnosis and approach to the evaluation of hyponatremia is shown in Fig. 17.1. The plasma osmolality may be used to exclude pseudohyponatremia. In patients with true hyponatremia, the next step is to assess the patient's volume status based on clinical criteria and to measure the urinary sodium and osmolality (or specific gravity). In patients with renal failure, the plasma osmolality may be misleading because of the presence of urea. In these patients, the plasma osmolality is best estimated by the formula: plasma osmolality = 2 × plasma sodium. The urine osmolality is useful to assess renal water excretion. A low urine osmolality (or a specific gravity less than 1.004) is consistent with primary polydipsia, whereas an osmolality greater than 100 mOsm/kg is indicative of abnormal water excretion (6). Measurement of the urine sodium is useful to assess patients with an abnormal renal water excretion as shown in Fig. 17.1. A low urinary sodium concentration (less than 10 mmol/L) is consistent with extrarenal disorders or effective circulating volume depletion. In contrast, renal diseases that produce hyponatremia exhibit a high urinary sodium concentration (more than 20 mmol/L).

Hypernatremia

Hypernatremia is defined as a serum sodium greater than 150 mmol/L. Patients with hypernatremia characteristically exhibit an increased osmolality. The symptoms of hypernatremia are mainly neurologic (metabolic encephalopathy) and depend on the severity of the associated hyperosmolality and the time interval over which the hypernatremia has developed. Acute hypernatremia may be symptomatic at a plasma sodium of 160 mmol/L, whereas in chronic hypernatremia, symptoms may not be evident at sodium concentrations significantly higher than this level.

In theory, hypernatremia can result from water losses or from a gain in sodium. Ordinarily the body responds to hyperosmolality by conserving water (ADH secretion) and by increasing water intake. Consequently, hypernatremia generally will not occur unless the thirst response is abnormal or the patient cannot increase water intake (e.g., a patient in coma or in infants). A selective differential diagnosis of hypernatremia is shown in Fig. 17.2.

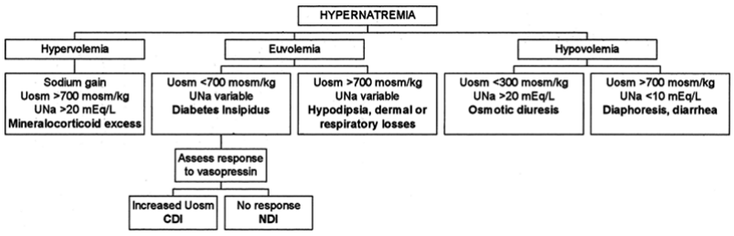


FIGURE 17.2. Selective differential diagnosis and laboratory approach to hypernatremia. Uosm: Urine osmolality; UNa: Urine sodium; NDI: Nephrogenic diabetes insipidus; CDI: Central diabetes insipidus.

Hypothalamic disorders may impair the thirst mechanism or produce abnormal osmoreceptor function (hyperaldosteronism, Cushing's disease), resulting in hypernatremia. Hypodipsia may be observed in patients with various CNS disorders. Rarely, excessive sodium intake may produce hypernatremia as may occur after infusion of hypertonic solutions or in infants fed hypertonic formulas.

Diabetes insipidus (DI) is an important cause of hypernatremia that results either from a defect in ADH secretion (central DI) or from an impaired renal response to ADH (nephrogenic DI). In either case, the result is an abnormal renal concentrating ability. Usually patients do not develop hypernatremia unless the thirst mechanism is abnormal the patient cannot increase oral water intake (6). DI usually presents with polyuria or polydipsia. The differential diagnosis of central DI includes a large number of CNS disorders, such as CNS ischemia or trauma, or may occur as an idiopathic form. Other causes include brain tumors, infections, and neurosurgery. Nephrogenic DI may occur in association with renal failure, hypokalemia, hypercalcemia, various drugs (diuretics, lithium), as a congenital form, or from other causes.

Laboratory Differential Diagnosis

Most adults with hypernatremia exhibit altered mental status or less commonly an abnormal thirst response. Hyponatremia may also be observed in infants. In patients with hypernatremia, the urine should be concentrated (more than 700 mOsm/kg) with a high specific gravity (more than 1.023). If this is not found, DI should be considered. The evaluation begins by assessing the patient's volume status followed by measurement of the urine osmolality and urine sodium. Central DI can be differentiated from nephrogenic DI by observing the response of the urine osmolality to ADH administration. A patient with central DI will exhibit an increase in the urine osmolality.

Hypokalemia

Signs and symptoms of hypokalemia generally do not occur until the plasma potassium has fallen below 3.0 mmol/L (7) and include neuromuscular disturbances (muscle weakness, paralysis), impaired urinary concentrating ability, cardiac arrhythmia, and increased digitalis sensitivity. The electrocardiogram may show flat or inverted T waves, prominent U waves, or S-T segment depression (8). Hypokalemia may result from decreased potassium intake; redistribution of potassium into cells; gastrointestinal, renal, or sweat losses; and hemodialysis. An approach to the evaluation of hypokalemia and a selected differential diagnosis is shown in Fig. 17.3.

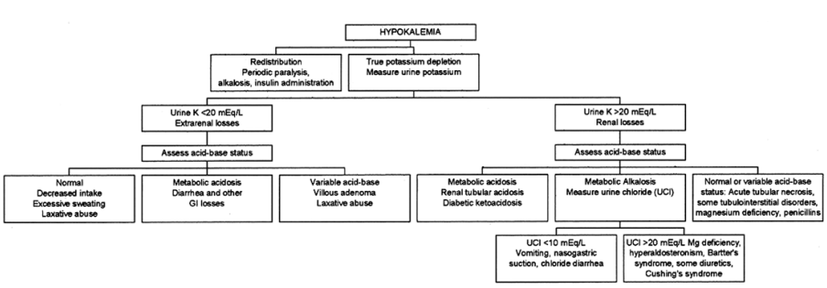


FIGURE 17.3. Selective differential diagnosis and laboratory approach to hypokalemia. UCl: Urine chloride; Mg: Magnesium.

A transient hypokalemia may occur by shifting potassium into cells as may occur in alkalosis, insulin, catecholamines, and increased ß-adrenergic activity (severe stress). Treatment of severe anemia may result in hypokalemia owing to uptake of potassium by newly formed cells. Periodic paralysis is a rare disorder that can be familial or acquired and can produce intermittent hypokalemia with muscle weakness or paralysis presumably owing to cellular uptake of potassium ions.

Hypokalemia may occur in patients on intravenous fluids without potassium supplementation or, rarely, owing to severe dietary deficiency (7).

Extrarenal potassium losses produce hypokalemia owing to potassium lost in gastrointestinal secretions (diarrhea, vomiting, nasogastric suction, Zollinger-Ellison syndrome) or through the skin (e.g., sweating, burns). Hypokalemia resulting from urinary losses may occur in hyperaldosteronism, Bartter's syndrome, acute tubular necrosis, tubulointerstitial disorders, renal tubular acidosis (RTA), and hypomagnesemia or in association with some drugs such as amphotericin or diuretics.

The initial evaluation of hypokalemia should consider the patient's potassium intake, potassium losses, and underlying conditions that may predispose to hypokalemia. Serum electrolytes, arterial blood gases, and urinary electrolytes may be useful in arriving at a correct diagnosis as shown in Fig. 17.3. Further tests (e.g., aldosterone, plasma renin activity, serum magnesium) can be requested when appropriate. The urinary potassium concentration is helpful because it will differentiate renal potassium losses from other causes of hypokalemia.

Hyperkalemia

The signs and symptoms of hyperkalemia include muscle weakness and abnormal cardiac conduction. Muscle weakness generally does not occur until the serum potassium is more than 7 mmol/L. Electrocardiographic changes can be seen with plasma levels greater than 6 mmol/L. More severe hyperkalemia (in excess of 8.0 mmol/L) is associated with a sine wave pattern on the electrocardiogram with imminent ventricular fibrillation and cardiac arrest (7).

An approach to the evaluation of hyperkalemia and a selected differential diagnosis is shown in Fig. 17.4. Hyperkalemia can result from increased intake, redistribution, and decreased renal potassium excretion. Normally the kidney excretes excess potassium and consequently chronic hyperkalemia implies impaired urinary excretion.

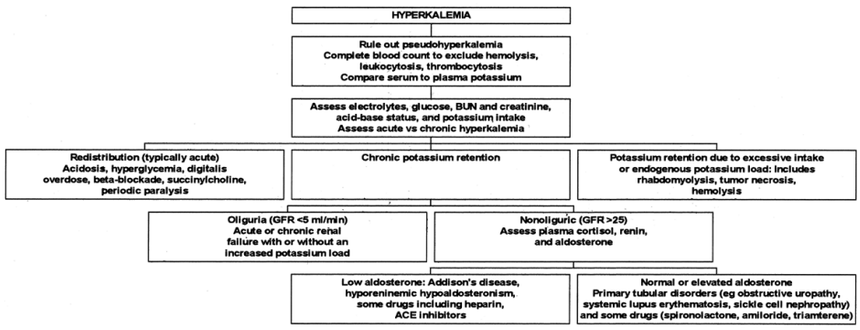


FIGURE 17.4. Selective differential diagnosis and laboratory approach to hyperkalemia. GFR: Glomerular filtration rate; ACE: Angiotensin converting enzyme.

Acidemia promotes the exchange of plasma hydrogen ions for cellular potassium and thus is a cause of hyperkalemia. The resulting hyperkalemia is greatest with hyperchloremic metabolic acidosis (8).

Hyperkalemia may occur despite a deficit of total body potassium stores. For example, in patients with diabetes mellitus, insulin deficiency enhances uptake of both glucose and potassium by cells. Rhabdomyolysis, or any state of tissue breakdown (e.g., trauma, hemolysis, or tumor lysis after chemotherapy) can release cellular potassium and cause hyperkalemia. Several drugs promote potassium loss from cells such as succinylcholine, whereas digitalis overdose interferes with the cellular sodium/potassium ATPase pump. Hyperkalemic periodic paralysis is a rare autosomal dominant condition characterized by episodes of paralysis with acute hyperkalemia.

Several conditions produce hyperkalemia by impairing renal potassium excretion including renal failure and hypoaldosteronism. The differential diagnosis of hypoaldosteronism is extensive and includes hyporeninemic hypoaldosteronism, several drugs (nonsteroidal antiinflammatory drugs, angiotensin-converting enzyme inhibitors, cyclosporine), adrenal insufficiency, congenital adrenal hyperplasia, and aldosterone resistance (potassium-sparing diuretics, pseudohypoaldosteronism, cyclosporine) (9).

Pseudohyperkalemia may occur in samples with hemolysis or in specimens with leukocytosis or thrombocytosis that have been allowed to stand for an excessive period before processing

Hyperchloremia and Hypochloremia

Chloride is quantitatively the most important extracellular anion. As such, abnormalities in the serum chloride may occur in a variety of settings as a component of acid-base, fluid, or electrolyte disorders. From a clinical perspective, the abnormality in the serum chloride itself is of little concern. Rather, attention is focused on the underlying disorder causing the hyperchloremia or hypochloremia. A selected differential diagnosis is shown in Table 17.3.

Of the various conditions listed in Table 17.3, only one, bromism, merits specific discussion here. Elevations of serum bromide once occurred in patients ingesting bromide ion in the

TABLE 17.3. Differential Diagnosis of Hyperchloremia and Hypochloremia

Hyperchloremia
Metabolic acidosis associated with gastrointestinal bicarbonate loss
Renal tubular acidosis
Hypoaldosteronism
Respiratory alkalosis
Hyponatremia with sodium losses in excess of chloride
Bromism
Administration of NH_4Cl, amino acids (hyperalimentation), saline, carbonic anhydrase inhibitors
Some cases of hyperparathyroidism
Hypochloremia
Gastrointestinal chloride losses (vomiting, nasogastric suction)
Anion-gap metabolic acidosis
Tubulointerstitial disorders
Hyperaldosteronism
Compensated respiratory acidosis
Metabolic alkalosis
Hyponatremia
Adrenocortical insufficiency

Adapted from Wallach J. *Interpretation of diagnostic tests: a handbook synopsis of laboratory medicine*. 4th ed. Boston: Little, Brown, 1986.

form of over-the-counter drug formulations or as an anticonvulsant. These drugs are not often used today, and, as a result, the incidence of bromism has declined. The one case we did observe was a foreign patient visiting the United States. Aside from its toxic properties, bromide ion causes falsely elevated chloride levels with ISE, coulometric, and, to a lesser degree, mercurimetric methods. Other conditions listed in Table 17.3 are discussed elsewhere in this chapter.

pH, PCO_2, AND BICARBONATE

Part of "17 - Electrolytes and Acid-Base Balance"

Physiology of Acid-Base Balance

An optimal pH is important for the functioning of cellular enzymes. The pH of the blood is normally maintained (Table 17.4) by a combination of the body's buffering systems and by renal and respiratory regulatory mechanisms. As a general rule, acidemia below pH 6.8 or alkalemia above pH 7.8 is not compatible with life. Less dramatic alterations in acid-base balance may cause significant morbidity and mortality, especially when superimposed on other serious illnesses.

TABLE 17.4. Reference Intervals for Arterial Blood pH and Gases

pH	7.36–7.44
Po_2	80–100 mm Hg
Pco_2	35–45 mm Hg
Oxygen saturation	>95%
Bicarbonate	22–26 mmol/L
Total CO_2	23–27 mmol/L
Base excess	-2 to +2 mmol/L

It is useful in discussing acid-base balance to distinguish between abnormalities in blood pH (acidemia and alkalemia) and conditions that tend to produce acidemia and alkalemia (acidosis and alkalosis). As are discussed, different types of acidosis and alkalosis can be present simultaneously, and the resultant blood pH may be either high, low, or normal.

Acidemia is defined as an arterial pH below 7.36 and alkalemia as a pH above 7.44. Acid-base disorders affect a variety of metabolic processes, ultimately leading to impairment of essential functions, but the clinical signs and symptoms of acidosis and alkalosis are relatively nonspecific. Therefore, the diagnosis requires laboratory confirmation by measurement of blood pH and gases.

Carbon dioxide is produced by cellular respiration and is the most important substance to challenge the acid-base balance of the body. Acid is produced by the association of CO_2 with water, forming carbonic acid, which dissociates to hydrogen ion and bicarbonate.

$CO_2 + H_2O$ [left harpoon over right harpoon] H_2CO_3 [left harpoon over right harpoon] $H^+ + HCO_3^-$

Almost all the CO_2 produced (volatile acid) is eliminated in the lungs. An additional acid load is created by the metabolism of proteins and the hydrolysis of phosphoester bonds, resulting in the production of nonvolatile acids, predominantly sulfuric and phosphoric acids. Other metabolic processes produce alkali in the form of bicarbonate, mainly by the oxidation of organic anions. Consequently, the body requires a finely regulated but flexible mechanism to cope with acid or base challenges.

Body Buffering Systems

The Henderson-Hasselbalch equation describes the relationship between pH and the acidic and basic forms of a buffer. The general relationship is

$$\mathrm{pH} = \mathrm{pK'} + \log\left(\frac{C_{\text{basic form of buffer}}}{C_{\text{acidic form of buffer}}}\right)$$

where C represents concentration and pK' represents the negative logarithm of the equilibrium constant for the buffer pair.

The principal buffering system of plasma is the bicarbonate/carbonic acid system. Carbonic acid in plasma is in equilibrium with dissolved CO_2. This in turn is directly measurable as PCO_2, and the two quantities are related by the solubility coefficient of CO_2 in plasma, which is 0.03 mmol + L^{-1} + mm Hg^{-1}.

CCO_2 (dissolved) = 0.03 PCO_2

For the bicarbonate/carbonic acid system, the Henderson-Hasselbalch equation can thus be written as

$$\mathrm{pH} = 6.1 + \log\left(\frac{C_{HCO_3^-}}{0.03\ P_{CO_2}}\right)$$

where $CHCO_3^-$ is expressed in millimoles per liter and PCO_2 is expressed in millimeters of mercury.

Normally the ratio of bicarbonate to dissolved CO_2 is 20/1. Changes in either the bicarbonate concentration or the PCO_2 will affect the ratio and thus change the pH of the blood. Blood gas analyzers measure pH, PCO_2, and PO_2, and the Henderson-Hasselbalch equation is used by the microprocessor built into the analyzer to calculate other quantities such as bicarbonate and total CO_2 concentrations.

The Henderson-Hasselbalch equation could also be written for any of the other buffer pairs in the body. Because all buffers must be in equilibrium, acid-base status can be expressed in terms of only the carbonic acid–bicarbonate system.

Respiration and Acid-Base Regulation

The rate of alveolar ventilation affects acid-base status through changes in the PCO_2.The rate of respiration is influenced by the arterial PO_2 and pH. Arterial pH is detected by chemoreceptors in the aortic arch and carotid bodies (peripheral) and in the brainstem (central). Acidemia stimulates respiration, resulting in a fall in the PCO_2 and an increase in arterial pH. Conversely, alkalemia induces hypoventilation with CO_2 retention and a decrease in pH toward normal. This forms the basis of the respiratory compensatory mechanism. Conversely, a pathologic alteration in alveolar ventilation may itself be the cause of an acid-base disorder by inducing CO_2 retention (respiratory acidosis) or excessive CO_2 elimination (respiratory alkalosis).

Renal Acid-Base Regulation

The kidney plays an important role in acid-base balance. The kidney regulates bicarbonate resorption and acid excretion, primarily in the form of ammonium ion. Plasma bicarbonate is filtered by the glomerulus and is resorbed to maintain acid-base balance. In alkalotic states, the kidney can also excrete bicarbonate to compensate for an elevated pH.

Normally the body produces a net excess of acid (50 to 100 mmol/day) (10). However, the minimum urinary pH is approximately 4.5 (9), and at this pH, only a small number of free hydrogen ions can be excreted. Consequently, the kidney requires an alternate mechanism to excrete the normal daily acid load. The two most important species in this regard are phosphate (HPO_4^{2-}) and ammonia, both of which can combine with hydrogen ions secreted by renal tubular cells and are excreted in urine as $H_2PO_4^2$ and NH_4^+. The amount of phosphate available for combination with H^+ cannot be increased significantly from its normal level; however, the production of ammonia can be greatly increased. Ammonia produced by the renal tubular epithelium diffuses into the tubular lumen where it combines with hydrogen ions to form ammonium ion, which is charged and cannot diffuse back into the cell. The hydrogen ions are thus trapped and are excreted in the urine. Therefore, ammonia formation is the most important mechanism of the kidney to increase acid excretion.

Measurement of pH, PCO_2, and Bicarbonate

Analytical Methods

Blood gas analyzers measure three quantities, pH, PCO_2, PO_2. All other quantities including bicarbonate, total CO_2, dissolved CO_2, and base excess are calculated by the microprocessor in the instrument.

pH Measurement

The method for measuring the pH of blood is essentially the same as that in any other liquid. The pH is determined by measurement of the voltage in an electrochemical cell containing a glass electrode that is selective for hydrogen ions. The principles of ISE measurements were discussed previously, so only those factors unique to blood pH measurement are included here.

The reference electrode is usually a saturated calomel electrode and the bridge solution is usually saturated, or nearly saturated, potassium chloride, although other solutions are sometimes used.

Calibration of the system is performed with two phosphate buffers that are traceable to the primary pH standards certified by the National Institute of Standards and Technology. The buffers most often used have a pH between 6.84 and 7.39 at 37°C. This narrow range is appropriate because the physiologic range of the blood pH is also quite narrow.

Blood pH, PCO_2, and PO_2 are all temperature dependent. By convention, measurements are made at 37.0°C, and the temperature is controlled (to within ±0.05°C) to ensure accurate results. This is important because the temperature coefficient of blood pH is approximately 10 times larger than that of the phosphate buffers used as calibrators. Therefore, if the instrument temperature drifts, it will not be detected (or compensated) by a shift in the pH of the calibrators and will result in a measurement error.

The other source of significant systematic bias in blood pH measurements is residual liquid junction potential. As explained earlier, this error arises in potentiometric measurements when the potential at the liquid junction between the salt bridge and the solution is not the same for the calibrators and the test solutions. In the case of blood pH, two different sources of bias arise at the liquid junction.

The first is known as a *suspension effect* and is a result of the presence of red cells at the interface with a concentrated KCl bridge solution. A bridge of 4 mol/L KCl causes the pH of whole blood to appear approximately 0.01 lower than the pH of the corresponding plasma. This effect can be eliminated by changing the bridge solution to sodium formate, but it is often simply ignored.

The second source of bias is a classic example of a residual liquid junction potential. The matrix of the phosphate calibration buffers is much different from whole blood. In particular, the ionic strength of plasma is much greater and the mix of ionic species is different. The liquid junction potential between these two kinds of solutions and a concentrated KCl salt bridge is not the same, and the difference makes the apparent pH of whole blood samples approximately 0.03 lower than would be the case in the absence of this bias.

Both of the systematic errors described above exist in commonly used instruments for blood pH measurement. Eliminating these errors is technically feasible, but the bias is nearly constant and has been included in the reference interval used for clinical interpretation of blood pH. The current consensus is that the costs of such a change outweigh the benefits (11).

PCO_2 Measurement

The partial pressure of CO_2 in whole blood (PCO_2) is measured using an electrode system first described by Stow et al. (12) and modified by Severinghaus and Bradley (13). The system consists of a pH-sensitive glass electrode in contact with a weak bicarbonate solution. This solution is separated from the sample by a membrane permeable to CO_2. When CO_2 from the sample diffuses across the membrane into the bicarbonate solution, the pH

change is sensed by the glass electrode and the change in pH is related to the P_{CO2} in the sample, which can be either a gas or a liquid. There is no bias owing to sample matrix or residual liquid junction potential because the sample is isolated from the cell by the gas-permeable membrane.

Blood gas partial pressures are quite dependent on the sample temperature. The solubility of gases in aqueous solutions increases as the temperature of the solution is lowered. Therefore, as temperature is increased, gases come out of the solution and the partial pressures of the gases increase. For CO_2, this increase is approximately 5% per degree Celsius. Therefore, the temperature of blood gas electrode systems is always held constant, conventionally at 37.0°C.

Calibration of Blood Gas Electrode Systems

The measurement of PCO_2 and PO_2 is performed simultaneously in most blood gas instruments. PO_2 measurement is discussed in Chapter 18. Calibration of both electrodes is done at two points, usually with gas mixtures containing known percentages of O_2 and CO_2. The two points chosen for CO_2 are usually 5% and 10%. The nature of the oxygen electrode permits a zero adjustment with an oxygen-free gas, and the second calibration point is usually either 12% or 20% O_2. These are merely conventions; other gas compositions can be used, but the composition must be accurately known, typically to within ±0.03%.

Although the composition of a gas mixture is stated as a percentage of each component, the partial pressures of O_2 and CO_2 in blood are expressed in units of millimeters of mercury (mm Hg) or kilopascals (kPa) (1 mm Hg equals 0.1333 kPa). To calculate the partial pressure of CO_2 and O_2 in a calibration gas, the vapor pressure of water (PH_2O) and the total (barometric) pressure (Ptotal) must also be known. The vapor pressure of water is needed because the (dry) gas mixture is saturated with water at 37°C before it reaches the electrodes. The partial pressure of gas G (PG) is calculated as

$$P_G = \frac{X_G}{100}(P_{total} - P_{H_2O})$$

where XG is the percentage of gas G in the dry gas mixture. The vapor pressure of water at 37°C is 47 mm Hg. Many blood gas analyzers have an automatic calibration feature with a built-in barometer, and the above calculation is performed internally.

Total CO_2

Total CO_2 is defined as the sum of the concentrations of the various CO_2-containing species in plasma (dissolved CO_2 gas, carbonic acid, bicarbonate ion, carbonate ion, and CO_2 associated with proteins as carbamates). Normally, bicarbonate ion and dissolved CO_2 account for approximately 94% and 5%, respectively, of the total CO_2. The total CO_2 is measured by adding acid or base to the sample to convert essentially all the CO_2 forms to either dissolved CO_2, bicarbonate, or carbonate ions and then measuring the concentration of that particular species.

The classic, and still the most accurate, methods begin with acidification of the sample to convert all forms of CO_2 to dissolved CO_2. Several techniques have been used to quantitate the CO_2 gas produced.

1. Manometry. This is an accurate method but requires special apparatus and is time-consuming.
2. Colorimetry, with the CO_2 gas being dissolved in a weak buffer containing an indicator such as phenolphthalein. This was very widely used in continuous-flow methods but is less common today.
3. A CO_2 electrode, whereby the increase in PCO_2 is measured directly using the same type of electrode used in blood gas analyzers.
4. Thermal conductivity, whereby a gas stream containing the CO_2 is passed over a "hot-wire" detector of the type used in gas chromatography.
5. Titrimetry, whereby the CO_2 released is recaptured in a basic solution, and the solution is titrated back to its original pH with standardized sodium hydroxide.

Another popular methodological principle begins by buffering the sample so that virtually all the CO_2 is in the form of bicarbonate, which is then measured enzymatically by the following reactions.

$$\text{Bicarbonate} + \text{Phosphoenol pyruvate} \xrightarrow{\text{PEPC}} \text{Oxaloacetate} + \text{Phosphate}$$

$$\text{Oxaloacetate} + \text{NADH} \xrightarrow{\text{MDH}} \text{Malate} + \text{NAD}$$

where PEPC is phosphoenolpyruvate carboxylase and MDH is malate dehydrogenase. The disappearance of NADH (nicotinamide adenine dinucleotide, reduced form) can be monitored spectrophotometrically at 340 nm, or this reaction can be coupled to another redox indicator, which can be monitored in the visible region of the spectrum.

A third method that has been used is based on raising the pH of the sample to shift the equilibria toward carbonate ion, which is then measured directly using a carbonate ISE.

Although most of the methods described above are capable of good precision, there is a need for a practical reference method. The manometric method has both high accuracy and precision but requires a specialized apparatus and careful technique. A titrimetric method based on capturing CO_2 in a $BaCl_2$ solution after liberation from an acidified sample has been proposed as a candidate reference method (14).

Calculations and Derived Quantities

Total CO_2 and Bicarbonate

Although the total CO_2 can be measured in serum, it is also common to calculate it in conjunction with direct measurement of pH and PCO_2 in whole blood. The relationship needed to calculate bicarbonate concentration is the Henderson-Hasselbalch equation, which may be written in the form

$$\log CHCO_3^- = pH - pK' + \log PCO_2 + \log \alpha$$

where pK′ for the equilibrium between the dissolved CO_2 and bicarbonate is 6.105 and α, the solubility coefficient of CO_2 in plasma, is $0.0307 \cdot mmol \cdot L^{-1} \cdot mm\ Hg^{-1}$. A close approximation of the total CO_2 can be made by adding the bicarbonate concentration and the concentration of dissolved CO_2 (i.e., ignoring the small amounts of carbonate ion and protein carbamates). This approximation is

Ctotal CO_2 = $CHCO_3^-$ + α PCO_2.

The preceding two equations are incorporated into most blood gas analyzers. There has been some controversy in the literature about the applicability of the Henderson-Hasselbalch equation in acutely ill patients. This has been carefully reviewed (15), with the conclusion that, although minor variations in pK′ certainly occur, they are not large enough to affect the clinical usefulness of a calculated total CO_2 (or bicarbonate). This calculation is usually a part of routine blood pH and gas analysis and facilitates the interpretation of blood gas data.

Base Excess

When blood pH and PCO_2 are abnormal, it is often not immediately obvious whether the abnormality is purely respiratory in nature or whether a metabolic component is present. Additional insight may be gained by using one of the many nomograms that have been constructed for this purpose. In addition, attempts have been made to define a quantity that can be readily calculated and that would reflect only the metabolic component of an acid-base imbalance.

Several quantities have been proposed, but base excess of ECF, CBE, is generally believed to be the most useful of these. Although in principle base excess can be measured, it is much more common to calculate it from the relationship (16)

CBE = $CHCO_3^-$ - 24.8 + 16.2 (pH - 7.40)

Base excess has also been defined for whole blood, which is calculated differently. Although there is agreement that this quantity is not as accurate in reflecting metabolic acid-base disturbances, its use preceded that of ECF base excess and became well established. Unfortunately, manufacturers of blood gas analyzers persist in making both types of base excess calculations available, so care must be taken not to use the base excess of whole blood. Many laboratories avoid the potential confusion by simply not reporting base excess. The rationale is that base excess, like any other calculated quantity, adds no essential information and is useful only to the extent that it facilitates interpretation of the measured quantities. The particular type of interpretive aid favored (if any) usually reflects how a person was trained to interpret acidbase data. There is no inherent reason to prefer calculation of ECF base excess to use of a nomogram or to some other approach.

Temperature Corrections

pH, PCO_2, and PO_2 are all temperature-dependent quantities, and by convention measurements are always made at 37.0°C. However, if the body temperature of the patient differs from 37°C, then the question arises whether to adjust the measured values to the temperature of the patient. There has always been some controversy about the desirability of making this adjustment. On the one hand, it is reasonable to do so because the adjusted values would correspond to the actual conditions *in vivo*. On the other hand, the reference intervals used to interpret the pH, PCO_2, and PO_2 were determined at 37°C, and the question of what values should be considered normal at other temperatures is unresolved. To minimize confusion, it is wise to report only the values measured at 37°C or to report both and the values adjusted to body temperature. Most blood gas analyzers include software to automatically calculate adjusted values if the body temperature of the patient is known, and algorithms suggested for these adjustments have been published (16).

Specimens and Specimen Handling

The collection and handling of specimens for blood pH and gas analysis have requirements unlike those for other clinical laboratory tests. These apply to the site and technique of collection, the specimen container, the anticoagulant, and the transport and storage of the specimen. All these topics and several others were recently reviewed by a subcommittee of the National Committee for Clinical Laboratory Standards (NCCLS), and a guideline document was published covering specimen collection as well as calibration and quality control issues (17). The essence of the NCCLS recommendations on specimen collection and handling are covered in this section.

Site and Technique

In the majority of requests for blood gas testing, there is a need to evaluate the degree of oxygenation of the blood in addition to the acid-base status. A specimen of arterial blood is therefore required. The technique of obtaining an arterial blood specimen is more difficult and more hazardous to the patient than venous sampling. Another NCCLS publication (18) contains detailed procedures and precautions to be observed in arterial blood sampling and is recommended.

If it is not practical to obtain arterial blood, capillary blood may be used, but with some constraints (19). The capillaries, usually of the foot or the fingertip, must be dilated by warming the skin so that the PO_2 will be close to the arterial level. Despite this, the correlation of capillary PO_2 to arterial PO_2 is not very good, in part owing to the difficulty in obtaining a capillary sample without exposing the blood to room air. The correlation is good, however, for PCO_2 and pH.

If arterial PO_2 is not a concern, the acid-base status of the blood can be evaluated by analysis of venous blood for pH and PCO_2 to avoid the discomfort and hazard of arterial puncture.

Container and Storage

Capillary specimens are collected in preheparinized glass capillary tubes. Recommended specifications for the tubes have been published (20). The blood must be well mixed in the tube to ensure homogeneity and dissolution of the anticoagulant.

Venous specimens (for pH and PCO_2 only) may be collected in either syringes or evacuated tubes, but the latter must be completely filled. Arterial specimens should be collected in a syringe. Both glass and plastic syringes are used for blood gas specimens, and two factors influence the choice: (a) the expense of maintaining a system to clean, sterilize, and reuse glass syringes and (b) the length of time that may elapse before the specimens are analyzed. Glass is superior because plastic syringes can alter the PO_2 (and, to a lesser extent, the PCO_2) of a blood specimen, presumably owing to room air dissolved in the plastic syringe barrel and plunger tip (21,22). The size of the error depends on the PO_2

and the temperature of the blood but is complicated because the buffering of PO_2 by hemoglobin changes with PO_2, as reflected by the oxygen-hemoglobin association curve. The error is increased when the specimen is cooled in ice water because oxygen solubility increases at lower temperatures.

Beginning at a PO_2 of 100 mm Hg, the oxygen level of a specimen in a plastic syringe stored in ice water may increase as much as 8 mm Hg in 30 minutes. This error is much smaller at lower PO_2 because hemoglobin is unsaturated and is also much smaller if the specimen is kept at room temperature instead of being iced. A complication of not chilling the specimen is that metabolic processes utilizing oxygen proceed at a faster rate in the blood specimen, which constitutes yet another source of error. This error is particularly significant in blood with an elevated leukocyte or platelet count.

It is recommended that if plastic syringes are used, the blood not be chilled but that the analysis be completed within 20 minutes to reduce the probability of error owing to metabolic changes. Conversely, if glass syringes are used, the specimens should be cooled in ice water so that metabolic changes will be insignificant even with moderate elevations in leukocytes or platelets, and most specimens can then be held at least an hour before analysis, if necessary.

Room air has a PO_2 of approximately 150 mm Hg and a PCO_2 of less than 1 mm Hg. Obviously, contact between the blood and room air must be prevented. No air bubbles may be present in the specimen, and the needle should be replaced with an airtight cap as soon as the specimen is drawn.

Anticoagulant

Sodium or lithium heparin is the anticoagulant to be used for blood pH and gas analysis. Both are available in dry and liquid forms.

If the liquid preparation is used, care must be taken not to use more volume than necessary because the errors caused by sample dilution can be significant. If a heparin strength of 1,000 IU/mL is used, the volume in just the syringe dead space (needle and hub) is sufficient to anticoagulate a specimen in a 3-mL syringe. More concentrated heparin solutions are not recommended.

Heparin in solid form is used in some capillaries and prepacked syringes, either as a pellet or as a coating on the container wall. Dilution errors are obviously avoided, but extra care must be taken to dissolve the heparin in the blood by vigorous mixing.

Quality Control and Proficiency Testing for Blood pH and Gas Analysis

The basic principles and issues of quality control and proficiency testing (QC/PT) that apply to all analytes in clinical chemistry apply also to blood pH and gases, but there are some additional aspects that are unique to blood gas QC/PT.

Materials and Matrix Effects

A universally accepted principle in QC/PT is that the sample used must have the same characteristics as fresh patient samples. However, this is often impractical. As a result, QC/PT materials may give results that are different from those expected on patient samples. Such differences are known as matrix effects and are of particular concern in blood gas QC/PT (23).

Commercially available controls for blood pH and gases usually are sold in the form of sealed glass ampules containing a solution in equilibrium with a gas phase above it. The solution matrix is one of three types: (a) aqueous; (b) blood based, either stabilized whole blood or a hemolysate; or (c) perfluorocarbon emulsion. In all cases, the solution is buffered to stabilize the pH. The most important differences are related to the PO_2 stability in the QC material when exposed to room air. This is an important issue because it is impossible to completely avoid contact between a QC sample for blood gas analysis and room air.

Blood gas analyzers have been carefully designed to give accurate results when whole blood is used as the sample. Matrix effects with other materials are primarily owing to differences in the solubility of oxygen and carbon dioxide in these materials, and the effect is especially noticeable for PO_2. The solubility of oxygen in whole blood is much greater than in aqueous solutions because of the presence of hemoglobin. Exposure of a whole blood sample to a small amount of room air will change the PO_2, but hemoglobin will tend to buffer this change as long as it is not fully saturated with oxygen. Aqueous solutions, having no hemoglobin, will show relatively large changes in PO_2 when exposed to room air. Perfluorocarbon emulsions are used because the solubility of oxygen in perfluorocarbons is much higher than in water, and this will minimize the error caused by contamination of the sample with room air.

In PT programs in which both perfluorocarbon emulsions and aqueous materials are used, data from perfluorocarbons show significantly better precision for PO_2 than do aqueous materials (23). Conversely, perfluorocarbons do not resemble a whole blood matrix, and this material may not be usable in analyzers containing a potassium ISE because of interaction between perfluorocarbons and the potassium ionophore.

Blood-based controls do not necessarily behave like fresh patient samples either. Often the hemoglobin present in such controls is fully saturated with oxygen at a very low PO_2, and the sensitivity to room air contamination approaches that of aqueous solutions. However, some blood-based materials, or fresh whole blood, can be used for blood gas QC using a procedure called tonometry. Tonometry is the process of equilibrating a liquid with a gas of known composition. The apparatus used to accomplish this is called a tonometer. Tonometry makes it possible to prepare QC materials with an accurately known PO_2 and PCO_2, one of the rare instances in clinical chemistry when a QC material can be used as a test of accuracy as well as precision. Tonometers are readily available and the procedure and sources of error have been well described (17,24), but the technique is underutilized because of the perception that it is too time-consuming to be performed routinely. Nevertheless, tonometry is the basis for the reference method for blood PO_2 and PCO_2.

Further details and recommendations for quality control programs in a blood gas laboratory have been published by the NCCLS (17).

Temperature and Altitude Effects

Control materials that are packaged in sealed ampules containing a gas bubble are subject to a source of preanalytical error not seen

with patient samples. As the temperature of the sealed ampule rises and falls, the solubility of oxygen and carbon dioxide in the liquid changes, and the amount of gas dissolved in the liquid phase changes. Thus, the measured PO_2 and PCO_2 will depend on the temperature of the ampule when it is opened, and the instructions for handling QC/PT materials must specify what this temperature should be. This effect does not occur in patient samples because care is taken to exclude bubbles from the specimen.

Another source of error in blood gas analysis that can be important in PT is the effect of ambient pressure. It has been observed that the PO_2 and, to a lesser extent, the PCO_2 measured in PT samples tend to be lower if the laboratory is at high altitude. The effect is too small to be important clinically. However, the bias can be a significant fraction of the allowable error in PT programs, and laboratories at high altitude should use a correction factor if available (25).

INTERPRETATION OF ACID-BASE DISORDERS

Part of "17 - Electrolytes and Acid-Base Balance"

Acid-base disorders are classified into one of four categories: respiratory acidosis, respiratory alkalosis, metabolic acidosis, and metabolic alkalosis. Each of these categories includes a number of possible causes. The evaluation of an acid-base abnormality requires the correct classification of the disorder, followed by a consideration of the differential diagnostic possibilities to obtain the specific etiology. The interpretation may be complicated by differences in acute versus chronic acid-base disorders, by the effects of renal and respiratory compensatory mechanisms, and by the presence of mixed acid-base abnormalities.

In response to an abnormal pH, the body initiates compensatory mechanisms to restore the pH. The renal compensatory mechanisms include the ability of the kidney to alter bicarbonate resorption and to increase acid excretion. The respiratory compensatory mechanism operates by changing the alveolar ventilation and thus retaining CO_2 (hypoventilation) or releasing additional CO_2 (hyperventilation). The effect of compensatory mechanisms is to create a mixed acid-base disorder in which the compensatory response is superimposed on the primary acid-base abnormality. Mixed disorders may also occur in a patient with two simultaneous but unrelated primary acid-base abnormalities.

The ability of the body to normalize the arterial pH is limited, because correction of the pH would shut off the physiologic mechanisms driving the compensatory response. For this reason, the arterial pH usually remains somewhat abnormal even in the compensated state.

Table 17.5 shows the various types of primary acid-base disorders, the principal mechanism causing the disorder, the expected laboratory abnormalities, and the effects of compensatory responses.

TABLE 17.5. Characteristics of Primary Acid-Base Disorders

Disorder	Primary Change	Compensatory Response
Metabolic acidosis	Decreased HCO_3	Hyperventilation
	Decreased pH	Decreased Pco_2
Metabolic alkalosis	Increased HCO_3	Hypoventilation
	Increased pH	Increased Pco_2
Respiratory acidosis	Increased Pco_2	Increased acid excretion and HCO_3 resorption
	Decreased pH	
Respiratory alkalosis	Decreased Pco_2	Decreased acid excretion and HCO_3 resorption
	Increased pH	

There are several general approaches to the evaluation of acid-base disturbances. One system utilizes an acid-base nomogram such as the one shown in Fig. 17.5. If the pH and the PCO_2 are measured, then a presumptive diagnosis may be made by locating the point corresponding to the two values on the nomogram to determine which of the seven diagnostic zones it falls into. For example, a patient with an arterial pH of 7.3 and a PCO_2 of 30 would be classified as having a metabolic acidosis. The advantages of this method are that it is simple, requires no calculations, and provides a reliable classification of the disorder, even with mixed acid-base abnormalities. The disadvantage of using a nomogram is that these charts are not always readily available, especially at the bedside in emergency situations. For this reason, it is essential to have a simplified system for interpreting acid-base disorders that does not rely on a nomogram or the use of complex calculations.

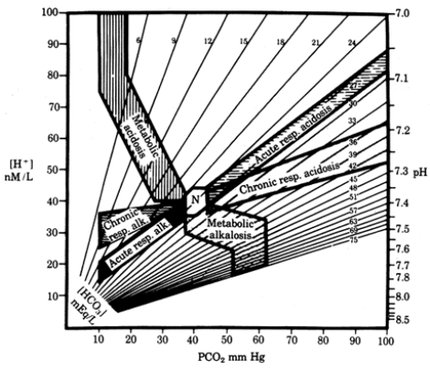


FIGURE 17.5. Diagram for interpretation of acid-base data. (From *JAMA*1973;223:269.)

Various textbooks describe sets of rules of thumb that permit a rapid interpretation of acid-base abnormalities (26,27 and 28). The major drawback to the use of these rules is that complex acid-base disorders are difficult to interpret. The clinical history must be considered whenever possible, rather than attempting a diagnosis based on laboratory data alone. Recognizing the need for a practical approach to the interpretation of acid-base data, the American Heart Association has recommended the use of a set of simplified rules that permit rapid interpretation of acid-base data (26,27). Two of these rules are shown in Table 17.6 and should be memorized and practiced. Most acid-base disorders can be interpreted by using only the first rule.

TABLE 17.6. Simplified Rules for the Interpretation of Acid-Base Disorders

Rule 1:	A change in Pco_2 of 10 mm Hg will change the pH by 0.08.
Rule 2:	A change in bicarbonate of 10 mmol/L will change the pH by 0.15.

Adapted from Comella L, Braen R, Olding M. Blood gases and acid-base disorders. In: *Clinician's pocket reference*, 4th ed. Laguna Niguel, CA: Capistrano Press, 1983:157–168.

The usefulness of rule 1 is best illustrated by example. Suppose a patient has a pH of 7.32 and a PCO_2 of 50 mm Hg. In this case, the PCO_2 is elevated by 10 mm Hg. Using rule 1, this should produce a decrease in pH of 0.08, and therefore the predicted pH is 7.32 (7.40 to 0.08). Because the measured pH and the calculated pH are the same (or close), all the change can be explained by respiratory dysfunction. The most likely diagnosis is therefore acute respiratory acidosis. Now consider a second patient with a pH of 7.35 and a PCO_2 of 52 mm Hg. The elevated PCO_2 again indicates a component of respiratory acidosis. However, rule 1 predicts a pH of approximately 7.30. The measured pH is greater (closer to normal) than predicted by 0.05, indicating that compensation has occurred. The correct diagnosis is respiratory acidosis with partial renal compensation (chronic respiratory acidosis).

Either the acid-base nomogram or these rules permit a classification of acid-base disorders into one of the categories shown in Table 17.5. The next step is to consider the differential diagnostic possibilities that comprise each category.

Respiratory Acidosis

Respiratory acidosis results from any condition that impairs CO_2 elimination by the lungs and is characterized by an increased

PCO_2 and a decreased pH. A large number of disorders may produce respiratory acidosis, as seen in Table 17.7. These disorders may be classified as neuromuscular diseases, airway obstruction, and cardiopulmonary-thoracic disorders. Neuromuscular diseases may affect the respiratory apparatus at any level, including the CNS respiratory center in the brainstem, the peripheral nerves innervating the diaphragm and chest wall muscles, or the respiratory muscles themselves. The result is hypoventilation and retention of CO_2. Several drugs produce respiratory acidosis by depressing the central respiratory center. Narcotic overdose is a classic example of this effect. Airway obstruction can occur at any level of the respiratory tract. Cardiopulmonary-thoracic disorders also include a variety of other conditions that impair gas exchange.

TABLE 17.7. Selected Differential Diagnosis of Acid-Base Disorders

Respiratory acidosis
Neuromuscular
Brainstem/spinal cord injury
Narcotic overdose
Guillain-Barré syndrome
Myasthenia gravis
Poliomyelitis
Botulism
Diaphragmatic paralysis
Myopathy of respiratory muscles
Cardiopulmonary
Airway obstruction (foreign body, tumor bronchospasm, COPD)
Cardiac arrest
Severe pneumonia or pulmonary edema
Pulmonary embolus
Mechanical underventilation
Plural effusion
Restrictive disease of thorax
Metabolic
Myxedema
Metabolic acidosis
High anion gap
Lactic acidosis
Ketoacidosis (diabetic or alcoholic)
Aspirin overdose (late)
Methanol/ethylene glycol intoxication
Paraldehyde
Renal failure
Rhabdomyolysis
Normal anion gap
GI loss of bicarbonate (diarrhea, pancreatic or biliary drainage)
Renal loss of bicarbonate (RTA, tubulointestitial diseases, hypoaldosteronism)
Acid gain (ammonium chloride, amino acid administration)
Respiratory alkalosis
Central nervous system
Anxiety
Aspirin overdose (early)
Fever
CNS infection, tumor, stroke
Hypoxia
Pulmonary
Pulmonary emboli
Pneumonia
Interstitial lung disease
ARDS
Pulmonary edema (mild)
High altitude
Mechanical overventilation
Metabolic/other
Hepatic failure
Pregnancy
Sepsis
Metabolic alkalosis
Chloride responsive
Vomiting, nasogastric suction
Diuretics
Villous adenoma
Cystic fibrosis
Posthypercapnia
Congenital chloridorrhea
Chloride resistant
Hyperaldosteronism
Cushing's syndrome
Bartter's syndrome
Excessive licorice ingestion
Severe potassium depletion
Miscellaneous
Alkalinizing agents
Milk alkali syndrome
Hypoparathyroidism
Drugs: carbenicillin, penicillin (large doses)

Adapted from Alder S, Lam M, Connors A. Acid-base and electrolyte disorders. In: A pocket manual of differential diagnosis. Boston, 1982: Little, Brown, 1–7.
CNS, central nervous system; ARDS, adult respiratory distress syndrome; COPD, chronic obstructive pulmonary disease; GI, gastrointestinal; RTA, renal tubular acidosis.

In acute respiratory acidosis, the bicarbonate level and total CO_2 are normal. The increase in PCO_2 will stimulate the central respiratory center to increase respiration. Renal compensatory mechanisms require more time to show an effective response. The effect of renal compensation is to convert CO_2 and water to hydrogen ion, which is excreted, and bicarbonate, which is reabsorbed into the plasma. When this occurs, the acid-base disorder is classified as chronic respiratory acidosis, characterized by an increased PCO_2, a decreased pH, and an increase in the plasma bicarbonate and total CO_2.

There are also effects on other electrolytes, notably the plasma potassium and chloride. Acidemia is associated with mild hyperkalemia followed by increased renal potassium excretion, and ultimately potassium depletion. The increase in renal bicarbonate resorption is balanced by increased excretion of chloride to maintain electroneutrality. Thus, hypochloremia may accompany respiratory acidosis.

Respiratory Alkalosis

Respiratory alkalosis results from hyperventilation and is characterized by a decreased PCO_2 and an increased pH. The differential diagnosis includes primary disorders of the central respiratory center, pulmonary diseases associated with hypoxemia, and a variety of miscellaneous causes of hyperventilation (Table 17.7). Symptoms of respiratory alkalosis include lightheadedness, paresthesias, tetany, and, in severe cases, syncope. The two most common causes are acute anxiety and hyperventilation in response to hypoxemia. A less obvious cause of respiratory alkalosis is the rapid correction of a metabolic acidosis, for example, by bicarbonate infusion. Although the blood pH may return to normal, the CNS system requires some time to equilibrate with the periphery. The temporary persistence of CNS acidosis continues to stimulate the respiratory control center with the resultant hyperventilation and peripheral respiratory alkalosis.

Acute respiratory alkalosis is characterized by a low PCO_2, a high pH, and a normal bicarbonate and total CO_2. Alkalemia depresses the respiratory control center, which tends to limit the

hyperventilation in some cases, but the primary compensation is renal. The kidneys respond by decreasing acid excretion and bicarbonate resorption, which causes the plasma bicarbonate to fall. In chronic respiratory alkalosis, therefore, the PCO_2 is low, the pH is high (but not as high as in the acute condition), and the plasma bicarbonate and total CO_2 are decreased.

Other electrolytes may show abnormalities as well. Alkalemia is associated with a mild hypokalemia. Increased renal bicarbonate excretion is accompanied by chloride retention, resulting in hyperchloremia.

Metabolic Alkalosis

Metabolic alkalosis is characterized by increased pH, bicarbonate, and total CO_2. The causes are often divided into two categories, designated chloride-responsive and chloride-resistant types (Table 17.7). The laboratory distinction is made by measuring the urine chloride; a level below 20 mmol/L is classified chloride responsive and above 20 mmol/L is chloride resistant.

One of the more common subcategories of chloride-responsive metabolic alkalosis is called contraction alkalosis. This condition results from a depletion of salt (NaCl) and water without a proportionate loss of bicarbonate (7). Contraction of the ECF volume results in an increase in the plasma bicarbonate concentration and thus metabolic alkalosis. Ordinarily the kidney would excrete the excess bicarbonate, but in the setting of salt and water depletion, the kidney conserves sodium to maintain the circulating volume and thus cannot excrete the bicarbonate (28).

Several causes of chloride-responsive metabolic alkalosis are associated with the loss of hydrogen ions from the gastrointestinal tract. Gastric secretions contain hydrochloride. Consequently, prolonged vomiting or nasogastric suctioning can produce metabolic alkalosis. Volume and salt depletion also contribute to the metabolic alkalosis as described above.

Villous adenomas of the intestine may secrete large amounts of sodium, chloride, and water (contraction alkalosis). In cystic

fibrosis, large amounts of sodium chloride may be lost in the sweat.

The major causes of chloride-resistant metabolic alkalosis are hyperaldosteronism, Cushing's syndrome, Bartter's syndrome, and severe potassium depletion. Hyperaldosteronism promotes renal hydrogen ion and potassium excretion. Cushing's syndrome may cause metabolic alkalosis because cortisol has weak mineralocorticoid activity. Bartter's syndrome, a rare disorder characterized by juxtaglomerular cell hyperplasia, produces a hypokalemic metabolic alkalosis as a result of increased renin production and resultant hyperaldosteronism. Licorice contains a steroid with weak mineralocorticoid activity (28).

Metabolic Acidosis

Metabolic acidosis is caused by an increased production of organic acids, decreased renal hydrogen ion excretion, or loss of bicarbonate. Any of these mechanisms causes a decrease in the pH, the bicarbonate concentration, and the total CO_2. The respiratory compensatory mechanism will partially correct the acidosis by increasing the rate of respiration, thereby lowering the PCO_2. The renal compensatory mechanism will increase hydrogen ion excretion and bicarbonate resorption. Metabolic acidosis is the most complicated of the various acid-base abnormalities; an excellent discussion is presented by Black (10).

Metabolic acidosis is usually classified based on whether the anion gap is increased. The anion gap (AG) is a calculated quantity that reflects the difference between the measured cations and the measured anions, as discussed earlier in this chapter. The AG is usually defined as

$AG = (C_{Na^+}) - (C_{Cl^-}) - (C_{HCO_3^-})$

The reference interval for the AG is 8 to 14 mmol/L. Although many conditions can alter the AG (5), its only clinical use is in the differential diagnosis of metabolic acidosis (Table 17.7).

There are five conditions that may cause a high-AG metabolic acidosis: renal failure, lactic acidosis, ketoacidosis, rhabdomyolysis, and some drugs and toxins.

Renal Failure

Renal failure can cause either a normal AG or a high-AG metabolic acidosis. Renal disease impairs hydrogen ion excretion due to a decrease in ammonia synthesis. Impaired hydrogen ion excretion necessitates a decrease in renal HCO_3^- resorption and thus the plasma HCO_3^- falls. Chloride replaces bicarbonate in the blood, resulting in a normal AG (hyperchloremic) acidosis. Conversely, severe renal failure is associated with renal retention of phosphate and sulfates (both unmeasured anions) and thus a high-AG acidosis.

Lactic Acidosis

Lactic acidosis is a common cause of high-AG acidosis. Lactic acid is a product of anaerobic metabolism. Major causes of lactic acidosis include tissue hypoxia (e.g., severe exercise, seizures, cardiac failure, hypoxemia); drugs and toxins (e.g., phenformin, catecholamines, salicylate, isoniazid, cyanide), congenital forms of lactic acidosis (such as defects in gluconeogenic enzymes), and a variety of severe illnesses, including sepsis, liver failure, and neoplasms (10). Lactate levels in excess of 5 mmol/L (reference interval 0.5 to 2.2 mmol/L) are diagnostic of lactic acidosis. The increase in the AG results from the unmeasured lactate anion.

Ketoacidosis

Ketoacidosis may occur in uncontrolled diabetes mellitus, starvation, or alcohol ingestion. Ketoacids (acetoacetic acid and β-hydroxybutyric acid) are overproduced by the liver, resulting in an increase in unmeasured anions and a high-AG acidosis.

Rhabdomyolysis

Massive destruction of muscle tissue releases organic acids from damaged myocytes, leading to a high-AG metabolic acidosis.

Drugs and Toxins

The four most common substances in this group are aspirin, methanol, ethylene glycol, and paraldehyde.

1. Salicylate overdose initially causes a respiratory alkalosis because the drug has a stimulatory effect on the central respiratory center. Eventually a high-AG acidosis occurs owing to the salicylate itself and to the effects of salicylate on peripheral metabolism, resulting in the production of a variety of organic acids, including lactate and ketoacids.
2. Methanol is metabolized to formaldehyde and formic acid. Accumulation of formate causes a high-AG acidosis. Methanol intoxication may also cause blindness.
3. Ethylene glycol, a component of antifreeze, is metabolized to glycolic, oxalic, and other organic acids. The toxic effects may include respiratory failure and acute renal failure, the latter owing to precipitation of calcium oxalate and hippurate crystals in the kidney. Urinalysis may therefore reveal oxalate and hippuric acid crystals.
4. Paraldehyde was once used to treat seizures. The drug is metabolized to acetic acid, although the metabolic acidosis may actually be caused by the presence of ketoacids in patients treated with this drug. The only case we recently encountered was observed in a foreign visitor to the United States.

The differential diagnosis of a high-AG acidosis may be simplified by calculation of the osmolal gap. The osmolal gap is the difference between the measured serum or plasma osmolality and the calculated osmolality as defined earlier in this chapter.

The osmolal gap is normally less than 10 mOsm/kg. If the measured osmolality exceeds the calculated osmolality by more than 10, the presence of a hidden osmotically active substance should be suspected (10). A high osmolal gap with a high-AG metabolic acidosis is suggestive of methanol or ethylene glycol poisoning.

Normal-AG Metabolic Acidosis (Hyperchloremic Acidosis)

A normal-AG acidosis can occur in several situations, including bicarbonate losses in the gastrointestinal tract, infusion of acids, and after recovery from ketoacidosis.

Pancreatic and biliary secretions contain bicarbonate that may be lost in patients with diarrhea, pancreatic or biliary drainage, or ureterosigmoidostomies.

A normal-AG acidosis can also result from a net gain in acid. This is usually owing to the use of acidifying agents such as ammonium chloride or to the administration of hyperalimentation fluids containing amino acids (10).

RTA may occur in a number of conditions associated with either decreased renal hydrogen ion excretion or renal bicarbonate wasting. Type 1 RTA (distal or classic RTA) results from an inability to secrete hydrogen ions in the distal tubule. The urine pH is high (above 5.3) despite acidemia. The causes include a primary idiopathic form(s), various autoimmune disorders, nephrocalcinosis, some drugs, several congenital conditions, pyelonephritis and urinary obstruction, and cirrhosis (10). Type 2 RTA (proximal RTA) results from renal bicarbonate wasting as evidenced by a rapid excretion of a bicarbonate load and fractional excretion of bicarbonate greater than 15%. Like type 1, RTA the differential diagnosis includes a long list of conditions (10). Type 4 RTA (there is no type 3 RTA) results from a deficiency of aldosterone or from aldosterone resistance. The causes include adrenal insufficiency, congenital adrenal hyperplasia, hyporeninemic hypoaldosteronism, and administration of angiotensin-converting enzyme inhibitors.

REFERENCES

1. Rose B. *Clinical physiology of acid-base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989.

2. Rose B. The total body water and the plasma sodium concentration. In: *Clinical physiology of acid-base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989:211–224.

3. Black R. Disorders of serum sodium and serum potassium. In: Irwin R, Alpert J, Fink M, eds. *Intensive care medicine*. Boston: Little, Brown, 1985:610–624.

4. Kumar A, Chapoteau E, Czech BP, et al. Chromogenic ionophore-based methods for spectrophotometric assay of sodium and potassium in serum and plasma. *Clin Chem* 1988;34:1709–1712.

5. Tietz N, Siggaard-Andersen O, Pruden E. Acid-base balance and acid-base disorders. In: Teitz N, ed. *Textbook of clinical chemistry*. Philadelphia: Saunders, 1986:1221–1254.

6. Rose B. Hypo-osmolal states—hyponatremia. In: *Clinical physiology of acid-base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989:601–638.

7. Delmez J. Fluid and electrolyte disturbances. In: Cambell W, Frisse M, eds. *Manual of medical therapeutics*, 24th ed. Boston: Little, Brown, 1983:23–43.

8. Tannen R. The patient with hypokalemia or hyperkalemia. In: *Manual of nephrology*, 2nd ed. Boston: Little, Brown, 1985:31–48.

9. Rose B. Hyperkalemia. In: *Clinical physiology of acid-base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989:715–756.

10. Black R. Metabolic acid-base disturbances. In: Irwin R, Albert J, Fink M, eds. *Intensive care medicine*. Boston: Little, Brown, 1985:596–609.

11. Burnett RW. On the use of phosphate buffers of physiological ionic strength for standardizing blood pH measurements. In: Siggaard-Andersen O, ed. Copenhagen: Private Press, 1981.

12. Stow RW, Baer RF, Randall BF. Rapid measurement of the tension of carbon dioxide in the blood. *Arch Phys Med Rehabil* 1957;38:646–650.

13. Severinghaus JW, Bradley AF. Electrodes for blood pO_2 and pCO_2 determination. *J Appl Physiol* 1958;13:515–520.

14. Dijkhuszen P, Fongers TME, Rispens P, et al. A new reference method for the determination of total CO_2 in biological fluids. *Clin Chim Acta* 1978;86:339–347.

15. Maas AHJ, Rispens P, Siggaard-Andersen O, et al. On the reliability of the Henderson-Hasselbalch equation in routine clinical acid-base chemistry. *Ann Clin Biochem* 1984;21:26–39.

16. Ehrmeyer S, et al. *Definitions of quantities and conventions related to blood pH and gas analysis*, 2nd ed. Document C12-T2. Villanova, PA: National Committee for Clinical Laboratory Standards, 1991.

17. Ehrmeyer S, et al. *Blood gas pre-analytical considerations: specimen collection, calibration and controls*. Document C27-A. Villanova, PA: National Committee for Clinical Laboratory Standards, 1993.

18. Wiseman JD, et al. *Percutaneous collection of arterial blood for laboratory analysis*, 2nd ed. Document H11-A2. Villanova, PA: National Committee for Clinical Laboratory Standards, 1992.

19. Bruck E, et al. *Procedures for the collection of diagnostic blood specimens by skin puncture*, 3rd ed. Document H4-A3. Villanova, PA: National Committee for Clinical Laboratory Standards, 1991.

20. Wiseman JD, et al. *Devices for collection of skin puncture blood specimens*, 2nd ed. Document H14-A2. Villanova, PA: National Committee for Clinical Laboratory Standards, 1990.

21. Mahoney JJ, et al. Changes in oxygen measurements when whole blood is stored in iced plastic or glass syringes. *Clin Chem* 1991;37:1244–1248.

22. Müller-Plathe O, Heyduck S. Stability of blood gases, electrolytes, and haemoglobin in heparinized whole blood samples: influence of the type of syringe. *Eur J Clin Chem Clin Biochem* 1992;30:349–355.

23. Burnett RW. Matrix effects in blood gas measurement. *Arch Pathol Lab Med* 1993;117:365–368.

24. Burnett RW, Covington AK, Maas AHJ, et al. IFCC Method (1988) for tonometry of blood: reference materials for pCO_2 and pO_2. *Clin Chim Acta* 1989;185:S17–S24.

25. Burnett RW, Itano M. An interlaboratory study of blood gas analysis: Dependence of pO_2 and pC_2 results on atmospheric pressure. *Clin Chem*1989;35:1779–1781.

26. Sladen A. Acid-base balance. In: McIntyre K, Lewis J, eds. *Textbook of advanced cardiac life support*. Dallas, TX: American Heart Association, 1983:135–142.

27. Comella L, Braen R, Olding M. Blood gases and acid-base disorders. In: *Clinician's pocket reference*, 4th ed. Laguna Niguel, CA: Capistrano Press, 1983:157–168.

28. Rose B. Metabolic alkalosis. In: *Clinical physiology of acid-base and electrolyte disorders*, 3rd ed. New York: McGraw-Hill, 1989:501–555.

18

Respiration and Measurement of Oxygen and Hemoglobin

Kent Lewandrowski

Michael Pins

Robert W. Burnett

- NORMAL PHYSIOLOGY
- MEASUREMENT OF OXYGEN AND HEMOGLOBIN SPECIES
- PATHOPHYSIOLOGY AND INTERPRETATION

NORMAL PHYSIOLOGY

Part of "18 - Respiration and Measurement of Oxygen and Hemoglobin"

Respiration is the exchange of oxygen and carbon dioxide between the environment and cells. In the lungs, gas exchange is accomplished by diffusion of oxygen from alveoli into the blood and of carbon dioxide from the blood into the alveoli. At the cellular level, respiration requires the transport of oxygen into the cells and the removal of carbon dioxide. A brief review of normal physiology is helpful to understand pathophysiology.

Components of Respiration

Respiration depends on the proper functioning of the following components:

1. Ventilation,
2. Gas exchange between the lungs and the blood,
3. Oxygen binding to hemoglobin,
4. Adequate cardiac output.

The exchange of oxygen and carbon dioxide in the lungs requires alveolar ventilation and the matching of ventilation to pulmonary blood flow. Gas exchange, therefore, depends on the rate and the depth of ventilation.

The airways are divided into two zones: a conducting and a respiratory zone. The conducting zone consists of the trachea, right and left mainstream bronchi, lobar bronchi, segmental bronchi, and terminal bronchioles. The respiratory zone comprises the structures in which gas exchange occurs. This includes the respiratory bronchioles, alveolar ducts, and alveoli. Passive diffusion of gases takes place primarily at the level of the alveolar-pulmonary capillary membrane. The pulmonary capillary membrane must be thin (0.3 μm) to permit efficient diffusion of gases. This membrane consists of the alveolar epithelium, the underlying basement membrane, and the pulmonary capillary endothelium.

A constant flow of blood is essential to allow proper gas exchange. Systemic (venous), deoxygenated blood carrying CO_2 is transported to the right side of the heart and from there to the pulmonary arteries. Gas exchange then occurs in the pulmonary capillary bed, and the oxygenated blood is transported via the pulmonary veins to the left side of the heart and then to the systemic arterial circulation. Ventilation can increase 20- to 30-fold with increased oxygen demand such as occurs with strenuous exercise.

The exchange of gases across the pulmonary capillary membrane is governed by gradients in the partial pressures of individual gases. Dry room air is composed of 78% nitrogen, 21% oxygen, 0.03% carbon dioxide, and 0.1% inert gases. Inspired air is rapidly warmed to 37°C in the upper respiratory passages and becomes saturated with water (PH_2O = 47 mm Hg at 37°C). Normal values for the partial pressures of gases in inspired air and in arterial and venous blood are shown in Table 18.1.

TABLE 18.1. Partial Pressures of Gases in the Respiratory Tract and Blood[a]

Compartment	Po_2	Pco_2	Ph_2o
	mm Hg	*mm Hg*	*mm Hg*
Inspired air	160	0.25	Variable
Arterial blood	100	40	—
Venous blood	40	46	—
Expired air	115	30	47

[a]For atmospheric pressure = 760 mm Hg.

The respiratory system is designed such that (a) forces exerted by the respiratory muscles counteract the elastic recoil and airway resistance of the lungs, (b) a minimal pulmonary capillary membrane thickness is maintained, (c) right and left heart output are balanced, and (d) ventilation ([V with dot above]) and perfusion ([Q with dot above]) within the lung are matched. Compliance is a measure of the stretchability of the lung and is inversely proportional to elasticity. Most of the energy required for expiration is derived from the elastic recoil properties of the lung. Conversely, inspiration requires energy expenditure from the respiratory muscles that is proportional to the pulmonary compliance. Low compliance results in difficulty with inspiration (restrictive disease), whereas a high compliance decreases the work of inspiration but impairs expiration. A similar concept applies at the level of the alveoli. Laplace's law describes the relationship between the alveolar volume (radius, *R*), surface tension of the alveolar fluid (*T*), and the pressure (*P*) required to maintain an inflated alveolus ($P = 2T/R$). In theory, the smaller the alveolar radius, the greater the pressure required to maintain the alveolar volume, assuming that the surface tension is the same. Pressure differences between alveoli might result in collapse of smaller alveoli were it not for the type II pneumocytes that secrete pulmonary surfactant. Surfactant is a mixture of phospholipids (chiefly lecithin) that lowers alveolar surface tension and thus prevents the collapse of small airspaces. A deficiency of surfactant results in stiff, noncompliant lungs. This condition underlies the pathogenesis of respiratory distress syndrome (RDS) in premature infants. The level of surfactant produced by the fetal lung increases during fetal development,

with a large increase after 35 weeks of gestation. Infants born between 32 and 36 weeks exhibit a 15% to 20% risk of developing RDS, compared with 5% born after 37 weeks.

Ventilation, Perfusion, and the Ventilation-Perfusion Ratio

Ventilation

Ventilation ([V with dot above]) is the exchange of gases between ambient air and the lungs. At the end of expiration, a certain amount of gas remains in the conducting airways and alveoli. Thus, with inspiration, fresh air mixes with retained gases in the lung. The inspired volume, termed the tidal volume, is normally 0.5 L. Therefore, an average respiratory rate of 12 to 15 breaths per minute accounts for a minute ventilation of 6 to 7.5 L/min. Ventilation to nonperfused areas is termed dead-space ventilation and includes anatomic dead space (conducting system) and alveolar dead space (nonperfused alveoli). Essentially no alveolar dead space exists in the normal physiologic state. The distribution of ventilation in the lungs depends on two variables: airway resistance and the ability of alveoli to accommodate pressure differentials (Laplace's law). Resistance to airflow is equal throughout the pulmonary tree in the normal lung, but variances in small airway resistance are a common cause of uneven ventilation in disease states. However, the ventilation distribution is not uniform even in healthy persons because of a gradient in pleural pressure between the uppermost and lowermost regions of the lung. As a consequence, alveoli are larger at the apex than at the base of the lungs.

Perfusion

Perfusion ([Q with dot above]) refers to the flow of blood through the lungs and specifically to the flow through the pulmonary capillaries that are in close proximity to alveoli. The distribution of perfusion depends on several variables: gravity, cardiac output, and pulmonary vascular autoregulation. The pulmonary vasculature normally operates under low pressure. Therefore, gravity has a significant impact on the distribution of perfusion. Perfusion is greater at the base of the lung (dependent area) than at the apex. As cardiac output increases, pulmonary vascular pressure increases, and nondependent areas are better perfused.

Ventilation-Perfusion Matching ([V with dot above]/[Q with dot above] Ratio)

Although both ventilation and perfusion are unevenly distributed in the normal lung, the critical feature is that ventilated areas are perfused and that perfused areas are ventilated. Most of the total ventilation, and an even greater proportion of the pulmonary blood flow, is directed to dependent areas of the lungs. The result is a slightly higher [V with dot above]/[Q with dot above] ratio in nondependent areas (the lung apices in a standing person) than in dependent areas (the lung bases). Because alveolar ventilation is approximately 4 L/min and cardiac output averages 5 L/min, the overall [V with dot above]/[Q with dot above] ratio is normally approximately 0.8. Pathologic mismatches of [V with dot above] and [Q with dot above] ([V with dot above]/[Q with dot above] mismatch) are the most common cause of hypoxemia in disease states.

Alveolar-Arterial PO_2 Difference (A-a O_2 Gradient)

The A-a O_2 gradient is the difference between the partial pressure of oxygen in alveolar air (PAO_2) and the partial pressure of oxygen in the arterial blood (PaO_2). The A-a O_2 gradient reflects the efficiency of gas exchange between the alveoli and pulmonary vasculature and is useful in evaluating patients with hypoxemia. The PaO_2 is measured, whereas the P_AO_2 is calculated from the alveolar gas equation:

$$PAO_2 = (Ptotal - PH_2O)\ FIO_2 - PaCO_2/RQ$$

where Ptotal is the barometric pressure (760 mm Hg at sea level), PH_2O is the water vapor pressure (47 mm Hg at 37°C), FIO_2 is the mole fraction of inspired oxygen (0.21 in room air), $PaCO_2$ is the arterial PCO_2 (normally 40 mm Hg), and RQ is the respiratory quotient (usually 0.8). Normally the A-a O_2 gradient is less than 15 mm Hg, rising slightly with age.

To determine the A-a O_2 gradient, the patient is normally placed on room air. Under these conditions, the inspired O_2 is 150 mm Hg, R = 0.08, and the $PaCO_2$ is measured (normally 40 mm Hg). The PAO_2 is then calculated as follows:

$$PAO_2 = 150 - (40/0.8) = 100 \text{ mm Hg}$$

If the PAO_2 is normally 100 mm Hg and the PaO_2 is, for example, 70 mm Hg, then the A-a gradient would be 30 mm Hg. A simplified equation for calculating the PAO_2 for use in determining the A-a gradient is: $PAO_2 = 7 \times \%$ inspired $O_2 - 1.25 \times PaCO_2$. For a patient breathing room air, the equation becomes $PAO_2 = 150 - 1.25 \times PaCO_2$.

The A-a gradient is useful to distinguish hypoventilation (normal A-a gradient) from other causes of hypoxemia in which the A-a gradient is increased.

Regulation of Respiration

Respiration is regulated by the interaction of the medullary respiratory center in the brain, peripheral chemoreceptors, and mechanoreceptors and by the effects of local mediators. Involuntary changes in the rate or depth of respiration are directed by the medullary respiratory center, which is influenced by peripheral and central chemoreceptors. Central chemoreceptors, located beneath the ventral surface of the medulla oblongata, respond to changes in the pH of the cerebrospinal fluid. Peripheral chemoreceptors, which include the carotid and aortic bodies, respond to both pH and PO_2. A decrease in pH stimulates both the central and peripheral chemoreceptors, whereas a decrease in PO_2 stimulates only the peripheral chemoreceptors. Stimulation

of these receptors results in an increase in ventilation with a subsequent increased gas exchange. To the extent that the PCO_2 affects the arterial pH, changes in the PCO_2 will also affect the rate and depth of respiration.

Local decreases in oxygen tension cause constriction of the pulmonary vasculature. The result is shunting of blood away from unventilated areas of the lung to ventilated areas. This phenomenon occurs in the normal physiologic state (such as in dependent areas of the lung where the [V with dot above]/[Q with dot above] ratio is relatively low) as well as in a number of disease states. The opposite effect (vasodilation) occurs in the systemic circulation, thus facilitating delivery of oxygen to peripheral tissues.

Mechanoreceptors (J receptors) in the lung respond to the stretching forces exerted on the lung during inspiration and reflexively excite expiration.

An additional sensor of hypoxia deserves brief mention. Juxtaproximal tubular cells in the kidney sense hypoxia and respond by stimulating erythropoiesis via erythropoietin. The result is an increased red blood cell mass and an increased O_2 carrying capacity of the blood.

Hemoglobin and the Oxygen Dissociation Curve

In addition to adequate ventilation and gas exchange with blood, respiration requires transportation of oxygen to cells. The solubility of oxygen in plasma, and therefore the oxygen content of plasma, is relatively low. Hemoglobin in erythrocytes provides the means, through binding of oxygen molecules, of transporting oxygen to cells. Hemoglobin is a protein made up of four subunits. One molecule of hemoglobin can therefore bind four oxygen molecules.

The binding of oxygen to hemoglobin alters the quaternary structure of the tetramer such that subsequent binding of additional O_2 molecules is facilitated. This effect results in the sigmoidal shape of the oxygen–hemoglobin equilibrium curve (also called the oxygen dissociation curve), which relates the PO_2 to the oxygen saturation of hemoglobin (SO_2), as shown in Fig. 18.1.

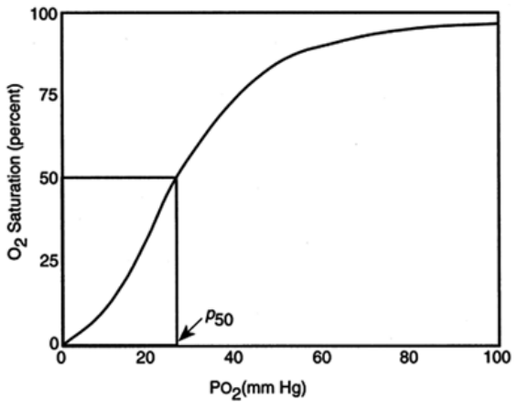


FIGURE 18.1. Oxygen–hemoglobin equilibrium curve.

At the PO_2 normally present in arterial blood, hemoglobin is more than 95% saturated with oxygen. At a PO_2 greater than 100 to 120 mm Hg, hemoglobin is virtually saturated. At this point, increases in inspired oxygen will increase only the dissolved oxygen content of blood and not the amount bound to hemoglobin. Thus, once hemoglobin is saturated, a higher FIO_2 will not appreciably increase the oxygen content of the blood (although the PO_2 may be markedly elevated).

The position of the oxygen dissociation curve reflects the affinity of hemoglobin for oxygen. Hemoglobin–oxygen affinity is influenced by a number of factors. Affinity is reduced by increases in temperature, decreases in pH, increases in PCO_2 (at constant pH), and increases in the erythrocyte 2,3-diphosphoglycerate (DPG) concentration.

An increase in hemoglobin–oxygen affinity shifts the oxygen dissociation curve to the left, whereas a decrease in affinity shifts the curve to the right. Thus, a right shift in the curve is observed with elevated DPG concentration, acidemia, and hypercapnia. The effect of these variables in normal physiology is important to the delivery of oxygen to the tissues. In the lung, where the PO_2 is high, hemoglobin takes up O_2 and releases CO_2, whereas in the tissues, PO_2 is low and the reverse occurs. The comparatively high pH and PCO_2 found in the tissues further enhances oxygen release from hemoglobin. Conversely, pathologic changes in pH, PCO_2, or temperature may impair oxygen delivery. DPG is an intermediate in normal glycolysis in red blood cells. The effect of DPG is to lower the hemoglobin–oxygen affinity and thus promote O_2 delivery to tissues. DPG concentrations increase in red blood cells in response to states of anemia or alkalosis, which tends to compensate for the primary disorder. DPG concentrations tend to decrease with length of storage of blood.

MEASUREMENT OF OXYGEN AND HEMOGLOBIN SPECIES

Part of "18 - Respiration and Measurement of Oxygen and Hemoglobin"

The Oxygen Electrode

The routine measurement of PO_2 in blood became practical with the development of the oxygen electrode by Clark in 1956. In contrast to ion-selective electrode methods used for hydrogen ion, sodium, potassium, and other ions, which are based on potentiometry, the oxygen electrode method is amperometric; that is, the PO_2 is related to the amount of current flowing in an electrochemical cell. In this cell, oxygen is reduced at a platinum wire cathode. The other half-cell is usually a silver–silver chloride electrode. When a voltage of approximately 0.6 V is applied to the cathode, oxygen is reduced according to the following reaction:

$O_2 + 2H^+ + 4e^- \rightarrow 2OH^-$

The current depends on the number of oxygen molecules reaching the cathode per unit time, which in turn depends on the concentration of dissolved oxygen in the sample, which is proportional to the PO_2. The area of the cathode is kept very small so that only a tiny fraction of the total dissolved oxygen is reduced, and the current then assumes a constant value proportional to the PO_2.

A diagram of the oxygen electrode is shown in Fig. 18.2. The platinum cathode is in contact with a very thin layer of buffer, which is separated from the blood sample by a gas-permeable membrane. The membrane serves to keep other reducible substances in blood from reaching the cathode, which would add to the cell current and could foul the electrode surface. The membrane itself adsorbs small amounts of protein from the blood samples and needs to be replaced periodically in most electrode designs.

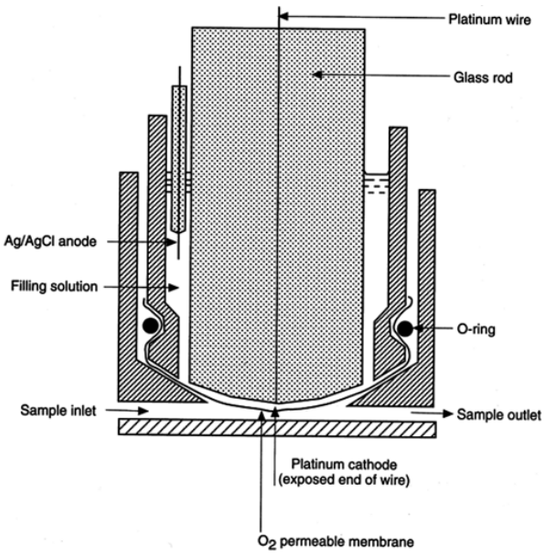


FIGURE 18.2. Schematic illustration of a PO_2 electrode. (Adapted from Siggaard-Andersen O. *The acid-base status of the blood*, 4th ed. Baltimore: Williams & Wilkins, 1974.)

Calibration of the oxygen electrode is usually done with two gas mixtures whose composition is known with very high accuracy. One contains no oxygen, which allows a zero point to be established, and the other contains a known oxygen percentage, usually either 12% or 20% (±0.03%) by volume. Further details about the calibration of blood gas analyzers and the preanalytical variables that affect PO_2 measurements are discussed in Chapter 17.

Spectrophotometry of Hemoglobin Species

The measurement of total hemoglobin is a routine part of a hematology profile. The measurement of oxyhemoglobin (O_2Hb), deoxyhemoglobin (Hb), carboxyhemoglobin (COHb), and methemoglobin (metHb) by spectrophotometry may be performed as an adjunct to blood gas measurements and is therefore discussed here.

Each of the hemoglobin species has a unique absorbance spectrum in the visible region, as illustrated in Fig. 18.3. After the spectrum of each species is established with solutions of known concentration, the analysis of a mixture is carried out by simply measuring the absorbance of a hemolysate at multiple wavelengths, one wavelength for each of the species to be determined. Beer's law applies to the absorbance at each wavelength.

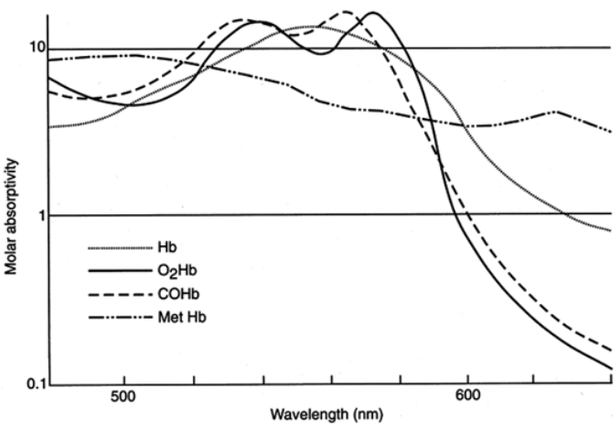


FIGURE 18.3. Absorbance spectrum of hemoglobin species.

$$A_{\lambda 1} = E_{\lambda}(1)\,C(1) + E_{\lambda 1}(2)\,C(2) + \ldots + E_{\lambda 1}(n)\;C(n)$$

$$A_{\lambda 2} = E_{\lambda 2}(1)\;C(1) + E_{\lambda 2}(2)\;C(2) + \ldots + E_{\lambda 2}(n)\;C(n)$$

$$A_{\lambda n} = E_{\lambda n}(1)\;C(1) + E_{\lambda n}(2)\;C(2) + \ldots + E_{\lambda n}(n)\;C(n)$$

where $E_{\lambda 1}$ (1) is the absorptivity at wavelength λ1 and C(1) is the concentration of the first component, and so forth.

To determine the concentrations of O_2Hb, Hb, COHb, and metHb in blood, the absorbance is measured at four wavelengths, giving four equations of the form above, which can be solved simultaneously to yield the concentrations. The absorptivities are constants that are permanently stored in the instrument. The absorbance scale may be calibrated with a total hemoglobin standard, and a microprocessor in the instrument then performs the calculations necessary to give concentrations of each hemoglobin species as well as calculations for the several derived quantities commonly reported. It should be recognized that if all four of the hemoglobin species mentioned are to be measured, a minimum of four different wavelengths must be used. Some older laboratory oximeters and oximeters used in point-of-care testing use two or three wavelengths and are intended

to measure only O_2Hb. With such devices, it is important to understand what errors will be seen if significant amounts of COHb or metHb are present (1,2).

Specimens and Specimen Handling

The measurement of O_2Hb and other hemoglobin species may be requested in conjunction with arterial blood gas measurements. In this case, the same specimen is used for both, namely, heparinized arterial blood.

If one is only interested in the measurement of COHb or metHb, then venous blood is preferred to avoid the discomfort and difficulty of an arterial puncture. In addition, the issues of the type of container and whether to store the specimen on ice are not relevant. Both COHb and metHb are relatively stable, so a specimen anticoagulated with heparin or ethylenediaminetetraacetic acid and transported at ambient temperature is satisfactory.

The most important aspect of specimen handling is to ensure that the blood is thoroughly mixed. Incomplete mixing of whole blood cannot be detected by eye but can produce significant errors in either direction in the total hemoglobin measurement.

Derived Quantities

There has long been confusion about definitions and conventions used in blood gas analysis. This has been addressed by two different organizations, the National Committee for Clinical Laboratory Standards and the International Federation of Clinical Chemists, and their documents are recommended for a more complete description of the various derived quantities that are in use (3,4).

O_2Hb Fraction and Oxygen Saturation

Results are usually expressed as a fraction (or percentage) of the total hemoglobin concentration, which is obtained by simply adding the concentrations of each of the components. For example, the O_2Hb fraction (in a percentage) would be calculated as

$$F_{O_2Hb}\,(\%) = \frac{100\,C_{O_2Hb}}{C_{O_2Hb} + C_{Hb} + C_{COHb} + C_{metHb}}$$

Fractions of COHb and metHb are calculated similarly.

Care must be taken not to confuse O_2Hb fraction with a related quantity, oxygen saturation (SO_2). The latter quantity, also called oxygen saturation of available hemoglobin, is defined as the concentration of O_2Hb expressed as a percentage of only the hemoglobin available for oxygen binding.

$$SO_2\,(\%) = \frac{100\,C_{O_2Hb}}{C_{O_2Hb} + C_{Hb}}$$

Thus, SO_2 can also be calculated from the measured concentrations of O_2Hb and Hb, and many oximeters perform this calculation automatically.

Comparison of these two equations shows that the two quantities are equal if no COHb or metHb is present. This is partly why the two are often confused, and the term *oxygen saturation* is unfortunately still loosely used to refer to either oxygen saturation or O_2Hb fraction. However, it is quite important for the laboratory to clearly identify which quantity is being reported so that clinicians are not confused when treating a patient with carbon monoxide poisoning or methemoglobinemia.

Oxygen Content

The concentration of total oxygen (CtO_2) in blood, also called oxygen content, is the sum of hemoglobin-bound oxygen and oxygen dissolved in blood. Oxygen content can therefore be calculated as

$CtO_2 = CO_2Hb + \alpha o_2\ PO_2$

where αO_2 is the solubility coefficient of oxygen in blood plasma.

Because 98% to 99% of the total oxygen is normally bound to hemoglobin, the term for dissolved oxygen is often ignored, in which case O_2 content can be closely estimated from oximeter data alone. Oxygen content is often expressed in milliliters of oxygen (at standard temperature and pressure) per deciliter of whole blood, and the measured quantities available are O_2Hb fraction and total hemoglobin. Oxygen content is then estimated as

CtO_2 (mL O_2/dL) = 1.39 FO_2Hb CtHb

where CtHb is total hemoglobin in grams per deciliter and 1.39 is the oxygen binding capacity of hemoglobin expressed as milliliters of O_2 per gram of hemoglobin.

P_{50}

The partial pressure of oxygen corresponding to an oxygen saturation of 50% is called P_{50} and is a measure of hemoglobin–oxygen affinity. In principle, if one assumes a normal shape of the oxygen dissociation curve and measures both the PO_2 and oxygen saturation, then this point on the curve can be used to calculate the PO_2 that would give 50% saturation. In practice, this can be done accurately as long as the oxygen saturation is less than approximately 90%. To interpret the calculated P_{50}, one must also measure the quantities mentioned above that affect the position of the oxygen dissociation curve. The determination of P_{50} is not commonly performed but can be useful, for example, in determining *in vivo* oxygen affinity or in testing for hemoglobin variants with abnormal oxygen affinity. Further details and the calculation procedure are given in reference 5.

Quality Control and Proficiency Testing

Quality control (QC) of PO_2 measurements is discussed with pH and PCO_2 in Chapter 17. This brief discussion relates to QC of oximetry measurements. One of the primary difficulties with QC and proficiency testing (PT) in blood gas analysis applies also to oximetry measurements, namely, the lack of a control material that closely resembles whole blood. Materials used for QC fall into three categories: (a) a dye mixture, (b) stabilized hemolysates, and (c) stabilized whole blood.

Dye mixtures are, in principle, adequate for QC of the spectrophotometer module of an oximeter. The principal drawbacks are that readings may not correspond to values encountered in patient samples and that different readings are usually obtained on various makes and models of oximeters. These problems are avoided by using a stabilized hemolysate. The main limitation of both these matrices is that they cannot check the functioning of the hemolyzer portion of the oximeter. This is important because incomplete hemolysis of samples can result in inaccurate readings that may not be noticed; for example, the errors may appear as moderate elevations of COHb or metHb. Although stabilized whole blood would seem to be a logical matrix, the products commercially available may not work satisfactorily in all oximeters. For example, stabilized erythrocytes may be more resistant to lysis than normal red cells, leading to erratic results on QC materials even when the instrument is functioning well with patient samples.

A useful adjunct to any QC program is to run occasional patient samples in duplicate on two different instruments. This is one way of using fresh whole blood for QC, and this should be done whenever two or more instruments are used. In addition, the accuracy of the total hemoglobin measured by an oximeter can be checked by using fresh whole blood together with the cyanmethemoglobin reference method for total hemoglobin.

Point-of-Care Testing

Some tests, including blood gas analysis and oximetry, can now be performed at the patient's bedside. This has several advantages but also raises some concerns (6,7 and 8). Advantages include minimizing turnaround time, limiting iatrogenic blood loss and patient discomfort, and eliminating problems relating to specimen transport (9,10). Areas of concern include cost, adequate quality control, adequate training of nonlaboratory personnel, and differences in reference ranges that may affect the interpretation of results. Several problems are particular to the acute-care setting. Precision, for example, may be more important than accuracy in situations in which an acute change in an analyte is more important than absolute values (11,12).

Important parameters in the critical care setting include the PO_2, the O_2 saturation, the PCO_2, pH, electrolytes (Na^+, K^+, Ca^+), and glucose. Table 18.2 summarizes the methodologies employed by a number of bedside instruments.

TABLE 18.2. Methodologies Used in Point-of-Care Instrument

Analytes	Electrochemical	Optical
Po_2	Amperometric Clark-type electrode	Optode based on: •Fluorescence quenching, or •Absorbance change of O_2 binding molecules
Pco_2	Potentiometric Severinghaus-type electrode	Optode based on pH sensor and gas permeable membrane
pH	Glass or polymer membrane electrodes	Optode based on fluorescence or absorbance measurement of immobilized pH indicator
Electrolytes	Ion selective electrodes (potentiometric)	Solution phase complexing reagents or colorimetric strip test
Hematocrit	Conductivity cell or ISEs (measure change in electrolytes after RBC lysis)	Microcentrifuge or cell counter
Total hemoglobin	—	Direct absorbance measurement or colorimetric measurement of pseudoperoxidase activity
Glucose	Amperometric enzyme electrodes based	Colorimetric measurement of glucose on immobilized glucose oxidase metabolite

Adapted from Misiano DR, Meyerhoff ME, Collison ME. Current and future directions in the technology relating to bedside testing of critically ill patients. *Chest* 1990 (Suppl);204S–214S.

Bedside testing instrumentation can be divided into four major groups:

1. Continuous noninvasive methods,
2. Discrete sample analyzers,
3. Extracorporeal sensing systems,
4. Continuous invasive methods.

Continuous Noninvasive Methods

Continuous noninvasive analyte testing is conceptually the ideal. Real-time continuous values are provided without the need to withdraw blood or use indwelling catheters. Currently, this technology is limited to the transcutaneous measurement of oxygen and carbon dioxide, and O_2 saturation by pulse oximetry. Future innovations may allow for noninvasive measurement of other analytes, including electrolytes and glucose.

Transcutaneous sensors use methods similar to conventional blood gas instruments (13,14). Clark-style amperometric oxygen and Stow-Severinghaus–style potentiometric carbon dioxide sensors have modified electrode designs to allow attachment of the monitor to the skin. Correlation studies comparing conventional PO_2 and PCO_2 measurements to transcutaneous values demonstrate significant biases, particularly in adults. Establishing appropriate reference ranges is therefore essential. Transcutaneous measurements are most commonly used for monitoring infants and neonates, in whom the discrepancies from established methods are less pronounced. Differences between the transcutaneous PO_2 and the PaO_2 are particularly pronounced in

shock and low-flow states. In some cases, this discrepancy is helpful to identify patients with these disorders (15).

Pulse oximetry measures the percentage of O_2 saturation of hemoglobin by the transmission of light at 660 nm (O_2Hb peak) and 940 nm (Hb peak) through a finger. The relative absorbance at each wavelength is measured and a ratio of O_2Hb to total hemoglobin is calculated. The measurement is conducted to coincide with the arterial pulse, thereby reflecting arterial blood (16). Advantages of pulse oximetry include simplicity, the performance of real-time measurements, and blood conservation. However, potential technical or mechanical problems may arise in patients with hypothermia, peripheral vascular disease, hyperbilirubinemia, or altered heme species or from stray ambient lighting (1,2,17,18 and 19). Nevertheless, pulse oximetry is one of the most widely used continuous noninvasive methodologies, particularly in settings in which early warning of sudden change is essential.

Discrete Sample Analyzers

Discrete sample analyzers include single-use, disposable analyzers (e.g., glucose meters) and multitest (e.g., PO_2, PCO_2, pH, Na, K, Ca, hematocrit, hemoglobin, glucose) analyzers. All these instruments require withdrawing blood from the patient, and therefore the test results are not true real-time values. The multitest analyzers may be placed in a central location in the critical care unit and function in a manner similar to that of a “stat lab.” Smaller, more portable analyzers may be moved closer to the patient for easier access and more rapid results. As instruments move from the central laboratory to the bedside, issues of QC and operator training and proficiency become more problematic (8). Establishing protocols for the use of bedside devices is mandated by the Clinical Laboratory Improvement Act (CLIA), and is often the responsibility of the clinical laboratory.

Extracorporeal Sensing Systems

Extracorporeal sensors are instruments that sample arterial blood, measure analytes, and return the blood to a vein (in-line) or pass the blood to a waste receptacle (on-line). The advantage of these systems is that they can provide real-time measurements of analytes by conventional methods. The disadvantages include the risk of infection, thrombotic complications, and mechanical damage to erythrocytes, all of which mandate short-term use of these devices. The use of extracorporeal monitoring systems is currently limited to dialysis and cardiac surgery settings.

Continuous Invasive (*In Vivo*) Methods

Continuous invasive monitoring is achieved by attaching the sensor directly to an indwelling catheter. Analytes are measured continuously and in real time at sites such as the pulmonary artery. Catheter tips with miniaturized O_2, CO_2, and pH sensors have been devised, and implantable sensors for glucose are being developed (20). Disadvantages include the complications usually associated with indwelling catheters, such as thrombosis and infection. The present technological obstacle is the development of miniaturized, durable, and biocompatible sensors that can function reliably at the end of an indwelling catheter.

Mechanical and technical variables remain the major obstacles to refining and implementing cost-effective point-of-care testing. Newer technologies, such as near-infrared spectroscopy (for the transcutaneous measurement of blood glucose) (21,22 and 23), magnetic resonance spectroscopy (24,25), and diffuse-sink sampling (26), as well as advances in instrumentation (27) have the potential to significantly advance point-of-care testing and improve patient care.

PATHOPHYSIOLOGY AND INTERPRETATION

Part of "18 - Respiration and Measurement of Oxygen and Hemoglobin"

Respiratory failure is a clinical condition that is best defined by abnormalities in arterial blood gas values. Although in some cases the diagnosis is obvious based on the patient's signs and symptoms (e.g., dyspnea, respiratory rate, cyanosis), the diagnosis usually requires analysis of arterial blood gases (PO_2, PCO_2, and pH). The clinical signs and symptoms of hypoxemia and hypercapnia are unreliable in predicting blood gas values. Reference intervals are shown in Table 17.4 of Chapter 17. As a general guideline, respiratory failure may be defined as a PaO_2 less than 50 to 60 mm Hg or a $PaCO_2$ greater than 50 mm Hg. In acute respiratory failure, the arterial pH may be decreased owing to retention of CO_2 (acute respiratory acidosis), whereas in chronic respiratory failure, the pH is less affected because of renal compensation. The evaluation of blood gas results is closely related to the topic of acid-base balance (see Chapter 17). This section is limited to a discussion of respiratory failure, and the differential diagnosis of hypoxemia (decreased PO_2) and hypercapnia (elevated PCO_2).

It is important to distinguish between the terms hypoxemia and hypoxia. Hypoxemia refers to a decreased blood PO_2, whereas hypoxia is a condition in which inadequate amounts of oxygen are available to the tissues. Hypoxemia is one cause of hypoxia, but hypoxia can exist with normal PO_2 (e.g., arterial thrombosis). Delivery of oxygen to the tissues depends on cardiac output, intact vasculature, and the oxygen content of the blood. Clinical assessment of these variables is an important adjunct to arterial blood gas measurements.

Respiratory failure may be viewed in terms of pathophysiologic mechanisms (hyperventilation, [V with dot above]/[Q with dot above] mismatch, shunts) or from an anatomic (pulmonary versus extrapulmonary) or etiologic perspective (e.g., pneumonia). Each approach has its advantages and is discussed below. An excellent discussion of this topic may also be found in reference 28.

Anatomic Analysis

One approach to respiratory failure is to consider whether the underlying problem is the lungs or outside the lungs. For example, extrapulmonary causes of respiratory failure include disorders of the respiratory center, neuromuscular system, chest wall, pleura, or upper airway. All these are associated with hypoventilation, hypoxemia, and hypercapnia. Respiratory failure may result from diseases of the lower airways, the pulmonary circulation or the pulmonary interstitium resulting in hypoxemia with or without hypercapnia. In a patient with hypercapnia, the main problem is inadequate ventilation, whereas in respiratory failure without CO_2 retention the main problem is with oxygenation.

Pathophysiologic Approach

This approach divides hypoxemia or hypercapnia according to the pathophysiologic mechanism. After the mechanism has been identified, the specific etiology can be determined.

Hypoxemia may result from one of five pathophysiologic mechanisms, as shown in Table 18.3. Breathing air with a low oxygen content may be seen in subjects at high altitudes or other unusual situations but does not occur in a typical patient population. A similar statement can be made regarding diffusion impairment as a cause of hypoxemia. In principle, diffusion impairment can occur whenever the alveolar-capillary membrane is thickened or infiltrated by fluid, fibrosis, or other substances. In practice, diffusion impairment contributes little to hypoxemia in most pulmonary diseases. This is because gas exchange occurs after only one third of the time that blood spends in transit in the alveolar capillary. Severe pulmonary interstitial fibrosis or edema, pulmonary amyloidosis, and some types of pneumoconioses may produce hypoxemia as a consequence of diffusion impairment, but this situation is uncommon. Arterial blood gas values in patients with significant diffusion impairment typically show hypoxemia often with hypocapnia and a greatly increased A-a gradient. There is often a coexisting ventilation–perfusion mismatch.

TABLE 18.3. Pathophysiologic Classification of Respiratory Failure

Mechanism	Blood Gas Values	A-a O_2 Gradient	Response to O_2
Low inspired oxygen	Low Po_2 and low/normal Pco_2	Normal	Increased Po_2
Hypoventilation	Low Po_2, elevated Pco_2	Normal	Increased Po_2
Ventilation/perfusion ([V with dot above]/[Q with dot above]) mismatch	Low Po_2, elevated Pco_2 when severe	Increased	Increased Po_2
Right to left shunt	Low Po_2 and normal Pco_2	Increased	Minimal improvement
Diffusion impairment	Low Po_2 and elevated Pco_2	Increased	Minimal improvement

The majority of patients with hypoxemia can be classified into one of three pathophysiologic mechanisms: hypoventilation, ventilation–perfusion mismatch, or shunting as shown in Figure 18.4. In some patients, more than one of these mechanisms may be operating simultaneously.

Hypoventilation

Hypoventilation produces an increase in arterial PCO_2 (CO_2 retention, or hypercapnia) and a decrease in arterial PO_2 (hypoxemia). Hypoventilation may result from a number of extrapulmonary disorders including diseases that affect the central respiratory center (heroin overdose, stroke), neuromuscular disorders (Guillain-Barré syndrome, myasthenia gravis, poliomyelitis), abnormalities of the chest wall or diaphragm (trauma, diaphragmatic paralysis, kyphoscoliosis), upper airway obstruction (foreign bodies, epiglottitis, tumors), or restrictive pleural disease. It is important to recognize that hypoventilation is not caused by diseases of the lung itself (excluding the pleura). Consequently, gas exchange is normal and the A-a O_2 gradient is also normal.

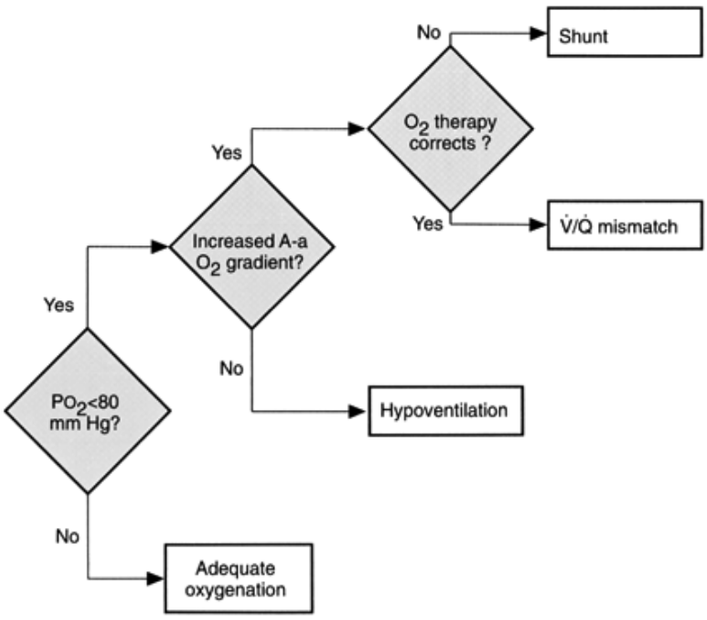


FIGURE 18.4. Algorithm for the evaluation of hypoxemia.

The treatment of hypoventilation includes maneuvers to improve ventilation to correct the hypoxemia and associated respiratory acidosis owing to hypercapnia (Fig. 18.4).

Ventilation–Perfusion Mismatch

Ventilation–perfusion mismatch is the most common cause of hypoxemia. Physiologic ventilation–perfusion mismatching occurs in normal subjects as described previously. There are two types of ventilation–perfusion mismatch: The first type is called low [V with dot above]/[Q with dot above] and refers to inadequate ventilation relative to perfusion. Conversely, high [V with dot above]/[Q with dot above] refers to inadequate perfusion relative to ventilation.

Areas of low [V with dot above]/[Q with dot above] result in a decrease in the PO_2 and the oxygen content of the affected pulmonary blood. This can be partially offset by a slight increase in oxygen content in areas of high [V with dot above]/[Q with dot above]. However, the compensatory effect is inadequate because of the fact that hemoglobin in the pulmonary capillaries is saturated with oxygen. Low [V with dot above]/[Q with dot above] decreases the oxygen content more than the areas of high [V with dot above]/[Q with dot above] can increase it. Mixing of the pulmonary capillary blood in the pulmonary venous system and left heart results in a low PO_2, a low oxygen content, and an increased A-a gradient. The PCO_2 may or may not be elevated. Retention of CO_2 in a patient with a V:Q mismatch is generally observed only with severe ventilation–perfusion mismatching. Part of the explanation for this is that diffusion of CO_2 across the alveolar-capillary membrane is more efficient than diffusion of oxygen. Consequently, the lung can usually eliminate excess CO_2 unless the ventilation–perfusion mismatch is severe.

The differential diagnosis of ventilation–perfusion mismatch is extensive and includes some of the most common pulmonary disorders encountered in clinical practice (e.g., asthma, emphysema, atelectasis, pneumonia, pulmonary embolus, adult RDS (ARDS), pneumothorax, lower airway obstruction, pulmonary edema). As with hypoventilation, hypoxemia owing to ventilation–perfusion mismatching may respond favorably to oxygen administration.

Shunting

Arteriovenous (right-to-left) shunts are defined as any process that permits unoxygenated venous blood to mix with oxygenated arterial blood. The term for mixing of blood from nonventilated (or markedly underventilated) airspaces with blood from ventilated areas is called venous admixture and can be expressed as a percentage of the cardiac output. Shunting of venous blood into the arterial circulation (including pulmonary arterial blood into the pulmonary venous system) reduces the arterial O_2 content and PO_2 and results in an increased A-a O_2 gradient. Shunts may be normal (anatomic) or pathologic.

Normal anatomic (physiologic) shunts contribute a small amount to venous admixture. The thebesian veins, which drain the myocardium, partially empty into the pulmonary veins, and intrapulmonary anatomic arteriovenous shunts bypass ventilated areas. Collectively, these normal shunts account for approximately 5% of right heart output.

Pathologic shunts may be intrapulmonary or extrapulmonary. Right-to-left cardiac shunts are examples of extrapulmonary shunts, as are shunts occurring in the great vessels. Most intrapulmonary shunts represent an extreme case of ventilation–perfusion mismatch and occur when blood perfuses areas of lung in which the alveoli are not ventilated (atelectasis, obstruction) or are filled with fluid or blood (e.g., pulmonary edema, pneumonia, hemorrhage). Thus, the differential diagnosis of

right-to-left shunts overlaps to some degree with conditions causing ventilation–perfusion mismatch, and many patients exhibit features of both. In contrast to patients with hypoventilation or ventilation–perfusion mismatching, patients with shunts exhibit a poor response to oxygen therapy.

Hypoxemia may occur in the presence or absence of arterial hypercapnia. Hypercapnia may result from one of three mechanisms.

1. Breathing air with a high PCO_2. This mechanism is not relevant in the usual clinical setting but may occur in some states of suffocation or in subjects rebreathing expired air (e.g., in the treatment of respiratory alkalosis).
2. Hypoventilation.
3. Severe ventilation–perfusion mismatch.

The pathophysiologic approach to respiratory failure can be extended to include three major subtypes (Table 18.4). The first is called type 1 oxygenation failure, characterized by a low PO_2, a normal or elevated PCO_2, and an elevated A-a gradient. Type 1 respiratory failure is caused by V:Q mismatch or shunts. In contrast, type 2 respiratory failure (or ventilation failure) is characterized by a low PO_2, an elevated PCO_2, and a normal A-a gradient. Finally, type 3 respiratory failure (or combined failure) is characterized by a low PO_2, an elevated PCO_2, and an increased A-a gradient. Type 3 failure occurs when ventilation cannot be increased sufficiently to maintain a normal PCO_2, in a patient with a marked V:Q mismatch. The result is hypoxemia, hypercapnia, and an increased A-a gradient. Type 3 respiratory failure most often occurs in patients with diffuse alveolar damage, severe asthma, or chronic obstructive pulmonary disease (COPD).

TABLE 18.4. Respiratory Failure: Pathophysiologic Approach

Failure Type	Po_2	Pco_2	A-a Gradient	Etiology
Oxygenation failure	Low	Normal/elevated	Elevated	V:Q mismatch/shunt
Ventilation failure	Low	Elevated	Normal	Hypoventilation
Combined failure	Low	Elevated	Elevated	Combined oxygenation/ventilation failure

Specific Diseases

The differential diagnosis of hypoxemia and hypercapnia encompasses a large number of disease entities. A detailed discussion of this topic is beyond the scope of this chapter, but some additional comments concerning the more common disease entities are given.

Patterns of Breathing

Various patterns of breathing may suggest specific disease states or result from compensatory mechanisms for ineffective respiration. Hyperventilation may be a pathologic primary event (resulting in respiratory alkalosis) or may result from stimulation by the chemoreceptors as an appropriate compensatory response to correct acidosis and hypoxia. Likewise, hypoventilation may be pathologic, resulting in respiratory acidosis, or may occur as a compensatory response in the setting of metabolic alkalosis.

Two examples of clinically significant abnormal patterns of breathing are Cheyne-Stokes and Kussmaul respirations. Cheyne-Stokes breathing is alternating hyperventilation and hypoventilation with periods of apnea. It may result from asynchronous sensory output from peripheral and central chemoreceptors and can occur in patients with congestive heart failure or cerebral vascular accidents. Kussmaul breathing is a regular, deep, and rapid breathing pattern seen in severe metabolic acidosis (such as diabetic ketoacidosis).

Restrictive Versus Obstructive Lung Disease

Spirometry is a frequently used method to assess the ventilatory function of the lungs, although it is usually performed in a location outside the clinical laboratory. Spirometry can be used to classify certain lung diseases into two categories termed restrictive and obstructive (see below). Spirometry is useful in monitoring patients with chronic pulmonary diseases to assess the progress of disease activity.

Chronic Obstructive Pulmonary Disease

COPD refers to a group of related disorders characterized by chronic, progressive airway obstruction resulting either from narrowing of the bronchial lumina as a consequence of inflammation and mucous hypersecretion (chronic bronchitis) or constriction (bronchial asthma), or loss of lung parenchyma and elastic recoil (emphysema). Proximal bronchial narrowing and mucous plugging leads to inadequate alveolar ventilation (ventilation–perfusion mismatch), hypoxemia, and, when severe, CO_2 retention. A component of pulmonary parenchymal shunting is also typically present. The major physiologic defect in COPD is an increase in resistance to airflow with limitation of expiratory airflow rates during expiration. COPD can be divided into two clinical types.

Type A patients have been called “pink puffers” or “fighters” because they increase ventilation to prevent hypercapnia. Type B patients are termed “blue bloaters” or “nonfighters” because their ventilatory response is blunted. The former generally have a mildly decreased or normal PO_2 and a normal PCO_2, whereas the latter exhibit a low PO_2 and an elevated PCO_2. Although type A is typically associated with emphysema and type B with chronic bronchitis, most patients with COPD have features of both. Hypoxia may be the primary mechanism driving ventilation in advanced COPD. Overzealous oxygen therapy may therefore aggravate hypercapnia and respiratory acidosis by decreasing the patient's respiratory drive. For this reason, oxygen is usually given judiciously to maintain a moderately low PO_2.

Diffuse Interstitial Diseases

In contrast to COPD, which results from airway obstruction, diffuse interstitial diseases are characterized by a restrictive process producing a reduced expansion of the lung and a decreased total lung capacity. Although patients with COPD cannot expire normally, patients with restrictive diseases typically exhibit a normal rate of airflow. Thus, the respiratory cycle as seen on spirometry is normal but with reduced volumes. The differential diagnosis of diffuse interstitial disease is extensive and includes disorders induced by occupational and environmental inhalants, drugs, and toxins. All are characterized by alveolitis and variable degrees of interstitial fibrosis. The primary underlying pathophysiology is inadequate ventilation, and hence interstitial diseases are characterized by ventilation–perfusion mismatch. Interstitial fibrosis results in decreased lung compliance, increased work of ventilation, and a rapid and shallow breathing pattern. A minor contributor to hypoxemia is impaired diffusion across a thickened alveolar-capillary membrane, but this occurs only in severe cases. Blood gas measurements may be normal in mild disease but may demonstrate a fall in the PO_2 with exercise. The A-a O_2 gradient is typically increased. Severe disease produces increasing hypoxemia and hypercapnia.

Atelectasis

Atelectasis is characterized by incomplete expansion of the lung (or part of the lungs) or collapse of previously inflated tissue. Atelectasis may result from many causes including airway obstruction with resorption of residual air, inadequate respiratory effort, extrinsic compression of the lung (e.g., pneumothorax), contraction of pleural or pulmonary scars, or loss of pulmonary surfactant (e.g., adult RDS).

Pulmonary Edema

Theoretically, pulmonary edema can result from alteration of any of the components of Starling's law. Increased hydrostatic pressure occurs in left ventricular failure, mitral stenosis, and pulmonary vein occlusion. Decreased plasma oncotic pressure can be seen in hypoproteinemic states such as nephrotic syndrome, cirrhosis, and massive crystalloid infusion. The capillary membrane permeability may be altered in septic shock, pneumonia, lung trauma, oxygen toxicity, disseminated intravascular coagulation, systemic lupus erythematosus, and many other conditions. Finally, compromise of lymphatic drainage may occur in lymphangitic carcinomatosis or radiation pneumonitis. The most common cause of pulmonary edema is left ventricular failure.

Generally, the interstitial compartment is the initial site of fluid accumulation. Eventually fluid accumulates in the pulmonary alveolar spaces as well. Compromised ventilation and subsequent hypoxemia usually are not present until alveolar filling occurs. In early (interstitial) pulmonary edema, the PO_2, PCO_2, and respiratory rate may be normal but exertion (e.g., climbing a flight of stairs) can cause dyspnea as the oxygen requirement is increased. Alveolar filling generally results in hypoxemia, an increased A-a O_2 gradient, a normal or decreased PCO_2, and dyspnea. The PCO_2 is increased only in very severe cases.

Adult Respiratory Distress Syndrome

ARDS (also called diffuse alveolar damage or shock lung) is a condition caused by alveolar capillary damage with rapid onset of severe respiratory failure. The differential diagnosis includes sepsis, shock, head injury, aspiration, smoke inhalation, drug overdose, pancreatitis, burns, oxygen toxicity, and diffuse pneumonia (especially viral) and a long list of other conditions. The pathophysiology of ARDS involves derangement of alveolar-capillary permeability, resulting in extravasation of fluid into the lung. The initial injury is to the capillary endothelium, but ultimately the epithelium is also affected (29). The mechanism of alveolar capillary damage depends on the etiology and may include aggregation of leukocytes in the pulmonary vessels, oxygen-derived free radicals, toxins, and other factors (29). The morphologic features of ARDS include congestion, edema, epithelial cell necrosis, and the formation of hyaline membranes. Clinically, patients exhibit hypoxemia, decreased pulmonary compliance, and diffuse pulmonary infiltrates. The hypoxemia is only minimally responsive to oxygen therapy because of pulmonary vascular shunting but may respond to positive end-expiratory pressure induced by mechanical ventilation.

Pulmonary Embolism

Pulmonary embolism is the principal cause of death in approximately 10% of hospitalized patients. Nonfatal emboli are even more common. Ordinarily pulmonary emboli do not cause infarction of the lung because the bronchial circulation is usually sufficient to prevent ischemic necrosis of the lung parenchyma. However, infarcts may occur in patients with preexisting cardiac or lung disease. Most emboli arise from thrombi that form in the deep veins of the legs, but some may originate from thrombi in the pelvic veins or other sites. Large pulmonary emboli are one cause of sudden death, whereas smaller emboli can produce a spectrum of symptoms from none at all to severe respiratory distress. Arterial blood gases show a low PO_2, often a low PCO_2 (owing to hyperventilation) and an increased A-a O_2 gradient (owing to severe ventilation–perfusion mismatch).

Pulmonary Infections

Many organisms may infect the lungs, and the degree of respiratory compromise is also quite variable. Inflammation, edema, congestion, and hemorrhage may produce hypoxemia, an increased A-a O_2 gradient, and a normal, elevated, or decreased PCO_2, depending on the degree of intrapulmonary shunting and ventilation–perfusion mismatch. Initially, a fall in the arterial PO_2 induces hyperventilation, which may lower the PCO_2, but as with other causes of respiratory failure, with severe disease, there may be CO_2 retention.

Methemoglobinemia

In normal hemoglobin, the iron atoms are in the reduced (ferrous) state. Oxidation of the iron to the trivalent ferric state produces a brown hemoglobin (metHb). metHb cannot combine with oxygen, and if the concentration of metHb is significantly elevated,

cyanosis and hypoxia (not hypoxemia) occur. MetHb is normally produced in small quantities and is reduced back to the ferrous state by various red blood cell enzymes (metHb reductases) or by reducing agents (e.g., glutathione, ascorbic acid). The normal concentration of metHb is less than 1.5% of the total hemoglobin, whereas levels above 10% produce cyanosis, levels above 35% produce symptoms, and concentrations above 60% may be fatal.

Increased levels of metHb may be hereditary (owing to enzyme deficiencies, most commonly an autosomal recessive metHb reductase deficiency) or may be acquired (owing to oxidant drugs or toxins, e.g., nitrites, nitrates, sulfones, aniline dyes, sulfonamides, phenacetin). Homozygous metHb reductase deficiency produces elevated metHb levels with cyanosis. Heterozygotes exhibit normal metHb concentrations unless subjected to an oxidant stress. Rarely, methemoglobinemia is caused by abnormal hemoglobins called (hemoglobin Ms), in which the iron atom cannot be reduced from the ferric state.

Acquired methemoglobinemia (and that occurring in reductase-deficient heterozygotes) results from ingestion of a variety of chemicals and drugs that oxidize hemoglobin.

Methemoglobinemia may be treated by the administration of methylene blue, which accelerates the activity of the reductase system, or reducing agents such as ascorbic acid, glutathione, cysteine, and British anti-Lewisite (dimercaprol).

Carboxyhemoglobinemia

COHb normally constitutes a small fraction (0.5%) of the total hemoglobin. It is formed by the reversible interaction of carbon monoxide with hemoglobin. Carbon monoxide binds hemoglobin with high affinity (210 times that of O_2), resulting in a left shift of the oxygen dissociation curve and a characteristic cherry red color of the blood. Because of the high affinity of CO for hemoglobin, COHb accumulates in acute or chronic CO exposure. Inhalation of CO may occur in subjects exposed to the products of incomplete combustion such as occurs in fires, charcoal grills, heating appliances, internal combustion engines, and tobacco products. Smokers may exhibit COHb levels of 4% to 8% of total hemoglobin.

In some patients, symptomatic CO toxicity (exertional dyspnea) can occur at levels as low as 10%, but overt symptoms usually do not occur until the blood level exceeds 30% (headache, syncope, confusion). Levels above 60% produce loss of consciousness or death. Oxygen therapy is effective treatment for CO poisoning because the CO–hemoglobin bond is reversible. The half-life of COHb in a patient breathing room air is approximately 4 to 5 hours, and this can be shortened to approximately 1 hour by breathing 100% oxygen.

REFERENCES

1. Barker SJ, Tremper KK, Hyatt J. Effects of methemoglobinuria on pulse oximetry and mixed venous oximetry. *Anesthesiology* 1989;70:112–117.

2. Barker SJ, Tremper KK. The effect of carbon monoxide inhalation on pulse oximetry and transcutaneous PO_2. *Anesthesiology* 1987;66:677–679.

3. *Definitions of quantities and conventions related to blood pH and gas analysis*, 2nd ed. NCCLS document C-12T2. Villanova, PA: The National Committee for Clinical Laboratory Standards, 1991.

4. Burnett RW, et al. Definitions of quantities and conventions related to blood gases and pH. *J Int Fed Clin Chem* (in press).

5. Wimberley PD, Burnett RW, Covington AK, et al. Guidelines for routine measurement of blood hemoglobin oxygen affinity. *Scand J Clin Lab Invest* 1990;50(Suppl)203:227–234.

6. Misiano DR, Meyerhoff ME, Collison ME. Current and future directions in the technology relating to bedside testing of critically ill patients. *Chest* 1990;97:204s–214s.

7. Zaloga GP. Evaluation of bedside testing options for the critical care unit. *Chest*1990;97:185s–190s.

8. Pemberton JO. The role of the laboratorian on the critical care team. *Med Lab Observer* 1991;23:16–19.

9. Chernow B. The bedside laboratory. A critical step forward in ICU care. *Chest* 1990;97:183s–184s.

10. Rock RC. Why testing is being moved to the site of patient care. *Med Lab Observer* 1991;23:2–5.

11. Thorson SH, Marini JJ, Pierson DJ, et al. Variability of arterial blood gas values in stable patients in the ICU. *Chest* 1983;84:14.

12. Zaloga GP. Monitoring versus testing technologies: present and future. *Med Lab Observer* 1991;23:20–31.

13. Severinghaus JW. Transcutaneous blood gas analysis. *Respir Care* 1982;27:152–159.

14. Tremper KK, Barker SJ. Transcutaneous oxygen-studies and adult application. *Int Anesthesiol Clin* 1987;25:67–96.

15. Tremper KK, Waxman K, Bowman R, et al. Continuous transcutaneous oxygen monitoring during respiratory failure, cardiac decompensation, cardiac arrest, and CPR. *Crit Care Med* 1980;8:377–381.

16. Mendelson Y. Pulse oximetry: theory and applications for noninvasive monitoring. *Clin Chem* 1992;38:1601–1607.

17. Barrington KJ, Finer NN, Ryan CA. Evaluation of pulse oximetry as a continuous monitoring technique in the neonatal intensive care unit. *Crit Care Med* 1988;16:1147–1153.

18. Mansouri A. Methemoglobinemia. *Am J Med Sci* 1985;289:200–209.

19. Eisenkraft JB. Pulse oximeter desaturation due to methemoglobinemia. *Anesthesiology* 1988;68:279–282.

20. Wilson GS, Zhang Y, Reach G, et al. Progress toward the development of an implantable sensor for glucose. *Clin Chem* 1992;38:1613–1617.

21. Robinson MR, Eaton RP, Haaland DM, et al. Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation. *Clin Chem* 1992;38:1618–1622.

22. Hall JW, Pollard A. Near-infrared spectrophotometry: a new dimension in clinical chemistry. *Clin Chem* 1992;38:1623–1631.

23. Rosenthal R. Research into noninvasive measurement of blood glucose by near-infrared technology. *Clin Chem* 1992;38:1645(abst).

24. Hornung PA, Schuff N. Noninvasive imaging and spectroscopy broad applications of magnetic resonance. *Clin Chem* 1992;38:1608–1612.

25. Otvos JD, Jeyarajah EJ, Bennett DW, et al. Development of a proton nuclear magnetic resonance spectroscopic method for determining plasma lipoprotein concentrations and subspecies distributions from a single, rapid measurement. *Clin Chem* 1992;38:1632–1638.

26. Wade SE. Less invasive measurement of tissue availability of hormones and drugs: diffuse-sink sampling. *Clin Chem* 1992;38:1639–1644.

27. Schembri CT, Ostoich V, Lingane PJ, et al. Portable simultaneous multiple analyte whole-blood analyzer for point-of-care testing. *Clin Chem* 1992;38:1665–1670.

28. Irwin R, Pratter M. A physiologic approach to managing respiratory failure. In: Rippe J, Irwin R, Alpert J, et al., eds. *Intensive care medicine*, 2nd ed. Boston: Little, Brown, 1991:449–454.

29. Cotran R, Kumar V, Robbins S. The respiratory system. In: Cotran R, Kumar V, Robbins S, eds. *The pathologic basis of disease*, 4th ed. Philadelphia: Saunders, 1989:755–810.

19

Nitrogen Metabolites and Renal Function

Alfred E. Hartmann

- RENAL PHYSIOLOGY
- WATER BALANCE, SPECIFIC GRAVITY, AND OSMOLALITY
- UREA
- CREATININE
- CLEARANCE TESTS
- CREATINE
- URIC ACID
- AMMONIA
- RENAL DISORDERS

RENAL PHYSIOLOGY

Part of "19 - Nitrogen Metabolites and Renal Function"

The kidneys are regulatory organs responsible for the maintenance of extracellular fluid volume and electrolyte and acid base balance, along with the excretion of toxic metabolites, drugs, and metabolic waste products. The kidney accomplishes this homeostasis by the formation of urine through ultrafiltration in the glomerulus, selective tubular reabsorption of water and solutes, and selective tubular secretion of solutes. The kidney also has an endocrine role; it produces erythropoietin, which regulates red cell mass, and renin, which affects sodium balance and regulates blood pressure, and metabolizes vitamin D, which influences calcium balance.

The two bean-shaped kidneys, which lie in the retroperitoneal space on either side of the aorta, are each approximately 10 to 12 cm in length and weigh approximately 150 g. Each kidney has an outer cortex that averages 40 mm in thickness surrounding the inner medulla. The medulla consists of a number of conical renal pyramids that terminate in a papilla. The papillae protrude into the renal pelvis, which collects urine. The renal pelvis drains into the ureter connecting the kidney to the urinary bladder. Within the cortex and medulla are the functional units of the kidney, called the nephrons. Each kidney contains approximately 1.25×10^6 nephrons, which operate in parallel with each other. A nephron is composed of a glomerulus, proximal convoluted tubule, loop of Henle, distal convoluted tubule, and collecting duct.

The glomerulus is composed of a cluster of capillaries formed from an afferent arteriole and is surrounded by an epithelial saccular expansion of the proximal tubule, forming Bowman's space. The proximal convoluted tubule coils extensively in the cortex and penetrates the deeper layers of the cortex to the outer medulla. The thin descending and thin and then thick ascending loops of Henle are located in the medulla in a hairpin configuration. This hairpin configuration allows a countercurrent multiplication concentrating process to occur. The distal convoluted tubule arises from the ascending thick loop of Henle and traverses through the cortex, terminating in a collecting duct near the original glomerulus. The collecting duct is formed in the outer cortex by the junction of two or more distal convoluted tubules and traverses through the cortex into the medulla, where several collecting ducts fuse to form the papillary ducts that terminate in the renal pyramid of the medulla. The total length of the nephron is 40 to 60 mm. The afferent arteriole leaves the glomerulus as the efferent arteriole and penetrates deep into the medulla, forming a network around the loop of Henle called the vasa recta. Tubular function depends on the anatomy of the loop of Henle and the surrounding vasa recta. The close proximity of the vasa recta to the loop of Henle is responsible for maintaining the medullary osmotic gradient, by a countercurrent exchange between the renal tubule and the corresponding arteriole (Fig. 19.1).

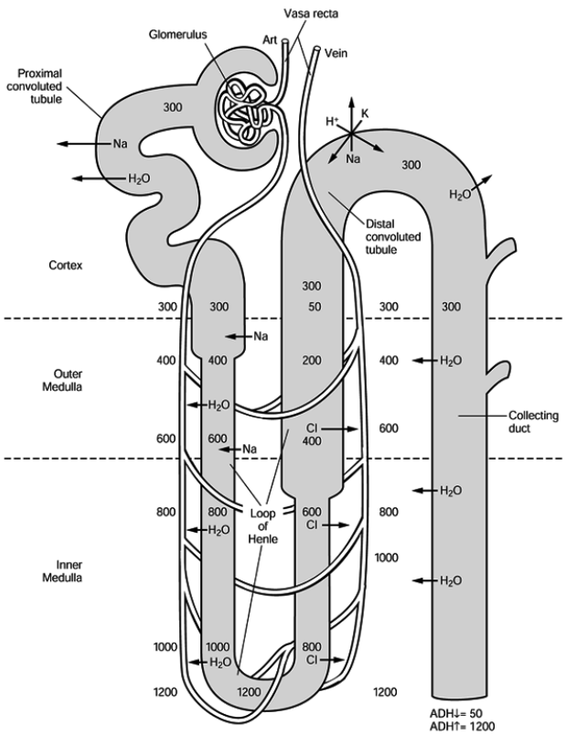


FIGURE 19.1. Diagram of the renal nephron. The numerals indicate the changes in osmolality (mOsm/kg H_2O) in the interstitium and the tubular fluid. The tubular exchanges of water and ions during the course of the production of urine are indicated.

Glomerular filtration is the initial event in the formation of urine. As blood flows through the glomerulus, ultrafiltration occurs owing to hydrostatic pressure within the glomerular capillaries, forcing plasma water and smaller solutes (<50,000 d) across the capillary membrane, opposed by the oncotic pressure of plasma proteins remaining in the glomerular capillaries and the hydrostatic pressure in Bowman's space. The glomerular filtration rate (GFR) depends on the balance of these Starling forces: the permeability of the capillary wall to water and solutes, the total surface area of the capillaries, and the rate of plasma flow through the capillaries. The GFR is the total glomerular filtrate formed by all the glomeruli per unit time. In the average-sized person, the GFR is 125 mL/minute or 180 L/day. The GFR can be determined with clearance studies using solutes that are filtered and subsequently neither reabsorbed nor secreted. The determination of the GFR provides an estimation of the functional renal mass and can be of considerable importance in the initial evaluation and follow-up of patients with renal disease. Certain extrarenal conditions will also cause the GFR to rise or fall significantly. Most important is the volume of the extracellular fluid. Expansion of the extracellular fluid volume is accompanied by an increase in the GFR rate, and an actual or perceived depletion of the extracellular fluid volume causes a marked reduction in the GFR rate. Postrenal obstruction of the outflow tract (urethra, bilateral ureters) will cause a decrease in the GFR proportional to the obstruction.

A large amount of ultrafiltrate (180 L/day) devoid of blood cells and protein containing 300 mOsm/kg H_2O of solute enters the tubules, and only 1% of that filtrate with a markedly altered solute content reaches the renal pelvis as urine. In the renal tubules, there is both selective reabsorption and secretion. Tubular reabsorption facilitates the conservation of substances essential for homeostasis. Examples are water, bicarbonate, glucose,

amino acids, and electrolytes. The reabsorption may be passive (e.g., urea and water) or active, defined as a net movement of the substance against an electrochemical potential gradient with the loss of metabolic energy (e.g., sodium or chloride). Tubular secretion involves the movement of substances from the tubular cell into the tubular fluid. Tubular secretion may also be passive or active. Many of the secreted substances are weak acids or bases, such as drugs, or metabolic end products that are eliminated from the body in the urine.

In the proximal convoluted tubule, 70% to 80% of the filtered solute and water are reabsorbed into the renal circulation. All the solutes are reabsorbed actively, except chloride, which is reabsorbed passively with sodium. Water is reabsorbed passively owing to the osmotic gradient from the active transport of sodium. At the end of the proximal convoluted tubule, the tubular fluid is sharply reduced in volume to approximately 20% of the original filtrate, whereas the osmolar concentration remains isotonic. As the tubular fluid flows through the loop of Henle, concentration of the filtrate first occurs owing to a countercurrent multiplication. In the thin descending loop of Henle, water diffuses out of the tubule into the hypertonic interstitium of the medulla, and a small amount of sodium diffuses into the tubule. At the hairpin curve deep in the medulla, the tubular fluid is maximally concentrated to nearly 1,200 mOsm/kg H_2O. In the thick ascending loop of Henle, the tubule is impermeable to water, whereas chloride is actively reabsorbed with sodium into the interstitium. The osmolar concentration decreases in the ascending loop of Henle until the tubular fluid is hypotonic (50 to 100

mOsm/kg H_2O) entering the distal convoluted tubule. The marked osmotic gradient established in the medullary interstitium is established by the active reabsorption of chloride followed by sodium in the ascending loop of Henle and the passive reabsorption of urea in the collecting duct. The rich capillary network (vasa recta) enveloping the tubule allows a countercurrent exchange to occur between the interstitial fluid and the renal blood flow. This exchange allows the reabsorption of water and electrolytes from the medullary interstitium into the renal circulation and is responsible for maintaining the hyperosmolality of the interstitium. In the distal convoluted tubule under the influence of aldosterone, 5% to 10% of the original filtered sodium is reabsorbed against a steep concentration gradient in exchange for hydrogen and potassium. In the collecting ducts, water reabsorption is regulated by antidiuretic hormone (ADH). Under the influence of ADH, the collecting ducts become permeable to water. Water is reabsorbed along with urea into the hyperosmolar medullary interstitium and tubular fluid becomes hypertonic. In the absence of ADH, the collecting duct is relatively impermeable to water and the tubular fluid remains hypotonic.

The renal nephron is also responsible for the maintenance of acid-base balance through the elimination of hydrogen ions and the regulation of plasma bicarbonate by conserving, regenerating, or eliminating bicarbonate. The renal tubular cells are capable of hydrogen ion secretion along much of the length of the nephron (Fig. 19.2). In the tubular cell, carbon dioxide is used as a source of hydrogen ion and bicarbonate. The hydrogen ion is secreted into the tubular lumen in exchange for a cation, usually sodium. In the proximal convoluted tubule, the hydrogen ion that is secreted into the tubular fluid is used up almost entirely in reabsorbing filtered bicarbonate, and the pH of the tubular fluid changes very little. In the distal convoluted tubule, under the influence of aldosterone, hydrogen ion and potassium ion are secreted into the tubular lumen in exchange for sodium. In addition to combining with bicarbonate in the distal tubular fluid, hydrogen ion can combine with other buffer ions such as hydrogen phosphate or can combine with ammonia to form ammonium ion. The secretion of hydrogen ion in the distal tubules changes the pH of the urine dramatically (less than pH 5.5) owing to the relatively small amount of bicarbonate remaining in the tubular fluid after proximal reabsorption of bicarbonate.

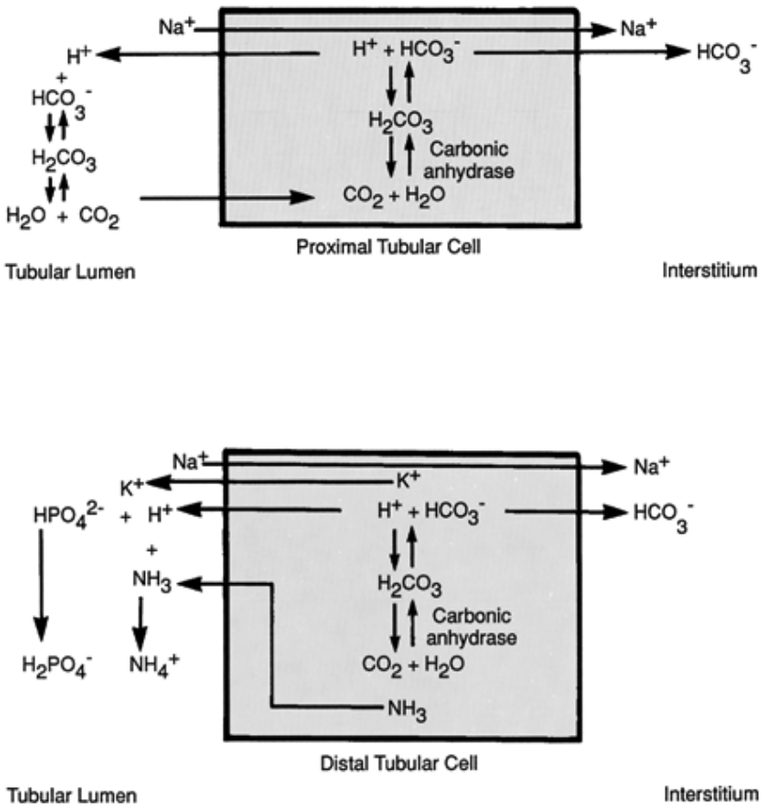


FIGURE 19.2. Hydrogen ion secretion in the proximal and distal renal tubular cells.

WATER BALANCE, SPECIFIC GRAVITY, AND OSMOLALITY

Part of "19 - Nitrogen Metabolites and Renal Function"

Water accounts for 50% to 60% of total body weight. Two thirds of total body water is intracellular, and one third is located in the extracellular fluid space. Water is freely diffusible across cell membranes. The distribution of water between the intracellular and extracellular spaces is determined by the concentration of osmotically active particles (osmolality) in each fluid. The osmolality of a fluid is roughly equal to the sum of the concentrations of individual solutes in the fluid. Total body water is regulated by thirst, which controls fluid intake, ADH, and renal function, which in turn controls renal conservation of water. With fluid deprivation,

there is a decrease in extracellular fluid volume and a corresponding increase in plasma osmolality. When the osmole receptors in the hypothalamus detect an increase in osmolality as little as 1% to 2%, both thirst and ADH pituitary release are stimulated. There is a linear increase in plasma ADH with an increase in plasma osmolality, with a maximum ADH response when the plasma osmolality reaches 290 mOsm/kg H_2O. Thirst increases water consumption and ADH increases water reabsorption in the collecting ducts of the kidney, resulting in water conservation and excretion of a concentrated urine. Urine volume is reduced, and urine osmolality is typically greater than 800 mOsm/kg. Conversely, when there is extracellular fluid volume expansion, receptors in the hypothalamus sense the fall in plasma osmolality and cause inhibition of both thirst and the pituitary release of ADH. The absence of ADH renders the collecting duct relatively impermeable to water, and hypotonic urine is produced with relatively high urine volume and osmolality typically less than 100 mOsm/kg. Both thirst and ADH release are also regulated directly by extracellular fluid volume in addition to osmolality. A decrease in extracellular fluid volume is sensed by carotid baroreceptors and stretch receptors in the left atrium of the heart, and signals are sent to the hypothalamus via the vagus nerve to stimulate thirst and ADH release. Extracellular fluid volume can override the regulation from plasma osmolality if there is a disparity between the two. For example, decreased extracellular fluid volume with low osmolality can actually cause stimulation of thirst and ADH release.

The concentrating and diluting function of the kidney is reflected in the solute concentration of the urine, which can be measured by specific gravity or osmolality. Specific gravity is defined as the ratio of the weight of a volume of urine to the weight of an equal volume of water. The specific gravity, like osmolality, is related to the concentration of dissolved solutes.

The specific gravity of a 24-hour urine specimen normally varies between 1.015 and 1.025, while the specific gravity of random urine samples can vary between 1.003 and 1.030, depending on fluid intake and state of hydration. After an overnight fast, a random urine specific gravity over 1.020 is considered to reflect normal renal concentrating ability.

There will be disparity between urine specific gravity and osmolality when significant amounts of compounds are present that contribute more to specific gravity than osmolality. These can include protein, glucose, and radiopaque dyes. The contribution to specific gravity is about 0.003 for each 1 g/dL of protein, and 0.004 for each 1 g/dL of glucose.

Specific gravity can be determined with a “urinometer,” a hydrometer specifically designed for urine samples, but this requires a large sample size and is seldom used.

Specific gravity is often estimated using a urine test strip or “dipstick.” The reagent includes a polymer with repeating carboxylic acid groups. When the polymer-impregnated strip is immersed into urine, the ionic species of the urine interact with the carboxylic acid groups and the pH change on the test strip is detected using an indicator such as bromthymol blue. The color on the test strip correlates with the specific gravity of the urine except in the presence of high concentrations of glucose or other nonionic species. The method also underestimates specific gravity in alkaline urine (pH greater than 6.5).

The specific gravity of urine can also be determined by refractometry. The principle of refractometry is based on the fact that light bends as it passes through a liquid. This angle of refraction is known as the refractive index and is related to the specific gravity of the solution being measured. A refractometer specially calibrated for use with urine is available.

The urine osmolality is a better measure of renal concentrating ability than the urine specific gravity. The sum of the active concentrations of all species dissolved in a solution determines the osmolality.

The colligative properties of a solution are directly related to the number of particles dissolved per unit mass of solvent, and the measurement of osmolality is based on one of these colligative properties. A 1-molal “ideal” solution will lower the freezing point by 1.86°C, lower the vapor pressure by 0.3 mm Hg, raise the osmotic pressure by 22.4 atm, and raise the boiling point by 0.52°C. The colligative properties most commonly used in the laboratory to measure osmolality are depression of the freezing point and depression of the vapor pressure. Real solutions usually have somewhat lesser effects on the freezing point and vapor pressure of the solvent than an ideal solution. The osmolality of a real solution is expressed as

Osmolality (mosm/kg H_2O) = φnC

where φ is the osmotic coefficient, which is 1.0 for an ideal solution and somewhat less than 1 for electrolytes that are incompletely dissociated or for solutes that tend to associate because of electrostatic or other forces, n is the number of particles formed by complete dissociation of the molecular species (mOsm/mmol), and C is the molality of the solute in mmol/kg H_2O.

Clinical Use of Urine Osmolality

The osmolality of urine can vary from 50 to 1,200 mOsm/kg H_2O, depending on the state of hydration. A random urine sample is usually in the range of 300 to 800 mOsm/kg H_2O. With a water load and corresponding decreased plasma osmolality, there is an inhibition of thirst and ADH release, which leads to diuresis with low urine osmolality. With water deprivation and a corresponding high plasma osmolality, thirst and ADH release are stimulated, and renal water conservation occurs, with a resultant low urine volume with high osmolality. Although the measurement of urine osmolality alone is of limited use in the diagnosis of water and electrolyte imbalance, a ratio of the urine osmolality to plasma osmolality can be extremely helpful in the differentiation of water imbalance states. In normal individuals, the urine osmolality to serum osmolality ratio is usually between 1.0 and 2.5. In chronic renal failure, the concentrating ability of the renal tubules is progressively lost. The urine-to-plasma osmolality ratio is less than 1.2 if there is predominantly tubular damage, but is more than 1.2 if there is predominantly a decrease in GFR rate. In acute tubular necrosis, the urine osmolality can be less than or equal to the plasma osmolality. Normal renal concentrating ability should result in a urine osmolality of more than 800 mOsm/kg H_2O after an overnight fast.

The urine-to-plasma osmolality ratio is also helpful in the diagnosis of polyuria. Polyuria can be caused by water diuresis induced

by insufficient ADH, suppression of ADH by excessive water consumption seen in such entities as psychogenic polydipsia, organic brain disease, and iatrogenic water ingestion, or failure of the kidney to respond to ADH, as seen in congenital nephrogenic diabetes insipidus, chronic renal disease, and acquired tubular defects from drugs such as lithium, phenytoin, ethanol, and amphotericin B.

Diabetes insipidus is a partial or total lack of the ability to produce ADH. Diabetes insipidus can be an inherited condition or can result from an insult to the hypothalamus or posterior pituitary gland. Nephrogenic diabetes insipidus is a different condition in which the distal nephron remains relatively impermeable to water reabsorption despite large amounts of ADH. The condition may also be inherited but is more commonly related to nephrotoxic drugs such as lithium, phenytoin, ethanol, and amphotericin B.

Either a lack of ADH, suppression of ADH, or a failure of the kidney to respond to ADH will result in copious volumes of dilute urine owing to the collecting duct being impermeable to water. There may be as much as 10 to 15 L of urine a day, with osmolalities of less than 100 mOsm/kg H_2O. Serious dehydration will occur unless water intake matches the urine output. Water diuresis states can be studied by concentration tests. The principle of a concentration test is that polyuria should be corrected by either fluid deprivation or ADH administration. With the patient fasting, plasma osmolality is determined every 2 hours. All urine excreted is also collected. The patient is adequately prepared when the plasma osmolality reaches 300 mOsm/kg or when the urine has a level plateau of osmolality. Baseline plasma and urine osmolality and plasma ADH levels are then determined. The patient is then given 5 mU of ADH subcutaneously, and the urine osmolality is measured 1 hour later. With diabetes insipidus, the result will be a high baseline plasma osmolality, low urine osmolality, absent plasma ADH level, and rise in urine osmolality after ADH injection. With suppression of ADH caused by psychogenic polydipsia, the results will usually be normal to low plasma osmolality with low urine osmolality, low plasma ADH level, and a rise in urine osmolality after administration of ADH. With nephrogenic diabetes insipidus, the results will be high plasma osmolality, low urine osmolality, high plasma ADH level, and absent rise in urine osmolality after ADH injection (Table 19.1).

TABLE 19.1. Laboratory Differentiation of Polyuria

Laboratory Parameter	Psychogenic Polydipsia	Diabetes Insipidus	Nephrogenic Diabetes Insipidus	Osmotic Diuresis
Posm	N, ↓	↑	↑	N, ↑
Uosm	↓	↓	↓	↑
Plasma ADH	↓	↓	↑	↑
Uosm after ADH	Increase	Increase	Nil	Slight increase

Posm, plasma osmolality; Uosm, urine osmolality.

Other concentration tests such as the Hickey-Hare test utilizing hypertonic saline can be used in the differential diagnosis of polyuria. In the Hickey-Hare test, samples for plasma osmolality and ADH levels are obtained at 20-minute intervals after hypertonic saline is started intravenously. In diabetes insipidus, there are undetectable ADH levels even when the plasma osmolality reaches 300 mOsm/kg H_2O.

The ratio of urine to plasma osmolality (U/P osm) is also helpful in the diagnosis of the syndrome of inappropriate ADH. In this syndrome, there is a pathologic amount of ADH, producing the opposite biochemical effects of diabetes insipidus. The excess ADH causes increased permeability of the collecting ducts to water, leading to water retention and hyponatremia. The urine osmolality will be elevated compared with the plasma osmolality (Uosm/Posm >2). The increased extracellular fluid volume leads to an increased GFR, increased tubular fluid flow, and a resultant sodium washout into the urine. There can be a substantial urinary loss of sodium in this condition, with urine sodium concentrations exceeding 20 mmol/L. The syndrome of inappropriate ADH can be seen in a variety of conditions, most commonly in pathologic lesions of the central nervous system or pulmonary system and in malignant tumors producing ADH, such as small cell carcinoma of the lung.

UREA

Part of "19 - Nitrogen Metabolites and Renal Function"

Urea is the major nitrogen-containing metabolite from the degradation of protein arising from the ammonia produced from the deamination of amino acids. The formation of urea occurs primarily in the liver, with minor amounts of urea formation occurring in the kidney and brain. The concentration of urea in the bloodstream depends on several factors. These are the rate of urea production, the volume of body water in which the urea is distributed, and the rate of urea elimination. The rate of urea production depends on the amount of protein in the diet, protein and blood in the gastrointestinal tract, catabolic states, and liver function. Urea is freely diffusible across cell membranes and is equally distributed in the total body water. Urea excretion depends on the GFR and the subsequent tubular reabsorption of urea. Urea is freely filtered across the glomerular membrane, with approximately 40% of the urea reabsorbed into the extracellular fluid from the tubule. The reabsorptive process is passive and depends on the fraction of filtered water that is reabsorbed.

Plasma urea will rise as the GFR decreases in patients with renal disease. The urea level is insensitive to a minimal fall in GFR and rises sharply only after 40% to 60% of the nephrons have ceased to function. Plasma urea increases progressively with a hyperbolic relationship to the GFR as renal disease progresses,

such that a 50% drop in the GFR results in an approximate doubling of plasma urea (Fig. 19.3).

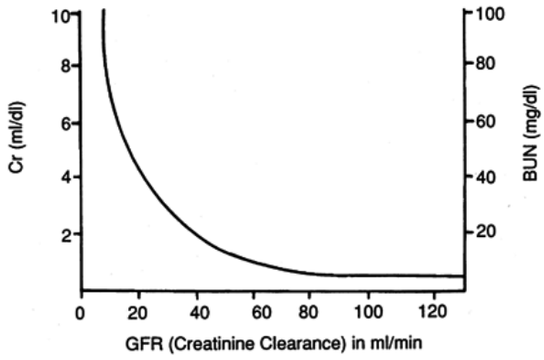


FIGURE 19.3. Relationship of serum creatinine or blood urea nitrogen to the glomerular filtration rate as measured by the creatinine clearance.

Etiologies of elevated urea level can be categorized as prerenal, renal, and postrenal (Table 19.2). Prerenal causes are related to increased production of urea or decreased renal perfusion. Renal causes are related to a loss of functioning nephrons with a corresponding decrease in GFR. Postrenal causes are related to obstruction of the urinary tract with a decreased urine flow.

TABLE 19.2. Etiologies of Elevated Serum Urea

Prerenal
Increased synthesis of urea
Catabolic states: fever, stress, burns
High-protein diet
Gastro intestinal bleeding
Hyperthyroidism
Cushing's syndrome
Hemolysis
Antianabolic drugs/tetracycline
Steroid therapy
Low calorie diet
Malignancy
Sepsis
Decreased perfusion of kidney
Congestive heart failure
Hypotension, shock
Renal vein thrombosis
Dehydration
Cirrhosis, ascites
Intrinsic renal disease
Glomerular disease
Tubular disease—acute tubular necrosis
Interstitial disease
Postrenal
Urinary tract obstruction
Benign prostatic hyperplasia
Prostatic carcinoma
Carcinoma of bladder of bilateral ureters
Retroperitoneal tumor
Calculi
Extravasation of urine into tissues

Urea may be abnormally low in several conditions in which there is a low rate or urea production, hemodilution, or increased rate of urea excretion. Decreased urea synthesis can occur with a low protein intake, administration of androgens or growth hormone, and severe liver dysfunction. Hemodilution can occur with overhydration, psychogenic polydipsia, and diabetes insipidus. Increased rates of elimination can be seen in hemodilution, pregnancy, and postdialysis states.

Analysis of Urea

1. *Urease (EC 3.5.1.5) with glutamate dehydrogenase (EC 1.4.1.3) coupled reaction*. This is the most common method for analysis of urea on automated systems. The highly specific urease enzyme converts urea into ammonia. A coupled reaction utilizing glutamate dehydrogenase with nicotinamide adenine dinucleotide (reduced form) (NADH) and α-ketoglutarate is monitored by measuring the decrease in absorbance at 340 nm as NADH is converted to NAD^+. Interferences may be seen in the assay if other enzymes that can also oxidize NADH are elevated in the specimen.

2. *Conductimetric urease*. There is an increase in conductivity of the specimen after the addition of urease (EC 3.5.1.5). The liberated bicarbonate and ammonium increase the conductivity of the solution. The rate of change of conductivity of the mixture as nonionized urea is transformed to the ionic species and can be monitored with a conductivity cell.

3. *Urease with quinolinium dye*. The urease (EC.3.5.1.5) enzyme liberates ammonia from urea. The ammonia reacts with merocyanine dye, which can be measured at 670 nm. This method is used in multilayer dry film assays and dry reagent strips where the reagents are incorporated into a matrix of hydrophilic polymers coated on top of a transparent base. The ammonia liberated transfuses through the base onto another layer containing the quinolinium indicator dye, which changes color with a change in pH.

4. *O-Phthalaldehyde*. Urea binds to O-phthalaldehyde to form a colored chromogen that can be monitored at 510 nm. The disadvantage of this method is interference by other primary amines.

5. *Diacetyl*. Diacetyl monoxime is hydrolyzed to form diacetyl and hydroxylamine. The diacetyl complexes with urea in an acid solution containing an oxidizing agent to form a yellow diazine compound (Fearon reaction), which can be measured at 550 nm, or fluorometrically at 415 nm. This method is easily automated. The disadvantage of this method is that nonspecific substances may produce chromogens and there is a low specificity compared with the enzymatic assays. The end color fades and usually a stabilizer such as thiosemicarbazide is added to enhance and stabilize the end product. An oxidizing agent may also be added to remove hydroxylamine and ensure a complete reaction.

6. *Berthelot reaction*. Urea is hydrolyzed with urease to carbonic acid and ammonia. Sodium phenate and hypochlorite react with ammonia in the presence of sodium nitroprusside to produce a blue indophenol complex with maximum absorbance at 630 nm. The disadvantages of this reaction are nonspecificity and long reaction times. The advantages are that protein precipitation is not required and turbidity is not a problem.

7. *Definitive Method*. Isotope dilution-mass spectrometry method is available as a definitive method.

The earliest methods for urea in the clinical laboratory used whole blood as the sample, and an analytical method that measured nitrogen in urea. This led to the convention of reporting concentrations in terms of milligrams of nitrogen per deciliter of sample, which is still used except in areas where SI units have been adopted. The misnomer BUN (for blood urea nitrogen) is also still commonly used today, even though plasma or serum urea nitrogen is meant.

The reference interval for plasma (or serum) urea nitrogen is 7 to 22 mg/dL.

CREATININE

Part of "19 - Nitrogen Metabolites and Renal Function"

Creatine is synthesized in the liver from arginine and glycine and is transported to muscle, which contains 98% of the total body stores of creatine. The creatine pool is related primarily to muscle mass, which is affected primarily by gender and age and to a lesser extent by dietary intake of creatine. Approximately 1.5% of the total body creatine is degraded to creatinine each day by nonenzymatic dehydration. Creatinine is not bound to protein but is freely distributed through the total body water and is excreted from the body by glomerular filtration, with an additional 10% to 20% of creatinine secreted through the tubule. Creatinine is not reabsorbed in the tubule.

The serum creatinine level is determined by creatinine production, state of hydration, and creatinine excretion. Increased creatinine production is seen in acute muscle disease such as dermatomyositis, and decreased creatinine production is seen in muscle-wasting diseases. Decreased renal perfusion, as in shock, hypotension, congestive heart failure, or cirrhosis, will decrease the GFR and raise the serum creatinine. Creatinine will also be elevated whenever the GFR is low because of renal disease or postrenal obstruction. Like urea, the etiologies of elevated creatinine level can be categorized as prerenal, renal, or postrenal and are summarized in Table 19.3.

TABLE 19.3. Etiologies of Elevated Serum Creatinine Level

Prerenal
Increased synthesis of creatinine
Muscle hypertrophy
Muscle necrosis
Anabolic steroid use
High-meat diet
Drug—phenacemide
Severe exercise
Decreased renal perfusion
Hypotension, shock
Congestive heart failure
Cirrhosis, ascites
Renal
Glomerular
Tubulointerstitial
Decreased tubular secretion
Drugs—cimetidine, trimethroprim, probeneoid
Postrenal
Urinary tract obstruction
Benign prostatic hyperplasia
Carcinoma of prostate, bladder, or ureters
Calculi
Retroperitoneal tumor

Also like urea, creatinine has a nonlinear relationship to the GFR and does not become significantly elevated until the GFR decreases by 40% to 60% (Fig. 19.3).

Simultaneous determination of the serum creatinine and urea levels and calculating the serum urea nitrogen:creatinine ratio may be of benefit in separating prerenal and postrenal conditions from renal diseases as the cause of an elevated serum creatinine or urea level. The normal serum urea nitrogen:creatinine ratio is in the range of 10:1 to 20:1. An elevated serum urea nitrogen:creatinine ratio is seen in prerenal and postrenal conditions, whereas a normal serum urea nitrogen:creatinine ratio is seen in renal disease. The ratio is elevated in prerenal or postrenal conditions due to either increased tubular reabsorption of urea or increased production of urea. In renal disease, the serum urea nitrogen:creatinine ratio remains in the normal range owing to both being similarly affected by the loss of functioning nephrons.

The amount of creatinine excreted in a 24-hour urine depends primarily on the muscle mass of the individual. A man excretes 14 to 26 mg of creatinine per kilogram per day, and a woman excretes 11 to 20 mg/kg per day in the urine, but the amount is fairly consistent for a given individual. The urine creatinine therefore provides a rough estimate of the completeness of a timed urine collection.

Creatinine Methods

The most common methods for the analysis of creatinine are based either on the Jaffé reaction or on one of several enzymatic reactions.

The Jaffé reaction, first described in 1886, is the reaction between creatinine and picrate in an alkaline medium, forming a red-orange complex that absorbs in the range of 480 to 520 nm. Optimal performance of the reaction depends on the concentrations of the various reagent components, especially the concentration of the alkali, which determines the rate of the reaction and the wavelength of maximal absorbance. Most manual methods are end-point assays with 10 to 15 minutes allowed for the reaction. The Jaffé reaction is not specific for creatinine, as a variety of other substances can also react with the picrate. These are called noncreatinine chromogens, and the most common are protein, acetoacetone, glucose, ascorbic acid, guanidine, acetone, α-ketoacids, urea, and cephalosporin antibiotics. Many modifications have been introduced to minimize these interferences. These include pretreatment with Lloyd's reagent (aluminum silicate) to remove most of the noncreatinine chromogens, dialysis or precipitation to remove protein, the addition of buffers to complex glucose and ascorbic acid, and the introduction of a kinetic approach rather than an end-point determination. The kinetic analysis is useful because there is a time interval (usually 25 to 60 seconds) during which the reaction rate is predominantly caused by the formation of a creatinine picrate complex. There is still some positive interference owing to acetoacetone, and bilirubin causes a negative interference. The kinetic rate method allows the Jaffé reaction to be more specific, faster, and easily automated.

Enzymatic methods. There are several enzymatic methods for the analysis of creatinine, which are usually based on one of two

enzymes. Creatinine deaminase (creatinine iminohydrolase) (EC 3.5.4.21) catalyzes creatinine degradation to n-methylhydantoin and ammonium ion. This is then followed by detection of the ammonium ion. Detection of the ammonium ion by reaction with bromphenol blue has been used in dry multilayer film methods with the ammonia diffusing through a semipermeable layer to react with the bromphenol blue. Creatininase (creatinine amidohydrolase) (EC 3.5.2.10) catalyzes hydrolysis of creatinine to creatine. The creatine can be coupled to other enzymatic reactions that, for example, usually culminate in measuring a change in NADH absorbance at 340 nm. The enzymatic methods are more specific than the Jaffé methods and give 10% to 20% lower serum creatinine values than a corresponding Jaffe method.

3,5-Dinitrobenzoic acid. This method involves creatinine reacting with the 3,5-dinitrobenzoic acid at an alkaline pH to give a purple color. The method correlates with the alkaline picrate methods.

High-pressure liquid chromatography. Several high-pressure liquid chromatography methods have been described involving the separation of creatinine from interfering substances using an isocratic reverse-phase method. Although this method is not practical for routine clinical use, it is useful for validating other methods because of its high specificity. A description of a definitive method using isotope dilution gas chromatography/mass spectrometry has also been published. This method is also extremely specific, and the creatinine values of the definitive method are approximately 20% lower than values from the alkaline picrate method when serum samples are used.

Analysis of creatinine in urine requires a dilution of 1:100 because the urine creatinine concentration is 100-fold that of serum.

CLEARANCE TESTS

Part of "19 - Nitrogen Metabolites and Renal Function"

The renal clearance of a substance can be thought of as the volume of plasma that can be completely cleared of that substance each minute by the kidney. The renal clearance of any substance can be calculated with the following relationship:

$$\text{Clearance (mL/min)} = \frac{\text{Urine concentration}}{\text{Plasma concentration}} \times \text{Timed urine volume (mL/min)}$$

The clearance is proportional to the renal parenchymal mass, which in turn is roughly proportional to the total body surface area. To compare clearance studies of individuals with varying body sizes, the clearance can be expressed in terms of a standard body surface area of 1.73 m^2. This is done by dividing 1.73 by the patient's body surface area and multiplying the observed clearance by this corrected surface area. Nomograms can be used to calculate body surface area from the patient's body weight and height.

The clearance of a substance that is freely filtered through the glomerulus and is not subsequently reabsorbed or secreted by the tubule is a measure of GFR. Inulin is such a substance. Inulin is an inert polysaccharide with a molecular weight of approximately 5,000 that is not bound to plasma proteins. The measurement of inulin clearance necessitates a continuous intravenous infusion of inulin to maintain a plasma inulin steady state. Intravenous administration of radiolabeled substances such as ^{51}Cr (creatinine)-EDTA (ethylenediaminetetraacetic acid), ^{125}I (iodine-125)-iothalamate, and ^{99m}Tc (technetium-99m)-DTPA (diethylenetriamine pentaacetic acid) can also be used to obtain clearances that directly measure the GFR.

Because determination of the clearance of these compounds is time-consuming and impractical to perform routinely, the endogenous creatinine clearance is commonly used for the estimation of GFR. Creatinine is produced in muscle at a relatively constant rate, is not bound to plasma protein, and is freely filtered through the glomerulus. Because a small amount (less than 10%) of creatinine is also secreted by the tubules, the creatinine clearance is slightly higher than a simultaneously measured inulin clearance, and thus slightly overestimates the GFR. If the creatinine analyses are performed with the Jaffé reaction, the presence of noncreatinine chromogens in the plasma but not in the urine will minimize this overestimation by partially cancelling out the secreted creatinine in the urine. More specific enzymatic creatinine methods will produce a creatinine clearance with a more pronounced overestimation of the GFR. With increasing renal failure, the creatinine clearance overestimates the GFR even more owing to increased tubular creatinine secretion.

The creatinine clearance averages 125 mL/min in adult males and 115 mL/min in adult females. The GFR is relatively stable from the first year of life through the fourth decade. After the fourth decade, the GFR decreases approximately 1 mL/min per year, and thus the normal range for creatinine clearance in elderly adults is lower than in younger individuals.

There are several other important factors to consider in interpreting the creatinine clearance. The coefficient of variation of creatinine clearance varies from 10% to 20% because of the imprecision of the creatinine measurement, biological variation of creatinine synthesis, variations in the urinary excretion of creatinine, and variations in the timing and collection of urine specimens. The creatinine clearance can also be estimated from the serum creatinine, taking into account the patient's age, weight, gender, and body size and using the Cockcroff and Gault formulas.

$$\text{Ccr} = \frac{(140 - \text{age}) \times \text{wt}}{\text{Pcr} \times 72} \text{ (males)} \quad \text{or} \quad \text{Ccr} = \frac{(140 - \text{age}) \times \text{wt}\ (0.85)}{\text{Pcr} \times 72} \text{ (females)}$$

where age is in years, weight in kilograms, and plasma creatinine in milligrams per deciliter. These formulas may not be accurate in estimating the GFR in patients with edema, obesity, cachexia, or muscle-wasting conditions. This estimation nonetheless may be helpful in the calculation of the dosage of many potentially toxic drugs.

The relationship between plasma creatinine and urea and the GFR is nonlinear (Fig. 19.3). It is evident that with normal to moderately low values of GFR, a large change in the GFR results in only a small rise in the plasma creatinine or urea level, whereas with lower values of GFR, the increase in urea or plasma creatinine level becomes increasingly greater as the GFR falls. Thus,

urea and creatinine are relatively insensitive tests of early renal disease; the GFR must fall by approximately 50% before these tests become significantly abnormal.

An abnormal creatinine clearance is not a specific indication of renal disease because of other factors that influence the GFR. With decreased renal perfusion (e.g., congestive heart failure, cirrhosis, dehydration, hemorrhage) or decreased tubular fluid flow seen in postrenal obstruction (e.g., benign prostatic hyperplasia, urinary bladder or prostatic carcinoma, calculi), the GFR will fall, with a corresponding decrease in the creatinine clearance. The GFR also increases in high renal perfusion states such as pregnancy and inappropriate ADH, leading to a corresponding increase in the creatinine clearance.

Other types of renal clearances can be used to determine parameters other than the GFR. For example, renal plasma flow can be estimated by *p*-aminohippurate clearance, tubular function by sodium clearance, and osmolar excretion by osmotic clearance.

CREATINE

Part of "19 - Nitrogen Metabolites and Renal Function"

Creatine can be measured directly by (a) the condensation of creatine with diacetyl-1-naphthol to form a pink product, (b) the use of creatine kinase enzyme with a fluorometric assay, and (c) the use of guanidine with ninhydrin in an alkaline solution to produce a fluorescent product. Creatine is commonly measured indirectly by converting creatine to creatinine through heating the sample with acid in the presence of a heavy metal. Creatinine is measured before and after the addition of the reagent, and creatine is calculated from the difference in the concentrations. Plasma creatine is increased with muscle necrosis and most muscle diseases, but the utilization of plasma creatine measurements has declined sharply in favor of the analysis of muscle enzymes such as creatine phosphokinase.

URIC ACID

Part of "19 - Nitrogen Metabolites and Renal Function"

Uric acid is the major end product of purine metabolism, specifically the catabolism of the purine nucleosides adenosine and guanosine. The metabolism of these nucleosides is as follows:

$$\begin{array}{r} \text{Adenosine} \rightarrow \text{Inosine} \rightarrow \text{Hypoxanthine} \\ \downarrow \text{Xanthine oxidase} \\ \text{Guanosine} \rightarrow \text{Guanine} \rightarrow \text{Xanthine} \\ \downarrow \text{Xanthine oxidase} \\ \text{Uric Acid} \end{array}$$

Uric acid production occurs primarily in the liver and intestinal mucosa because of the high xanthine oxidase enzyme activity in these two tissues. Daily synthesis of uric acid is approximately 700 mg, with dietary nucleoprotein contributing 300 mg to the daily total uric acid production. The body pool of uric acid is approximately 1.2 g, and approximately 60% of this pool turns over every day. Since the pKa of uric acid is 5.75, the majority of uric acid in the body exists as the urate ion.

The only effective route for excretion of uric acid from the body is through the kidney. Uric acid excretion is complex. Of filtered uric acid, 98% to 100% is actively reabsorbed in the proximal portion of the proximal convoluted tubule, is secreted in the distal portion of the proximal convoluted tubule, and finally is actively reabsorbed in the distal convoluted tubule. Only 6% to 12% of the original filtered uric acid is excreted in the urine, which amounts to 400 to 800 mg excreted in the urine daily.

Plasma uric acid levels show day-to-day and seasonal variations in the same patient. The uric acid level is higher in the summer than in the winter in most individuals. Plasma levels are lower during childhood, reaching adult levels by age 18. The reference range depends on the method that is used to analyze uric acid and is in the range of 4.0 to 7.5 mg/dL for adult males and 2.7 to 6.0 mg/dL for adult females when a uricase method is used. Uric acid levels tend to increase gradually with age, with an average 10% increase from age 20 to 60 years. There is a significant rise in uric acid in women after menopause. Several factors influence the uric acid level, most of which are poorly understood. Increases in exercise, stress, and weight, along with hypertension, diabetes mellitus, and type A personalities are associated with higher values of uric acid. Drugs can affect uric acid levels by altering the rate of synthesis and/or excretion. Drugs that have been reported to elevate uric acid are ethanol, acetaminophen, androgens, low-dose aspirin, furosemide, and mercurial diuretics; the following drugs have been reported to lower uric acid: high-dose aspirin, allopurinol, chlorpromazine, cortical steroids, warfarin (Coumadin), estrogens, and phenothiazines.

Hyperuricemia is seen in several pathologic conditions. These conditions can be owing to either an increased production of uric acid or decreased renal excretion of uric acid. Increased production of uric acid is seen with increased nucleic acid turnover, which leads to increased catabolism of purines. This condition can be seen if there is a rapid proliferation of cells, such as in lymphoproliferative disorders or malignancy, in hemolytic anemia, with a rich purine diet, or with increased *de novo* synthesis, such as in primary gout. A decreased urinary excretion of uric acid can be seen in acute or chronic renal disease of any type (Table 19.4).

TABLE 19.4. Etiology of Hyperuricemia

Increased uric acid synthesis
Gout
Purine-rich diet (organ meats)
Tissue catabolism (necrosis, radiation)
Myeloproliferative disorders
Lymphoproliferative disorders
Chemotherapy of malignancy
Lesch–Nyhan syndrome
Decreased renal excretion of uric acid
Renal failure
Drugs
Diuretics
Small-dose aspirin (< 2 g/day)
Metabolic acidosis
Toxemia or pregnancy
Miscellaneous etiologies
Intoxications (ethanol, methanol)
Endocrine (acromegaly, hyperparathyroidism)
Psychosocial, hypertension, type A personality
Psoriasis

Gout is a disorder of purine metabolism or renal excretion characterized by an increase in the total body uric acid and the precipitation of monosodium urate as deposits (tophi) in and around joints, bursae, periarticular cartilage, bone, and subcutaneous tissue. There can be recurrent attacks of arthritis and nephropathy with nephrolithiasis. The deposits of urate crystals are responsible for the clinical signs and symptoms of the disease. There is thought to be a genetic disposition, although fewer than one third of patients have a positive family history. Men make up 90% of all cases, and gout is uncommon in women before menopause. The peak age of onset of gout is in the fifth decade. The plasma level of uric acid correlates roughly with the clinical severity of gout; however, there is extreme variability from individual to individual with the same level of uric acid producing different degrees of clinical symptoms.

Hypouricemia, defined as a uric acid level below 2.0 mg/dL, can be seen in severe liver disease with a decreased synthesis of purines, a hereditary deficiency of xanthine oxidase, administration of allopurinol (a xanthine oxidase inhibitor) or uricosuric drugs, and a defect in renal tubular reabsorption of uric acid seen in Fanconi's syndrome.

Analytical methods for uric acid fall into two categories: phosphotungstic acid and uricase methods.

Phosphotungstic acid oxidizes uric acid to allantoin. The reduction product of phosphotungstate in an alkaline medium is blue and can be quantitated by measuring absorbance at approximately 700 nm. Phosphotungstic acid methods differ in the reagents used to intensify the blue color. The reaction is nonspecific, and positive interferences can be seen for glucose, ascorbic acid, glutathione, cystine, and drugs such as aspirin, acetaminophen, caffeine, and theophylline. Protein must be removed to avoid turbidity and quenching of the color development. Phosphotungstic acid methods give plasma uric acid values 0.2 to 0.4 mg/dL higher than uricase methods.

Uricase methods are based on the oxidation of uric acid to allantoin and hydrogen peroxide by the enzyme uricase (urate oxidase, EC 1.7.3.3) and have the advantage of greater specificity and accuracy. There are many adaptations of this method. Uric acid is maximally absorbed between 290 and 295 nm at an alkaline pH. This absorbance disappears as uric acid is oxidized with uricase. The concentration of uric acid can therefore be calculated from the difference in absorbance before and after the action of uricase. In another approach, coupled uricase methods utilize the hydrogen peroxide produced as a substrate for a chemical indicator reaction catalyzed by peroxidase or catalase. A coupled uricase method is utilized in the dry multilayer film technology methods. The reference method for uric acid analysis is based on the uricase reaction, and there is also a definitive method using isotope dilution/mass spectrometry.

AMMONIA

Part of "19 - Nitrogen Metabolites and Renal Function"

Ammonia is a product of amino acid metabolism, primarily in the liver, where oxidative deamination of L-glutamate occurs to form α-ketoglutarate and ammonia. The major source of ammonia, however, is absorption from the gastrointestinal tract where bacteria proteases, ureases, and amine oxidases act on protein contents in the colon, and simultaneous hydrolysis of glutamine occurs in the small and large intestines to form a large amount of ammonia. Ammonia is absorbed into the portal blood system and is ultimately metabolized to urea in the liver through the urea cycle. A small amount of ammonia is converted into glutamine and stored in the liver, brain, and skeletal muscle. The ammonia concentration in the systemic circulation is low, in the range of 10 to 60 μgN/dL. Elevated ammonia levels can be seen in congenital deficiencies of any of the enzymes of the urea cycle, in Reye's syndrome, and in severe liver disease. In severe hepatic failure, hepatic encephalopathy and coma are associated with high blood ammonia and glutamine, but the severity of the clinical signs does not correlate well with the blood ammonia level.

The most common method for the analysis of ammonia uses the enzyme glutamate dehydrogenase (GLDH) (EC 1.4.1.3) in the following reaction:

$$NH_4^+ + \alpha\text{-ketoglutarate} + NADPH \xrightarrow{GLDH}$$

$$\text{L-glutamate} + NADP^+ + H_2O$$

The ammonia level is determined by measuring the decrease in absorbance at 340 nm as NADPH is converted to $NADP^+$.

Two-stage assays are also available. In the first stage, the plasma is mixed with a cation exchange resin that binds ammonium ion. The ammonium is eluted off the resin and quantitated with Nessler's reagent or the Berthelot reaction (phenolhypochlorite).

Ammonia analysis in the clinical laboratory is particularly subject to preanalytical sources of error because of the very low concentration of ammonia that is normally present. The ammonia content of blood rises rapidly after collection because of the *in vitro* metabolism of nitrogenous compounds and the degradation of glutamine. Hemolysis can lead to an increased ammonia level because of a higher concentration of ammonia in red cells than the plasma. These can be minimized by keeping the sample on ice and beginning the analysis within 30 minutes. Ammonia is stable for at least 2 days if plasma is frozen at -20°C. Cigarette smoking can cause a marked contamination of ammonia in the sample.

RENAL DISORDERS

Part of "19 - Nitrogen Metabolites and Renal Function"

Acute Renal Failure

An abrupt decline in renal function is a common clinical problem. This is manifested clinically by the onset of oliguria (less than 500 mL of urine per day) and an acute rise in serum creatinine and urea. Approximately 70% to 75% of cases have a prerenal etiology associated with decreased renal perfusion, and 25% of cases are owing to intrinsic renal disease from ischemia or toxins. Decreased renal perfusion can be the result of volume depletion from hemorrhage or gastrointestinal tract fluid loss, hypotension associated with septic shock, congestive heart failure, and cirrhosis. The most common renal disease associated with acute renal failure is acute tubular necrosis (ATN). ATN is manifested by sloughing of renal tubular cells and occlusion of the tubules by cellular debris. ATN is associated with hypoxic injury of the tubule by renal ischemia or tubular injury caused by toxins

or drugs, such as aminoglycosides, radiologic contrast media, heme pigments, amphotericin B, and cisplatin. The oliguric phase of ATN lasts 7 to 21 days and is followed by a diuretic phase. Azotemia decreases 1 to 3 weeks after the onset of diuresis.

The mortality rate is 40% to 60%. It is important to differentiate between decreased renal perfusion and ATN as the cause of acute renal failure because the treatment for the two conditions is quite different. In general, the treatment for decreased renal perfusion is fluid administration to expand blood volume, whereas ATN requires careful management of fluid to control water and electrolyte imbalance.

In acute renal failure caused by decreased renal perfusion, serum urea tends to be elevated proportionately more than creatinine because of increased reabsorption of urea in the renal tubule. The urea-to-creatinine ratio is generally greater than 20. The decreased plasma volume stimulates ADH and aldosterone, resulting in production of smaller urine volume. The urine is relatively concentrated with a high osmolality but with a low sodium concentration.

Renal tubular damage seen in ATN results in marked tubular dysfunction with a loss of concentrating ability and inability to conserve sodium. The urine osmolality is similar to plasma osmolality, and the urine sodium is more than 40 mEq/L (refer to Table 19.5 for differentiation between ATN and decreased renal perfusion).

TABLE 19.5. Laboratory Differentiation of Acute Tubular Necrosis and Renal Hypoperfusion

Value	Hypoperfusion	ATN
UN/cr	> 20	< 15
Uosm	> 500	< 350
U/Posm	> 1.5:1	1:1
U/Pcr	> 40	< 15
UNa	< 20	> 30–40
SG	> 1.015	< 1.010
FE_{Na}	< 1	> 2
CH_2O	-20 to -100	-20 to +#

ATN, acute tubular necrosis; UN/cr, urea nitrogen:creatinine ratio; FE_{Na}, fractional excretion of sodium; SG, specific gravity; UNa, urinary sodium (mEq/liter); Uosm, urinary osmolality; U/Pc, ratio of urinary to plasma creatinine; U/Posm, ratio of urinary to plasma osmolality; CH_2O, free water clearance.

The fractional excretion of sodium (FE_{Na}) may also be of help in differentiating between ATN and decreased renal perfusion. The FE_{Na} is an index of the ability of the kidney to conserve sodium and represents the percentage of filtered sodium that is excreted in the urine. The FE_{Na} can be calculated by simultaneously measuring urinary and plasma sodium (U_{Na}, P_{Na}) and creatinine (U_{cr}, P_{cr}) and using the following formula:

$$FE_{Na}\,(\%) = \frac{U_{Na} \times P_{cr}}{U_{cr} \times P_{Na}} \times 100$$

In decreased renal perfusion, the FE_{Na} will be less than 1%, whereas in ATN, it will generally be greater than 2%.

Free water clearance (CH_2O) determination also is a valuable test to differentiate between prerenal azotemia and ATN. Water excretion by the kidney can be divided into that needed to excrete the solutes, which is called the osmolar clearance (Cosm). Free water clearance is the volume of solute free water that has been added to the solute water (hypotonic urine) or that subtracted from the solute water (hypertonic urine). Free water clearance can be calculated by the following formula:

$$\text{Cosm} = \frac{\text{Uosm} \times \text{Uvol}}{\text{Posm}}$$

$$\text{Uvol} = \text{Cosm} + CH_2O$$

$$\text{thus } CH_2O = \text{Uvol} - \text{Cosm}$$

In prerenal azotemia with a concentrated urine, the free water clearance will be markedly negative (-20 to -100), whereas with ATN, the free water clearance will usually be a positive number because of a lack of tubular concentration.

Chronic Renal Disease

Patients with early renal disease are usually asymptomatic. However, there may be complaints relating to the urinary tract, such as urinary frequency, hematuria, and pain, or there may be systemic complaints, such as edema, hypertension, lassitude, and weakness. The urinalysis is one of the most important tests to detect early renal disease. The kinds of abnormalities seen in the urinalysis are related to the type of renal disease present.

Renal disease can be divided into glomerular, tubular, and interstitial diseases. Glomerular diseases are usually caused by immunologic factors and are defined by specific pathologic changes seen on microscopy. Usually the first manifestation of glomerular disease is altered permeability of the glomerulus, causing proteinuria or hematuria. Additional glomerular damage leads to a decreased GFR, seen as a progressive decrease in the creatinine clearance, and finally an elevation in plasma urea and creatinine. Because of the hyperbolic relationship between plasma urea or creatinine and the GFR, as mentioned earlier, these measurements are poor indicators of early renal disease.

Tubular and interstitial damage can be secondary to multiple myeloma, ATN, nephrotoxic drugs, and pyelonephritis. Tubular dysfunction can be detected by abnormal tests of urinary concentration such as changes in specific gravity, urine and plasma osmolality ratios, inability to conserve sodium, defects in acidification of the urine, increased fractional excretion of sodium, and positive free water clearance.

Progressive and irreversible deterioration of renal function will lead to the kidney's progressive inability to perform its excretory, secretory, and regulatory functions. When more than 90% of the nephrons are nonfunctioning, signs and symptoms known as the uremic syndrome can develop. The uremic syndrome consists of anorexia, nausea and vomiting, lassitude, anemia, bleeding tendency, altered endocrine function, neuromuscular symptoms, and hypertension. Characteristic laboratory findings include anemia, abnormal urinalysis, elevated urea and creatinine, low creatinine clearance, elevated phosphorus, hyperkalemia, hypocalcemia, hyperuricemia, hypermagnesemia, elevated alkaline phosphatase, metabolic acidosis, and a loss of concentrating ability. Diabetes mellitus and essential hypertension are contributing factors in a sizable portion of all chronic renal disease patients. When the GFR decreases to less than 25 mL/min, most patients require dialysis or renal transplantation to prevent the onset of the uremic syndrome.

Renal Tubular Acidosis

Renal tubular acidosis is a term applied to a group of diverse disorders in which the kidneys are unable to acidify urine normally. The common feature is a mild hyperchloremic metabolic acidosis caused by abnormal renal acid-base handling. These disorders show little alteration in the GFR rate. There are two major types of renal tubular acidosis involving either the proximal or the distal convoluted tubules and a third subtype of distal disorder.

Type I renal tubular acidosis was the first renal tubular acidosis described. It occurs more often in female children than in adults. Type I renal tubular acidosis usually presents with growth retardation, muscle weakness, and renal abnormalities. It is caused by a reduced ability of the distal convoluted tubule to acidify urine. Patients with type I distal renal tubular acidosis cannot produce urine with a pH less than 6.0. The exact mechanism is uncertain.

The presence of a urine pH less than 5.2 excludes the diagnosis of distal renal tubular acidosis. There may be an associated abnormal loss of potassium, phosphorus, and calcium in the urine, with the formation of renal calculi. Type I renal tubular acidosis may be familial, autosomal-dominant, or secondary to toxic or metabolic etiologies such as hyperparathyroidism, hypergammaglobulinemia, autoimmune disorders, vitamin D intoxication, use of amphotericin B, and hydronephrosis (Table 19.6). Diagnosis can be made with an ammonium chloride (100 mg/kg) challenge test, in which the urine pH (greater than 5.5) remains inappropriately high in relation to the degree of induced systemic acidosis. A bicarbonate loading test will usually be normal (less than 10%) in this disorder.

TABLE 19.6. Causes of Distal Renal Tubular Acidosis (Type I)

Hereditary
Diseases associated with gammopathy
Amyloidosis
Systemic lupus erythematosus
Sarcoidosis
Macroglobulinemia
Tubulointerstitial disease
Analgesic nephropathy
Sickle cell disease
Chronic pyelonephritis
Obstruction of the urinary tract
Nephrocalcinosis
Hyperparathyroidism
Vitamin D intoxication
Drugs
Amphotericin B
Lithium
Analgesics

TABLE 19.7. Causes of Proximal Renal Tubular Acidosis (Type II)

Idiopathic or hereditary
Vitamin D deficiency
Hyperparathyroidism
Fanconi's syndrome
Chronic renal disease
Amyloidosis
Systemic lupus erythematosus
Drugs
Tetracycline
Carbonic anhydrase inhibitors
Streptozotocin
Gentamicin
Heavy metal intoxication
Wilson's disease
Multiple myeloma
Cystinosis
Galactosemia

Type II renal tubular acidosis is a proximal renal tubular acidosis that usually presents in male children with symptoms of growth retardation. Type II renal tubular acidosis is caused by an inability of proximal renal tubular cells to conserve bicarbonate. In the renal tubule, the majority of renal hydrogen ion secretion occurs in the proximal convoluted tubule, which is responsible for reabsorbing 85% of the filtered bicarbonate. In this disorder, the proximal reabsorption of bicarbonate is impaired, and thus bicarbonate is lost in the urine. Distal hydrogen ion secretion is normal in these patients. In some patients with type II renal tubular acidosis, the net acid secretion may be normal if the serum bicarbonate is low. When the serum bicarbonate falls, a level is reached at which the proximal tubular cells are able to recover all of the filtered bicarbonate. The distal tubule then can acidify the urine in a normal manner. A bicarbonate loading test will reveal the fractional bicarbonate excretion will be more than 15%. Other significant laboratory findings include hypokalemia and increased urinary excretion of calcium, sodium, and potassium. Type II proximal renal tubular acidosis may be an isolated hereditary defect but is more commonly associated with other proximal defects, such as aminoaciduria, glycosuria, and phosphaturia (Fanconi's syndrome). Type II renal tubular acidosis can be secondary to multiple myeloma, cystinosis, galactosemia, heavy-metal intoxication, carbonic anhydrase inhibitors, tetracycline use, hyperparathyroidism, and amyloid infiltration (Table 19.7). The treatment is large dosages of bicarbonate.

Type IV renal tubular acidosis is a type of distal renal tubular acidosis associated with a hyporenin–hypoaldosterone state. Because of a decreased response or absence of aldosterone, there are a failure of potassium and hydrogen ion secretion in the distal tubule and a failure of the distal convoluted tubular cells to produce ammonia will result in metabolic acidosis with hyperkalemia. This type of renal tubular acidosis may also be seen in generalized renal disease.

Laboratory differentiation of the different types of renal tubular acidosis is tabulated in Table 19.8.

TABLE 19.8. Differentiation of Type I, II, and IV Renal Tubular Acidosis

Clinical or Laboratory Features	Type I	Type II	Type IV
Urinary pH with			
HCO_3> 20	> 5.5	> 5.5	> 5.5
HCO_3 < 15	> 5.5	< 5.5	> 5.5
Potassium	↓	↓	↑
Calculi	Yes	No	No
NH_4Cl loading	> 5.5	< 5.5	> 5.5
Fanconi's	No	Yes	No
% HCO_3 excretion	< 10	> 15	< 10
Bone disease	Yes	No	No

SUGGESTED READINGS

Renal Physiology

Harrington AR, Zimmerman SW. *Renal pathophysiology*. New York: John Wiley & Sons, 1982.

Kassirer JP. Clinical evaluation of kidney function—glomerular function. *N Engl J Med* 1971;285:365–385.

Kassirer JP. Clinical evaluation of kidney function—tubular function. *N Engl J Med* 1971;285:499–501.

Pitts RF. *Physiology of the kidney and body fluids*, 3rd ed. Chicago: Year Book Medical Publishers, 1974.

Preuss HG. Basics of renal anatomy and physiology. *Clin Lab Med* 1993;13:1–11.

Renkin EM, Robinson RR. Glomerular filtration. *N Engl J Med* 1974; 290:785–792.

Sands JM, Kokko JP. Current concepts of the countercurrent multiplication system. *Kidney Int* 1996;57:893–899.

Seldin DW. *The kidney: physiology and pathophysiology*, 2nd ed. New York, Raven Press, 1992.

Renal Tubular Acidosis

Battle D. Renal tubular acidosis. *Med Clin North Am* 1983;67:859–878.

Lash JP, Arruda JA. Laboratory evaluation of renal tubular acidosis. *Clin Lab Med* 1993;13:117–129.

Maher ER, Scoble JE. Renal tubular acidosis. *Br J Hosp Med* 1989; 42:116–119.

Roth KS. Diagnosis of renal tubular transport disorders. *Clin Pediatr* 1988; 27:463–470.

Rothstein M, Obialo C, Hruska KA. Renal tubular acidosis. *Endocrinol Metab Clin North Am* 1990;19:869–887.

Water Metabolism

Carroll HJ, Oh MS. *Water, electrolyte and acid-base metabolism: diagnosis and management*. Philadelphia: Lippincott, 1978.

Kleeman CR, Fichman MP. The clinical physiology of water metabolism. *N Engl J Med* 1967; 277:1300–1306.

Robertson GL. Differential diagnosis of polyuria. *Annu Rev Med* 1988; 39:425–442.

Valtin H. *Mechanisms preserving fluid and solute balance in health*, 2nd ed. Boston: Little, Brown, 1983.

Zerbe RL, Robertson GL. A comparison of plasma vasopressin measurements with a standard indirect test in the differential diagnosis of polyuria. *N Engl J Med* 1981; 305:1539–1546.

Specific Gravity

Burkhardt AE, Johnston KG, Waszak CE, et al. A reagent strip for measuring the specific gravity of urine. *Clin Chem* 1982;28:2068–2072.

Rubini HE, Wolfe AV. Refractometric determination of total solids and water of serum and urine. *J Biol Chem* 1957;225:869–874.

Wolf AV, Fuller JB, Goldman EJ, et al. New refractiometric methods for the determination of total proteins in serum and urine. *Clin Chem* 1962;8:158–165.

Osmolality

Dorwart WV, Chalmers L. Comparison of methods for calculating serum osmolality from chemical concentrations and the prognostic value of such calculations. *Clin Chem* 1975;21:190–194.

Gennari FJ. Serum osmolality. *N Engl J Med* 1984;310:102–105.

Jacobson MH, Levy SE, Kaufman RM, et al. Urine osmolality: a definitive test of renal function. *Arch Intern Med* 1962;110:83–86.

Weisberg HF. Osmolality. *Lab Med* 1981; 12:81-85.

Urea

Baum N, Dichoso CC, Carlton CE. Blood urea nitrogen and serum creatinine—physiology and interpretations. *Urology* 1975;5:583–588.

Jurado R, Mattix H. The decreased serum urea nitrogen-creatinine ratio. *Arch Intern Med* 1998;158:2509–2511.

Kessler A, Siekmann L. Measurement of urea in human serum by isotope dilution mass spectrometry, a reference procedure. *Clin Chem* 1999;45:1523–1529.

Lum G, Leal-Khouri S. Significance of low serum urea nitrogen concentrations. *Clin Chem* 1989;35:639–640.

Lyman JL. Blood urea nitrogen and creatinine. *Emerg Med Clin North Am* 1986;4:223–233.

Passey RB, Gillum RL, Fuller JB, et al. Evaluation of three methods for the measurement of urea nitrogen in serum as used on six instruments. *Am J Clin Pathol* 1980;73:362–368.

Spayd RW, Bruschi B, Burdich B, et al. Multilayer film elements for clinical analysis: applications to representative chemical determinations. *Clin Chem* 1978;24:1343–1350.

Ward PC. Renal dysfunction: urea and creatinine. *Postgrad Med* 1981;69:93–104.

Creatinine

Bjorkhem I, Blomstrand R, Ohman G. Mass fragmentography of creatinine proposed as a reference method. *Clin Chem* 1977;23:2114–2121.

Bowers LD, Wong ET. Kinetic serum creatinine assays. II. A critical evaluation and review. *Clin Chem* 1980;25:555–561.

Fossati P, Prencipe L, Berti G. Enzymic creatinine assay: a new colorimetric method based on hydrogen peroxide measurement. *Clin Chem* 1983;29:1494–1496.

Levey AS, Perrone RD, Madias NE. Serum creatinine and renal function. *Annu Rev Med* 1988;39:465–490.

Narayanan S, Appleton HD. Creatinine: a review. *Clin Chem* 1980;26:1119–1126.

Perakis N, Wolff CM. Kinetic approach for the enzymic determination of creatinine. *Clin Chem* 1984;30:1792–1796.

Perrone RD, Madias NE, Levey AS. Serum creatinine as an index of renal function: new insights into old concepts. *Clin Chem* 1992;8:1933–1953.

Rosano TG, Ambrose RT, Wu A, et al. Candidate reference method for determining creatinine in serum: method development and interlaboratory validation. *Clin Chem* 1990;36:1951–1955.

Spencer K. Analytical reviews in clinical biochemistry: the estimation of creatinine. *Ann Clin Biochem* 1986;23:1–25.

Weber JA, Van Zanten AP. Interferences in current methods for measurement of creatinine. *Clin Chem* 1991;37:695–700.

Welch MJ, Cohen A, Hertz HS, et al. Determination of serum creatinine by isotope dilution mass spectrometry as a candidate definitive method. *Anal Chem* 1986;58:1681–1685.

Creatinine Clearance

Apple FS, Benson P, Abraham PA, et al. Assessment of renal function by insulin clearance: comparison with creatinine clearance as determined by enzymatic methods. *Clin Chem* 1989;35:312–314.

Bauer JH, Brooks CS, Burch RN. Clinical appraisal of creatinine clearance as a measurement of glomerular filtration rate. *Am J Kidney Dis* 1982;2:337–346.

Bennett WM, Porter GA. Endogenous creatinine clearance as a clinical measure of glomerular filtration rate. *BMJ* 1971;4:84–86.

Camara AA, Arnk D, Reimer A, et al. The twenty-four hour endogenous creatinine clearance as a clinical measure of the functional state of the kidneys. *J Lab Clin Med* 1951;37:743–763.

Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976;16:31–41.

Herget-Rosenthal S, Kribben A, Pietruck F, et al. Two by two hour creatinine clearance—repeatable and valid. *Clin Nephrol* 1999;51:348–354.

Lemann J, Bidani AK, Bain RP. Use of the serum creatinine to estimate glomerular filtration rate in health and early diabetic nephropathy. *Am J Kidney Dis* 1990;16:236–240.

Rhodes PJ, Rhodes RS. Evaluation of eight methods for estimating creatinine clearance in man. *Clin Pharm* 1987;6:399–406.

Sokol LJ, Russell RM, Sadouski JA, et al. Establishment of creatinine clearance: reference values in older women. *Clin Chem* 1994;40:2276–2281.

Van Lente F, Suit P. Assessment of renal function by serum creatinine and creatinine clearance: glomerular filtration estimation by four procedures. *Clin Chem* 1989;35:2326–2330.

Uric Acid

Bendersky G. Etiology of hyperuricemia. *Ann Clin Lab Sci* 1975;5:456–467.

Carroll JJ, Coburn H, Douglass R, et al. A simplified alkaline phosphotungstate assay for uric acid in serum. *Clin Chem* 1971;17:158–160.

Duncan PH, Gochman N, Cooper T, et al. A candidate reference method for uric acid in serum: 1. Optimization and evaluation. *Clin Chem* 1982;28:284–290.

Elin RJ, Johnson E, Chester R. Four methods for determining uric acid compared with a candidate reference method. *Clin Chem* 1982;28:2098–2100.

Ellerbe P, Cohen A, Welch MJ, et al. Determination of serum uric acid by isotope dilution mass spectrometry as a new candidate reference method. *Anal Chem* 1990;62;2173–2177.

Paulus HE, Coutts A, Calabro JJ, et al. Clinical significance of hyperuricemia in routinely screened hospitalized men. *JAMA* 1970;211:277–281.

Price CP, James DR. Analytical reviews in clinical biochemistry: the measurement of urate. *Ann Clin Biochem* 1988;25:484–498.

Ammonia

Flannery DB, Hsia VE, Wolf B. Current status of hyperammonemic syndromes. *Hepatology* 1982;2:495–506.

Forman DT. Rapid determination of plasma ammonia by an ion-exchange technic. *Clin Chem* 1964;10:497–499.

Gerron GG, Ansley JD, Isaacs JW, et al. Technical pitfalls in measurement of venous plasma NH_3 concentration. *Clin Chem* 1976;22:663–666.

Miller GE, Rice JD. Determination of the concentration of ammonia nitrogen in plasma by means of a simple ion exchange method. *Am J Clin Pathol* 1963;39:97–103.

Pesh-Iman M, Kumar S, Willis CE. Enzymatic determination of plasma ammonia: evaluation of Sigman and BMC kits. *Clin Chem* 1978;24:2044–2046.

Sundberg MW, Becker RW, Esders TW, et al. An enzymic creatinine assay and a direct ammonia assay in coated thin films. *Clin Chem* 1983;29:645–649.

Willems D, Steenssens W. Ammonia determined in plasma with a selective electrode. *Clin Chem* 1988;34:2372.

Renal Disease

Bricker NS, Kirschenbaum MA. *The kidney: diagnosis and management*. New York: John Wiley & Sons, 1984.

Cohen EP, Lemann J Jr. The role of the laboratory in evaluation of kidney function. *Clin Chem* 1991;37:785–796.

Duarte CG, Preuss HG. Assessment of renal function-glomerular and tubular. *Clin Lab Med* 1993; 13:33–52.

Espinel C. The FE_{NA} test: use in the differential diagnosis of acute renal failure. *JAMA* 1976;236:579–581.

Espinel C. Diagnosis of acute and chronic renal failure. *Clin Lab Med* 1993;13:89–102.

Finn WF. Diagnosis and management of acute tubular necrosis. *Med Clin North Am* 1990;74:873–891.

Harrington JT, Cohen JJ. Acute oliguria. *N Engl J Med* 1975;92:89–91.

Jacobson HR. Chronic renal failure: pathophysiology. *Lancet* 1991;338:419–423.

Pillay VRG. Clinical testing of renal function. *Med Clin North Am* 1971;55:231–241.

Schrier RW. *Renal and electrolyte disorders*, 4th ed. Boston: Little, Brown, 1992.

Wolfson M. Laboratory values in renal failure. *Lab Med* 1985;16:107–110.

20

Calcium, Magnesium, and Phosphate

John G. Toffaletti

The electrolytes calcium, magnesium, and phosphate are the principal inorganic constituents of bone and are vital in the function of membranes, hundreds of enzymes, genetic regulation, muscle contraction, and energy utilization. Their distribution between bone, soft tissues, and extracellular fluids is regulated by parathyroid hormone and vitamin D. As with most electrolytes, the kidneys are a most important regulator of the concentration of these ions in blood.

With the exception of parathyroid dysfunction and occult causes of hypercalcemia, the usefulness of these electrolytes in diagnostic tests was of only moderate interest through the 1970s. During the past 15 to 20 years, the discovery and awareness of the importance of calcium and magnesium in cellular regulation, energy metabolism, ion regulation, ischemia, and contraction, especially of myocardial and smooth muscle cells, has led to a much more prominent role for these tests in monitoring during and after surgery, in critical care, and in the neonate. The role of intracellular calcium ions and protection by magnesium ion in reperfusion injury following ischemia, the potentially fatal complications from hypomagnesemia in cardiovascular disease, and the importance of adequate phosphate for regulation of oxygen delivery and energy utilization are some examples of the vital importance of these electrolytes in critically ill patients.

Although the importance of these electrolytes in health and cellular survival is well established, the measurement of their physiologically active forms remains challenging. Measurements of ionized calcium are reliable and readily available on several analyzers. While measurements of ionized magnesium have achieved acceptable reliability, they are less reliable than measurements of ionized calcium and are in lower clinical demand. An even greater challenge arises in measuring and interpreting intracellular concentrations of magnesium and phosphate. Techniques that measure intracellular concentrations of any of these electrolytes remain relatively crude and too slow to be of value for clinical decision making in an acute care setting, but they are the subject of continuing research and development.

CALCIUM

Part of "20 - Calcium, Magnesium, and Phosphate"

Biochemistry and Physiology

In 1883, Ringer showed that calcium was essential for myocardial contraction (1). In the 1930s, McLean and Hastings studied the actions of bound and free forms of calcium on frog heart contraction. They showed that the ionized (free) calcium concentration was proportional to the amplitude of frog heart contraction, while protein-bound and citrate-bound calcium had no effect (2). From this observation, they concluded that the free ionized calcium was the physiologically active form of calcium in blood and used isolated frog hearts to develop the first assay for ionized calcium. While the method had poor precision by today's standards, they were able to show that blood-ionized calcium was both closely regulated and had a mean concentration in humans of about 1.18 mmol/L.

Of all the calcium in the body, 99% is in the bone. The remaining 1% is mostly in the blood and other extracellular fluids. The amount of calcium inside of cells is relatively low, with an extremely low concentration of free ionized calcium (5,000 to 10,000 lower than blood) maintained in the cytosol of many cells, such as cardiac or smooth muscle cells. Calcium circulates in the blood is several forms:

45% to 50% free ionized

40% to 45% bound to protein, mostly albumin

10% to 15% bound to anions such as bicarbonate, citrate, phosphate, and lactate

As shown in Figure 20.1, the bound forms of calcium ions are in equilibrium with the free calcium ions. pH has a significant effect on calcium ion binding to protein, with each 0.1 unit decrease or increase in pH causing ionized calcium to change inversely by about 0.05 mmol/L. The distribution of these forms of calcium can change, especially during surgery or critical care situations, where concentrations of citrate, bicarbonate, and lactate can change rapidly.

Because decreased ionized calcium may impair myocardial function, maintaining ionized calcium at a clinically acceptable concentration is an important goal both during surgery and in critically ill patients (3). The flow of calcium ions into cells in the myocardium helps control cardiac contraction and rhythm. Calcium binds to contractile proteins, which initiates the contractile process. The rate at which calcium ions flow into smooth muscle cells influences the tension of arterioles, which regulate blood pressure. A diagram showing these basic intracellular movements of calcium ions is shown in Figure 20.2.

Because the intracellular concentration of calcium ions is an important factor in muscle contraction, the regulation of calcium

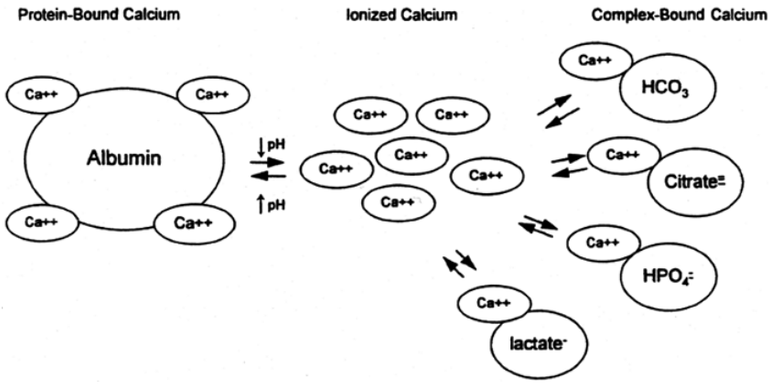


FIGURE 20.1. Equilibria between protein-bound, complex-bound, and ionized calcium in blood. (From AACC Self-Study Course: Understanding the Critical User of Blood Gases and Electrolytes. Washington DC: AACC Education Programs, 1998:41. Used with permission.)

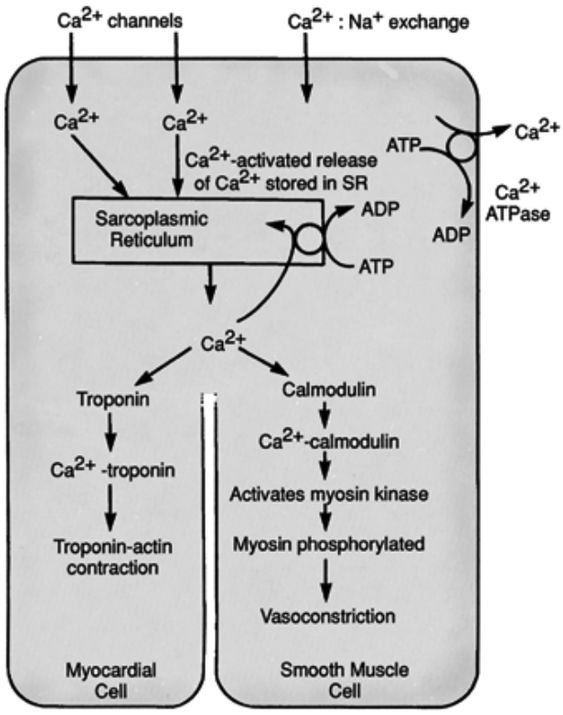


FIGURE 20.2. Basic intracellular movements of calcium ions that result in the contraction of myocardial or smooth muscle cells. The reactions characteristic of a myocardial cell are shown in the bottom left, while the reactions in a smooth muscle cell are shown in the bottom right of the figure.

ion flow into and out of cells is extremely important. Two types of calcium channels on the cell membrane are known to regulate this flow: a voltage-dependent and a phosphorylation-dependent channel. These gates function as follows: the voltage-dependent gate opens when the membrane is depolarized, while the phosphorylation-dependent gate requires a cAMP-activated protein kinase. The cardiotropic drugs epinephrine and isoproterenol facilitate transport of calcium ions through these channels, while acetyl choline hinders the transport of calcium ions (4). In addition, magnesium ion helps to stabilize these channels.

Note that this inward flux of calcium ions does not directly trigger contraction of the myocardial cell. Rather, these ions amplify intracellular calcium by releasing calcium ions from both the sarcoplasmic reticulum and the inner cell membrane.

To remove the excess calcium ions that have accumulated inside the cell, ATPases are necessary: a membrane-bound ATPase extrudes calcium ions from the cell, while a magnesium ATPase transports calcium ions back into the sarcoplasmic reticulum for future release.

Calcium ions also are important as “second messengers” in controlling the secretion of many hormones such as insulin, aldosterone, vasopressin, and renin. Following stimulation at the cell surface by a specific molecule (first messenger), the inward flux of calcium ions (second messenger) initiates intracellular events, such as production of a hormone.

Decreased ionized calcium concentrations in blood can cause both cardiovascular disorders, such as cardiac insufficiency and arrhythmias, and neuromuscular irritability, which may become clinically apparent as irregular muscle spasms, called tetany. Studies have shown that the rate of fall of ionized calcium in blood initiates tetany as much as the absolute decrease in concentration of ionized calcium (5).

Regulation of Calcium in the Blood

Three hormones are known to participate in serum calcium regulation, and have rates of secretion that depend on ionized calcium

concentration. These hormones are parathyroid hormone (PTH), 1,25 dihydroxyvitamin D, and calcitonin. Their actions are depicted in Figure 20.3. Parathyroid hormone-related protein (PTHRP) has hypercalcemic actions similar to PTH, however, the role of PTHRP in normal calcium homeostasis has yet to be clearly defined.

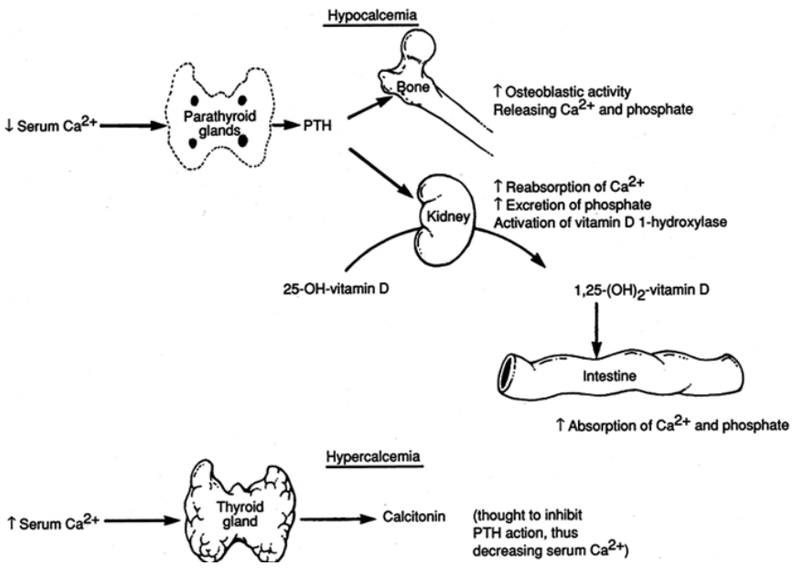


FIGURE 20.3. Hormonal responses to hypocalcemia and hypercalcemia.

The parathyroid gland responds rapidly to a decrease in ionized calcium, with a fourfold increase in PTH secretion stimulated by a 5% decrease in ionized calcium (6, 7). An increase in ionized calcium will reduce circulating PTH concentrations: a recent study found that when oral calcium increased serum ionized calcium by 0.07 mmol/L, serum intact PTH decreased (8).

PTH circulates as intact hormone and several fragments of varying biological activity. Both intact PTH and amino-terminal PTH have biological activity by an ability to bind to a PTH receptor on cell membranes (9).

PTH acts on both bone and kidney to increase calcium in blood. In the bone, PTH activates osteoclasts that break down bone and subsequently release calcium and phosphate into the extracellular fluid. In the kidneys, PTH conserves calcium by increasing tubular reabsorption of calcium ions and lowers phosphate by inhibiting tubular reabsorption of phosphate. PTH also activates a specific hydroxylase enzyme in the kidney that increases renal production of active vitamin D.

Vitamin D_3, a cholecalciferol, is most often obtained either in the diet or from exposure of the skin to sunlight. Vitamin D_3 from either source is converted in the liver to 25-hydroxycholecalciferol(25-OH-D_3), an inactive form of vitamin D. In the kidney, 25-OH-D_3 is specifically hydroxylated by a 1[α]-hydroxylase to form 1,25-dihydroxycholecalciferol($1,25-(OH)_2-D_3$), the most active form. This form of vitamin D increases calcium and phosphate absorption in the intestine and enhances the effect of PTH on bone resorption and calcium retention by the kidneys (9).

Calcitonin is synthesized and secreted by the parafollicular cells of the thyroid gland in response to an increased concentration of calcium in blood (9). While an acute infusion of calcium stimulates calcitonin secretion, calcitonin may not contribute to normal calcium homeostasis, because calcitonin does not change in response to a 1% to 2% increase in ionized calcium. Chronic hypercalcemia has little effect on calcitonin secretion. A study on hemodialysis patients showed that ionized calcium had to increase by about 10% to elicit a marked calcitonin response (10). Calcitonin inhibits osteoclast activity and apparently exerts its calcium-lowering effect by inhibiting the actions of both parathyroid hormone and vitamin D.

Distribution Of Calcium in Blood and Cells

Over 99% of the calcium in the body is part of bone. The remaining 1% is mostly in the blood and other extracellular fluid, with very little in intracellular stores. In fact, the concentration of ionized calcium in blood plasma is 5,000 to 10,000 times higher than in the cytosol of cardiac or smooth muscle cells. Maintenance of this large gradient is vital to allowing the rapid inward flux of calcium ions that is essential for the initiation of muscle contraction.

Calcium in blood is distributed among several forms. About 45% circulates as ionized calcium (i.e., free calcium ions), 40% is bound to anionic sites on protein (mostly albumin), and 15% is bound to anions such as bicarbonate, citrate, phosphate, and lactate (11). This distribution can change in disease because concentrations of citrate, bicarbonate, lactate, phosphate, and albumin can change dramatically, especially during surgery or in critically ill patients. This is the principal reason why ionized calcium cannot be accurately calculated from total calcium measurements. In addition, calcium binding to protein is dependent on pH. As pH increases, hydrogen ions dissociate from anionic sites on protein and more calcium ions bind to protein, which lowers the ionized calcium concentration. As a rough guide, an increase of pH by 0.1 pH unit will decrease ionized calcium by about 0.06 mmol/L (12).

Clinical Interpretation of Calcium Measurements

Most cases of abnormal serum calcium concentration are related to one of the following:

A defect in parathyroid function

A defect in vitamin D metabolism

Malignancy

Renal disease

An iatrogenic cause (e.g., administration of citrate, calcium, or saline)

In addition, calcium measurements are useful in the monitoring of neonates, critically ill patients, and patients undergoing surgery.

Both total calcium and ionized calcium measurements are available in many laboratories. The relative merits of each test are briefly mentioned in this section. As emphasized many years ago by Ladenson et al. (13), ionized calcium cannot be accurately calculated from a total calcium concentration, although a recent editorial mentions that ionized calcium can be calculated from total calcium, total protein, albumin, and anion gap, a total of six to seven measurements (14). Given the simplicity, speed, and accuracy of current ionized calcium analyzers, it would seem more efficient to measure ionized calcium directly.

Ionized calcium measurements also may be reported as a pH-adjusted ionized calcium parameter ($[Ca^{2+}]_{7.4}$). This was introduced as a convenience to allow the measurement of ionized calcium in serum collected for the performance of routine chemistry tests. These samples usually are exposed to room air, and the pH-adjusted parameter attempts to correct for loss of carbon dioxide, which increases pH. The increased pH increases calcium-ion binding by proteins, which lowers the ionized calcium concentration. Because the pH correction may be less accurate in patients undergoing surgery, patients who are critically ill, patients with renal disease, or neonates, ionized calcium measured on anaerobically collected blood should be used.

Hypocalcemic Disorders

Hypocalcemia occurs most commonly in patients receiving citrated blood products during major surgery, in patients with renal disease, in patients with parathyroid gland insufficiency following surgery, in neonates, in patients with a magnesium deficiency, and in patients with pancreatitis. A flow diagram for the evaluation of hypocalcemic disorders is shown in Figure 20.4.

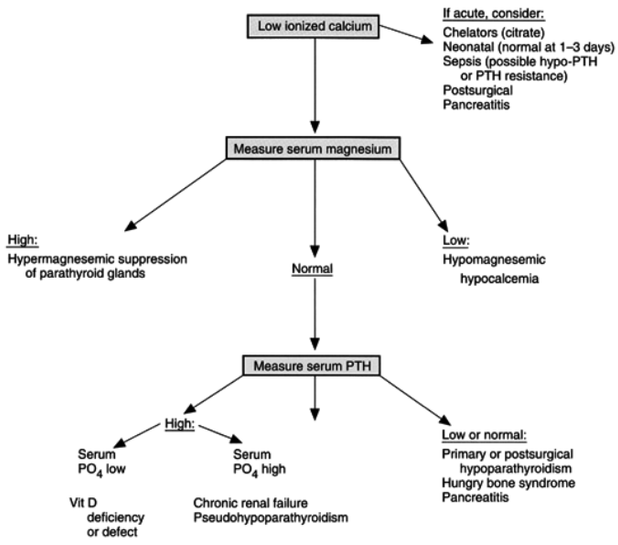


FIGURE 20.4. Evaluation of hypocalcemia.

HYPOPARATHYROIDISM

Part of "20 - Calcium, Magnesium, and Phosphate"

While hypoparathyroidism may be acquired by autoimmune or other chronic processes, most cases of hypocalcemia from hypoparathyroidism result from surgical removal of tissue during parathyroid or thyroid surgery. Hypocalcemia usually is transient (lasting less than 5 days) unless surgery has removed too much parathyroid tissue or has interfered with the parathyroid blood supply. Zaloga and Chernow recommend that serum ionized calcium should be monitored every 12 hours after neck surgery until a rise in the ionized calcium concentration indicates recovery of the parathyroid gland (4).

Pseudohypoparathyroidism is a condition in which renal cells do not respond to PTH. The combined effects of PTH resistance and hyperphosphatemia may suppress renal 1[α]hydroxylase activity, leading to vitamin D deficiency and hypocalcemia.

RENAL DISEASES

Part of "20 - Calcium, Magnesium, and Phosphate"

Patients with renal glomerular disease often have altered concentrations of calcium, phosphate, albumin, magnesium, and hydrogen ion (pH). Because these conditions tend to change ionized calcium independently of total calcium, ionized calcium is usually better for accurately monitoring calcium status in renal disease (15). In chronic renal disease, a marked increase in serum PTH levels (secondary hyperparathyroidism) compensates for hypocalcemia caused either by hyperphosphatemia, in which phosphate binds calcium, or by altered vitamin D metabolism. Interestingly, among end-stage renal disease patients beginning dialysis, initial PTH levels above 200 pg/mL were associated with half the mortality of similar patients with PTH levels below 65 pg/mL. Controlling ionized calcium by monitoring its concentration can avoid problems from hypocalcemia, such as osteodystrophy, unstable cardiac output, unstable blood pressure, and problems arising from hypercalcemia such as renal stones or other soft-tissue calcifications (7, 16).

NEONATAL MONITORING

Part of "20 - Calcium, Magnesium, and Phosphate"

The ionized calcium concentration in the blood of neonates typically is high at birth, then declines by 10% to 20% after 1 to 3 days. After about a week, ionized calcium concentrations in the neonate stabilize at concentrations slightly higher than in adults (16), and no treatment usually is needed.

Neonatal hypocalcemia may be either of two types. The first type, which normally develops within 1 to 3 days after birth, is associated with parathyroid immaturity and usually resolves by the first week of life. This type of neonatal hypocalcemia in healthy infants may be a normal stimulus to activate parathyroid

gland function. The second type of neonatal hypocalcemia usually is seen about 1 week after birth and is associated with both hyperphosphatemia from a high phosphate intake from milk and hypomagnesemia caused by decreased intestinal absorption of magnesium (4). Other factors that increase the incidence and severity of hypocalcemia during this period are prematurity, maternal diabetes mellitus, complications during delivery, and birth asphyxia (4). Hypocalcemia even may become life-threatening if it is severe, prolonged, or accompanied by seizures, hypotension, hypoglycemia, or sepsis. Neonatal seizures occur especially in infants of diabetic mothers, with hypocalcemia and hypomagnesemia being possible causes (18).

Ionized calcium measurements are preferred over total calcium in monitoring calcium status in the neonate (17). Rather large changes in the concentration of ionized calcium may occur in the early neonatal period, because calcium may be lost rapidly and not readily reabsorbed. Several possible etiologies have been suggested, including abnormal PTH and vitamin D metabolism, hypercalcitoninemia, hyperphosphatemia, and hypomagnesemia (17).

MONITORING IN SURGERY AND ACUTE CARE

Part of "20 - Calcium, Magnesium, and Phosphate"

Largely because of the importance of both transcellular and intracellular movement of calcium ions in myocardial and vascular smooth muscle cells, the maintenance of a normal ionized calcium concentration in blood is important to the patient in either surgery or intensive care. Adequate calcium concentrations promote good cardiac output and maintain adequate blood pressure. Monitoring and adjusting calcium concentrations may be most critical in open-heart surgery when the heart is restarted. Normalizing ionized calcium by administering calcium as a cardiotropic agent is preferred before giving drugs such as epinephrine or isoproterenol (3). Monitoring of ionized calcium is important during liver transplantation, because large volumes of citrated blood are given at a time when the function of the liver (the major organ for metabolizing citrate) is compromised or absent.

Because patients in surgery may receive large amounts of citrate, bicarbonate, calcium salts, and fluids, clear discrepancies between total calcium and ionized calcium are common in these patients. In these cases, ionized calcium measurements should be used wherever possible.

CRITICALLY ILL PATIENTS

Part of "20 - Calcium, Magnesium, and Phosphate"

In critically ill patients, including those with sepsis, thermal burns, renal failure, and/or cardiopulmonary insufficiency, hypocalcemia occurs in as many as 70% of cases (19). Sepsis appears to cause hypocalcemia and may correlate with the degree of

hypocalcemia (20). The inflammatory response and mediators of sepsis are involved in reducing ionized calcium concentrations (21).

In patients with hypocalcemia, calcium administration can increase cardiac output and blood pressure (21). A direct relationship has been shown between serum ionized calcium and mean arterial pressure in a study of 112 patients admitted to a medical intensive-care unit (22). While it may seem that correction of hypocalcemia will always benefit cardiovascular status, this is not clearly established. Hypocalcemia may serve as a protective mechanism against cell death during cellular hypoxia. One report suggests that when calcium administration is slowly titrated to no higher than normal concentrations by measurement of blood ionized calcium, adverse side effects were not observed (21).

HYPOMAGNESEMIA

Part of "20 - Calcium, Magnesium, and Phosphate"

Chronic hypomagnesemia has become recognized as a frequent cause of hypocalcemia. There are three proposed mechanisms:

1. Inhibition of transport of PTH across the parathyroid gland membrane (23)
2. Impairment of PTH action at its receptor site on bone (4)
3. Interference with the action of vitamin D (24).

Hypercalcemic Disorders

A flow chart for the evaluation of hypercalcemia is presented in Figure 20.5. While a total calcium results of 3 mmol/L or above virtually always indicates hypercalcemia, a total calcium lower than 3 mmol/L should be confirmed by measurement of ionized calcium, if available. Malignancy and primary hyperparathyroidism are the most common causes of hypercalcemia, accounting for 80% to 90% of cases. Hypercalcemia as a result of malignancy is more common in hospital populations, while primary hyperparathyroidism is more commonly found in the outpatient setting.

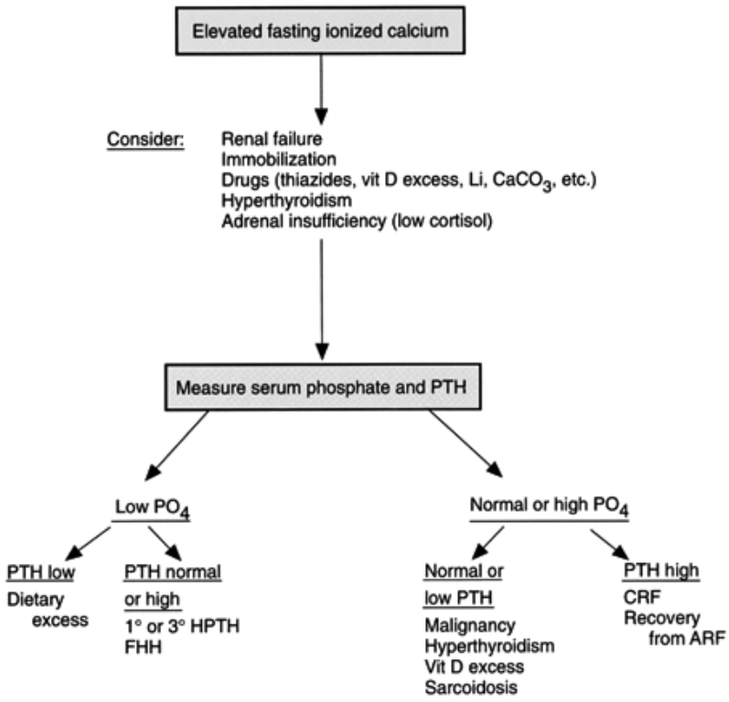


FIGURE 20.5. Evaluation of hypercalcemia. Abbreviations: 1° HPTH, primary hyperparathyroidism; 3° HPTH, tertiary hyperparathyroidism; FHH, familial hypocalciuric hypercalcemia; ARF, acute renal failure; CHF, chronic renal failure.

While hyperparathyroidism refers to excessive secretion of parathyroid hormone, this excess can range from slight to a very large amount. Hyperparathyroidism therefore may either show obvious clinical signs or be asymptomatic. Several studies have shown that, while both total and ionized calcium measurements are elevated in severe hyperparathyroidism, ionized calcium more frequently is elevated in subtle or asymptomatic hyperparathyroidism. Ionized calcium measurements have been reported to be elevated in 90% to 95% of cases of hyperparathyroidism, while total calcium is elevated in 80% to 85% of cases (25). A recent report found that of 60 cases of confirmed hyperparathyroidism, 59 had elevated ionized calcium, while 47 had an elevated corrected total calcium (26).

A new clinical use of PTH measurements has developed from the need to measure PTH rapidly during operations on the parathyroid glands to determine if sufficient parathyroid tissue has been removed to correct hyperparathyroidism (27).

Hypercalcemia is present in up to 20% of various types of malignancy, with calcium measurements sometimes serving as biochemical markers for otherwise inapparent disease (Table 20.1). The increase may be caused by direct lysis of bone by tumor

or by the secretion of PTHRP. PTHRP appears to be expressed in tumors derived from lung, breast, kidney, and many squamous cell carcinomas and T-cell lymphomas (9). PTHRP, being very similar to PTH in its first 13 amino acids, binds to and activates the PTH receptor and mimics the actions of PTH on bone, kidney, and intestine. While measurements of PTHRP ultimately may have clinical diagnostic value (28), they only rarely are used and measurement of PRHRP is available in very few clinical laboratories.

TABLE 20.1. Interpretation of Laboratory Tests in Differentiating Primary Hyperparathyroidism from Malignancy

Test	Favors HPTH	Favors Malignancy
Total Ca (mmol/liter)	<3.13	>3.13
Serum Cl (mmol/liter)	>103	<103
Intact PTH	Elevated	Low or normal
Serum PO_4	N to low	Normal to high
Hematocrit	Normal	Low
Urine Ca	High	Very high
1,25-$(OH)_2$ vitamin D	High	Low
PTHRP	Normal (undetected)	Elevated

PTH, parathyroid hormone; PTHRP, PTH-related peptide.

Although ionized calcium may have slightly higher diagnostic sensitivity, ionized and total calcium measurements have about equal utility in the detection of occult malignancy (29). Carcinoma of the bronchus, breast, urogenital tract, and head and neck area, and multiple myelomas account for 75% of hypercalcemia in malignancy.

Proper Collection and Handling of Samples

Total calcium determinations may be performed on either serum or heparinized plasma. Because anticoagulants such as EDTA and oxalate bind calcium very tightly and interfere with its measurement, even by atomic absorption, they must not be used.

The proper collection of samples for either total or ionized calcium measurements requires both care and appreciation of the problems that can occur. Tourniquet application should be as brief as possible (less than 1 minute) before a specimen is collected: the combined effects of hemoconcentration (causing hyperproteinemia) and localized lactic acid production can alter total calcium by up to 10% (30) and ionized calcium by about 2% to 3% (30, 31). Because loss of CO_2 will increase pH, all samples for ionized calcium measurements must be collected anaerobically. The metabolic activity of cells during storage affects pH and therefore ionized calcium. Although the effects of the lactate anion and the hydrogen cation partly offset each other (31), blood should be centrifuged within 1 hour to prevent acidosis from affecting the ionized calcium concentration (32, 33). While anticoagulated whole blood can be analyzed more rapidly, serum from sealed, evacuated blood collection tubes is satisfactory for ionized calcium if clotting and centrifugation are done within 1 hour. Because of dilutional effects, no liquid heparin products should be used. Heparin also binds a small amount of calcium and tends to lower ionized calcium concentrations. However, if the heparin concentration is less than 30 IU/mL, the interference should be less than 4%. Dry heparin products are available that virtually eliminate the interference by heparin. These products are available either as heparin titrated with small amounts of calcium or zinc ions or as a small amount of heparin dispersed in an inert soluble material. For analysis of calcium in urine, an accurately timed urine collection is preferred. The urine should be acidified with 6 mol HCl/L, with approximately 1 mL of the acid added for each 100 mL of urine.

Methods for Measurement of Total and Ionized Calcium

While atomic absorption spectroscopy remains the reference method for total calcium, accurate measurements require care, both in maintenance and calibration of the instrument and in preparation of samples. In routine testing of serum, many automated analyzers give results that are comparable to atomic absorption. However, for analysis of total calcium in urine and other fluids, atomic absorption still is preferred.

Atomic absorption involves introducing a diluted sample into an air-acetylene flame and measuring the absorption of light at 422.7 nm.

Many automated methods for total calcium measurement are based on the complexometric reaction between calcium and the dye ortho-cresolphthalein complexone, often with 8-hydroxyquinoline added to prevent magnesium interference (33). This methodology was used on the duPont aca, Beckman Astra 8 and Technicon SMAC. Some automated analyzers, including the J and J Vitros and Beckman analyzers use another dye, Arsenazo III.

Ionized calcium is measured using an ion-selective electrode system. These electrodes use membranes impregnated with molecules (ionophores) that selectively, but reversibly, bind calcium ions (35). As calcium ions bind to these membranes, an electric potential develops across the membrane that is related to the ionized calcium concentration. The potential is measured with a sensitive voltmeter in the same way that pH is measured.

Reference Ranges for Calcium

Ionized calcium concentrations can change rapidly during the first 1 to 3 days of life. Following this, they stabilize at relatively high levels with a gradual decline through adolescence. For total calcium, the reference range varies slightly with age. Calcium concentrations are relatively high through adolescence when bone growth is most rapid (see Table 20.2).

TABLE 20.2. Reference Ranges for Calcium[a]

	mmol/liter	mg/dl
Total calcium		
Child	2.20–2.68	8.8–10.7
Adult	2.10–2.55	8.4–10.2
Ionized calcium		
At birth	1.30–1.60	5.2–6.4
Neonate	1.20–1.48	4.8–5.9
Child	1.20–1.38	4.8–5.5
Adult	1.16–1.32	4.6–5.3

[a] Urine reference ranges will vary with diet. Adults on an average diet typically excrete 12 to 75 mmol/day (50 to 300 mg/day).

MAGNESIUM

Biochemistry and Physiology

Awareness of the physiological role and the clinical importance of magnesium has developed relatively recently. The first clinical symptoms of hypomagnesemia were not reported until 1934 by Hirschfelder and Haury. In 1969, Shils reported that magnesium depletion was associated with hypomagnesemia, hypokalemia, and hypocalcemia (36). While hypomagnesemia has been overlooked frequently in the past, there now is a great awareness of the clinical importance of magnesium. This is the result of many articles describing the effects of magnesium on myocardial function and blood pressure (37), the role of magnesium as a calcium-channel–blocking agent (38), the clinical consequences of hypomagnesemia (39), and the implications of magnesium depletion during open-heart surgery (40) and critical care (41).

The clinical importance of magnesium reflects its profound biochemical importance. Magnesium is an essential activator of over 300 enzymes, including those important in glycolysis, gene replication, transcellular ion transport, muscle contraction, and oxidative phosphorylation. In fact, the magnesium complex of adenosine triphosphate (ATP) probably is the true substrate in energy production, rather than ATP alone. It can be said fairly that magnesium is equal in importance to any of the other electrolytes in the body, many of which have been much more intensively studied and are more commonly measured.

Regulation of Magnesium in the Blood

The average dietary intake of magnesium is 10 to 15 mmol/day. Rich sources are green vegetables, meat, grains, and seafood. The small intestine may absorb from 20% to 63% of the dietary magnesium, depending on the need. The overall regulation of body magnesium is controlled largely by the kidney, which can avidly reabsorb magnesium in deficiency states or can readily excrete excess magnesium in overload states. The renal threshold for magnesium is about 0.60 to 0.85 mmol/L. Because this is close to the normal serum concentration, slight excesses of magnesium in serum are excreted rapidly by the kidneys. Because renal magnesium reabsorption competes with reabsorption of calcium in the ascending limb of the loop of Henle, the loss of calcium by diuretics, hypercalcemia, or saline infusion also will cause loss of magnesium (42).

Although magnesium and calcium appear to have several common factors in their regulation, regulation of magnesium is not as well characterized as that of calcium. Parathyroid hormone increases renal reabsorption of magnesium and enhances absorption of magnesium in the intestine. However, for equivalent decreases in ionized calcium and magnesium in blood, changes in ionized calcium have a far greater effect on PTH secretion (7, 43). Paradoxically, chronic or severe acute hypomagnesemia can depress secretion of PTH, one mechanism by which hypomagnesemia can cause hypocalcemia. Aldosterone apparently has the opposite effect of PTH, increasing the renal excretion of magnesium. Insulin appears to increase both intestinal and renal absorption of magnesium, yet also increases intracellular concentrations of magnesium.

Distribution of Magnesium in Tissues and Blood

The human body contains about 1 mol (24 g) of magnesium, with about 50% in the skeleton and relatively high concentrations in skeletal muscle, liver, and myocardium. Magnesium is primarily an intracellular ion, with only about 1% being in the blood and extracellular fluid (44). Magnesium in serum exists as protein-bound (24%), complex-bound (10%), and ionized (66%) forms (data derived from reference 43). Of course, these proportions vary with the concentration of proteins and other anions in blood, and, as with calcium, the extent of binding is affected by blood pH.

Clinical Importance of Magnesium Measurements

Spontaneous hypermagnesemia rarely is observed in clinical practice, with most cases resulting from administration of magnesium-containing antacids, enemas, or parenteral nutrition to patients with renal insufficiency (42). In contrast, hypomagnesemia can arise from a number of causes (45, 46 and 47), which are listed in Table 20.3 (48). Magnesium should be measured during the initial examination of acutely or chronically ill patients, especially those with poor food intake, malabsorption disorders, hypokalemia, or hypocalcemia, and in those taking diuretics (45). Hypomagnesemia has been defined as a serum magnesium of less than 0.75 mmol/L (49). The incidence of hypomagnesemia in hospitalized patients has been reported to be from 5% to 50%, with variations in reference ranges, methods, and populations studied accounting for some of this variation.

TABLE 20.3. Causes of Hypomagnesemia

Diarrhea
Diuretics, especially loop diuretics (furosemide)
Diabetes
Dietary
Alcoholism
Drugs, especially cyclosporin and cisplatin
Cellular hypoxia
Toxemia or eclampsia of pregnancy
Skin loss by burns

CRITICAL CARE

Part of "20 - Calcium, Magnesium, and Phosphate"

Magnesium deficiency is found in a large percentage of critically ill patients (50, 51). In patients admitted to intensive care, the presence of hypomagnesemia is associated with an increased mortality rate (52), and may warrant magnesium replacement therapy if serum magnesium is below 0.5 mmol/L (41). In acute-care patients, diagnoses commonly associated with hypomagnesemia include coronary artery disease and coronary bypass surgery, malignancy, chronic obstructive pulmonary disease, and alcoholism. Among chronic diseases, alcoholism, liver disease, and carcinoma were commonly associated with hypomagnesemia (53).

Magnesium was the most common electrolyte abnormality found among pediatric intensive-care unit patients (54). There

also is evidence that ionized magnesium may be decreased in many critically ill pediatric patients who have normal total magnesium concentrations (55).

CARDIAC DISORDERS

Part of "20 - Calcium, Magnesium, and Phosphate"

Magnesium ion plays an important role in energy metabolism, calcium-channel regulation, myocardial contraction, and other cardiac functions. The heart, with its high metabolic activity, is particularly vulnerable to magnesium deficiency, which has been associated with coronary vasospasm, arrhythmias, acute infarction, and sudden death (36, 57). These conditions may arise because hypomagnesemia promotes loss of myocardial potassium and increases the ratio of calcium to magnesium, both of which alter myocardial electrophysiology (58). The extracellular concentration of ionized magnesium apparently controls arterial tone and blood pressure by regulating magnesium-calcium and sodium-calcium exchange sites on the membrane (59) and by modulating the vasoconstrictive effects of hormones such as norepinephrine and angiotensin II: a high magnesium concentration antagonizes their effects, while a low magnesium concentration enhances their activity (60). A reduction of ionized magnesium in the extracellular fluid produces sustained constriction of arterioles and venules, coronary vasospasm, and increases the potency of vasoconstrictive agents.

One of the most important consequences of hypomagnesemia in cardiovascular disease may be that cellular loss of magnesium uncouples oxidative metabolism that, by depleting ATP, leads to disruption of mitochondrial function and structure. Moreover, in congestive heart failure (CHF), cardiac output is diminished and is inadequate for the metabolic needs of tissues. Because deficiencies of magnesium and potassium lead to high blood pressure, causing increased vascular resistance, the problems of diminished cardiac output of CHF are further intensified (59).

An area that has received increasing attention over the years is the use of prophylactic magnesium to minimize complications in both myocardial infarction (61, 62) and after cardiac operations (63).

In patients undergoing surgery with cardiopulmonary bypass, plasma ionized (but not total) magnesium was decreased significantly by 24 hours after bypass (64). Because magnesium promotes vasodilation and inhibits platelet aggregation and adhesion, hypomagnesemia may promote coronary vasoconstriction and thrombosis in the early postoperative period (63). To avoid this, two to four grams of magnesium now are routinely administered after cardiac operations at our institution. Measurements of ionized magnesium may prove to be clinically useful during this therapy.

DRUG EFFECTS ON MAGNESIUM CONCENTRATION

Part of "20 - Calcium, Magnesium, and Phosphate"

Several drugs, including diuretics, gentamicin and other aminoglycoside antibiotics, cisplatin, and cyclosporine increase renal loss of magnesium and result frequently in hypomagnesemia. The loop diuretics, such as furosemide, especially are potent in increasing renal loss of magnesium, while thiazide diuretics usually require chronic use to develop hypomagnesemia (65). Gentamicin inhibits reabsorption of magnesium in the renal tubule. Because hypomagnesemia intensifies the toxic side effects of gentamicin, hypomagnesemia especially should be avoided (66). Cisplatin also is nephrotoxic, and profound hypomagnesemia can be caused by cyclosporine, an immunosuppressant widely used following organ transplantation. Cyclosporine also severely inhibits the renal tubular reabsorption of magnesium in addition to its other side effects, which include hypertension and hepatotoxicity (67).

DIABETES MELLITUS

Part of "20 - Calcium, Magnesium, and Phosphate"

Hypomagnesemia has been reported in from 25% to 75% of patients with diabetes (68). The mechanism of magnesium deficiency appears to be magnesium loss secondary to ketoacidosis and glycosuria (49), along with possible abnormal intracellular-extracellular distributions of magnesium caused by hormonal imbalances. Hypomagnesemia may lead to insulin resistance, diabetic retinopathy, hypertension, cardiovascular disease, increased platelet activity and thrombosis, and decreased PTH and altered vitamin D metabolism. The American Diabetic Association issued a statement about magnesium and diabetes. Among the conclusions were:

1. Hypomagnesemia often results from glycosuria
2. Hypomagnesemia probably does not play a primary role in the pathogenesis of diabetes
3. A strong relationship exists between hypomagnesemia and insulin resistance
4. Magnesium supplements should be given to all patients with documented hypomagnesemia, but not to all patients with diabetes (69).

Acute hypomagnesemia can result from intracellular shifts of magnesium following the administration of glucose or amino acids (46, 47). This effect is pronounced following starvation or insulin treatment for hyperglycemia.

ALCOHOLISM

Part of "20 - Calcium, Magnesium, and Phosphate"

Many studies have found a marked deficiency of total magnesium in chronic alcoholism (70, 71). Both acute and chronic ingestion of alcohol are associated with renal magnesium losses and decreased muscle magnesium content (72). Other factors that may cause hypomagnesemia in alcoholic patients include dietary deficiency, ketosis, vomiting, diarrhea, and hyperaldosteronism. Alcoholic patients may become hypomagnesemic after treatments such as dextrose infusion, which induces insulin secretion and may cause a shift of magnesium back into cells to correct a chronic intracellular magnesium deficiency. A recent report found that both total and ionized magnesium increased after three weeks of abstinence (71). While total magnesium was lower in alcoholics than controls, ionized magnesium was either similar to or decreased relative to controls, depending on the type of ionized magnesium analyzer used.

OTHER DISEASES

Part of "20 - Calcium, Magnesium, and Phosphate"

Because magnesium requirements increase during pregnancy, hypomagnesemia may develop if intake also is not increased. Premature labor and preeclampsia or eclampsia are clearly associated with the development of hypomagnesemia. Treatment with magnesium salts has long been an accepted practice in these conditions, although hypermagnesemia sometimes results from excessive doses. The fetus also may be affected by hypermagnesemia in such cases.

Neonatal seizures occur especially in infants of diabetic mothers, with hypocalcemia and hypomagnesemia being possible causes (18). Hypomagnesemia is common in many critically ill children, with more having decreased ionized magnesium than decreased total magnesium (55).

Magnesium deficiency may be an important risk factor in renal stone formation and other renal calcification (73). Magnesium ion probably inhibits calcification by pairing with anions that would otherwise form insoluble salts with calcium.

Cerebrospinal fluid magnesium, but not serum magnesium, has been found to be significantly lower in acute schizophrenic patients than in schizophrenic patients in remission (74). A similar association of CSF (but not serum) magnesium also has been reported to be a prognostic indicator of the intensity of the neurologic disorder in stroke victims (75).

Patients with Paget's disease who also have hypomagnesemia tend to have more active disease as measured by increased serum alkaline phosphatase and urinary hydroxyproline. Because urinary magnesium is not significantly increased in Paget's disease, the low serum magnesium apparently is caused by increased uptake of magnesium into bone. For these reasons, patients with Paget's disease should increase their intake of magnesium to about 10 mmol per day (76).

Proper Collection and Handling of Samples

The patient should fast before collection of blood. Because of the threefold higher concentration of magnesium in erythrocytes, serum should not remain in contact with the clot for an extended period of time. Hemolyzed samples are not acceptable. The anticoagulants citrate, oxalate, and EDTA bind magnesium tightly and must not be used.

Measurement of Magnesium

While measurement of total magnesium concentrations in serum remains the usual diagnostic test for the detection of magnesium abnormalities, serum concentrations have two limitations. First, about 30% of magnesium is protein bound. Therefore, as with total and ionized calcium, total magnesium may not reflect the physiologically active magnesium. Second, and probably of greater significance, because magnesium is primarily an intracellular ion, serum concentration will not necessarily reflect the status of intracellular magnesium. Even when cellular magnesium is depleted by 20%, serum magnesium concentration may remain normal.

A magnesium load test may detect body depletion of magnesium in patients with a normal concentration in serum. However, a magnesium load test requires over 48 hours to complete. After collecting a baseline 24-hour urine, $MgSO_4$ is administered while another 24-hour urine sample is collected. Individuals with adequate body stores of magnesium will excrete 60% to 80% of the magnesium load within 24 hours, while magnesium-deficient patients excrete less than 50% (42).

Of the numerous methods developed for measurement of magnesium over the years, at least four have accuracy and precision that are acceptable for clinical laboratories (77, 78). These methods are atomic absorption spectroscopy, colorimetric methods using either calmagite or methylthymol blue, and a dry-slide colorimetric method using a formazan dye and a calcium chelator.

Calmagite is a metallochromic derivative of naphthol sulfonic acid that can be used to measure magnesium without deproteinization of the sample. The blue-colored reagent forms a pink magnesium-calmagite complex, with the absorbance measured at around 532 nm. The use of ethylene glycol tetraacetic acid (EGTA) prevents interference by calcium, and cyanide is used to bind heavy metals, preventing them from binding to the dye. The reaction is completed within 60 seconds, which allows the calmagite method to be used on automated analyzers (79).

Methylthymol blue is also widely used for determination of magnesium, with calcium chelators added to increase the specificity. This method is used on the duPont aca, with the absorbance measured at both 510 and 600 nm (79). The dry-slide method uses a multilayered reagent with both a calcium chelator and a magnesium-sensitive formazan dye. The magnesium-formazan complex is measured at 630 nm (78).

The reference method for magnesium measurement is atomic absorption spectroscopy. This method is capable of excellent accuracy and precision (77) in the hands of an experienced operator. Neutron activation analysis with the radioactive isotope magnesium-27 has been proposed as a definitive method (79).

The routine measurement of ionized magnesium now is possible by the development of magnesium ion-selective electrodes. The sensors for these electrodes were developed from studies on over 200 ionophores by the late Dr. Wilhelm Simon and his colleagues (80). Two different companies, each with their own ionophore, have introduced such electrodes that are approved by the Food and Drug Administration (81). Both electrodes also are sensitive to calcium ions, an effect that is corrected for by simultaneous measurement of the calcium ion concentration. Free magnesium has been quantitated as ultrafiltrable magnesium (43, 82). While this may give a reasonable estimate of ionized magnesium in many cases, ultrafiltrable measurements can significantly overestimate the concentration of both ionized magnesium and ionized calcium following administration of citrate (6).

Reference ranges for total magnesium in serum and other fluids are shown in Table 20.4, along with reference ranges for ionized magnesium in blood. Note the slightly different reference ranges for ionized magnesium as determined by the two brands of analyzers currently available (83).

TABLE 20.4. Reference Ranges for Magnesium[a]

	mmol/liter
Total magnesium	
Serum (newborns)	0.50–0.90
Serum (adults)	0.65–1.05
Erythrocytes	1.65–2.65
CSF	1.0–1.40
Urine	1–5 mmol/day
Ionized magnesium (83)	
Blood	0.44–0.60 (AVL analyzer) 0.39–0.64 (Nova analyzer)

[a] Concentrations of magnesium in serum are lower at birth than in older children or adults.

PHOSPHATE

Part of "20 - Calcium, Magnesium, and Phosphate"

Biochemistry and Physiology

Compounds of phosphorus are in all cells and participate in many biochemical processes. The genetic materials deoxiboxynucleic acid (DNA) and ribonucleic acid (RNA) are complex phosphodiesterases. Phosphate is a component of nucleotides, phospholipids, and most coenzymes. The most important reservoirs of biochemical energy are ATP, creatine phosphate, and phosphoenol pyruvate. One explanation for the widespread occurrence of phosphate in biochemical systems is that phosphate has three oxygen radicals. This allows linkage with two other molecules while still leaving one basic oxygen, which can be protonated or can participate in other reactions (84). Inorganic phosphate is important for maintaining adequate cellular stores of compounds such as ATP, and 2,3-DPG (85). Phosphate deficiency can lead to ATP depletion, which ultimately is responsible for many of the clinical symptoms observed in hypophosphatemia.

About 80% of the 700 to 800 grams of phosphate in the body are contained in bone, mostly in the form of hydroxyapatite [$Ca_{10}(PO_4)_6(OH)_2$]. Phosphate in blood is either absorbed from dietary sources or resorbed from bone. Most phosphate is found within cells, and the transport of glucose into cells is accompanied by an influx of phosphate. Intracellular phosphate then is used in the synthesis of phosphorylated compounds (86).

The kidney plays an important role in the regulation of serum phosphate concentration, and the renal reabsorption of phosphate is affected by several factors, including acid-base status, vitamin D, and PTH. PTH inhibits the normal tubular reabsorption of phosphate, increasing loss in the urine and decreasing serum phosphate concentrations. Therefore, PTH can have a dramatic effect on phosphate regulation, because the renal tubules normally reabsorb over 90% of phosphate filtered at the glomerulus.

While PTH lowers serum phosphate, vitamin D increases phosphate concentrations in the blood by promoting both phosphate absorption in the intestine and phosphate reabsorption in the kidney. In fact, phosphate may have a direct effect on parathyroid cells to enhance PTH secretion (87). Growth hormone, which helps regulate skeletal growth, also can affect circulating concentrations of phosphate. With excessive secretion or administration of growth hormone, phosphate concentrations in the blood tend to increase because of decreased renal excretion. However, growth hormone administered together with parathyroid hormone increases renal excretion of phosphate (88).

Although the total concentration of phosphate (expressed as phosphorus) in blood is about 12 mg/dL, most of this is organic phosphate; only about 3 to 4 mg/dL is inorganic phosphate. HPO_4^{2-} and $H_2PO_4^-$ are an important buffer pair; at pH 7.4 inorganic phosphate is about 75% HPO_4^{2-} and 25% $H_2PO_4^-$.

Clinical Interpretation of Phosphate Measurements

Both hyperphosphatemia and hypophosphatemia are seen frequently in clinical medicine, and both conditions are important to recognize. Flow diagrams for the evaluation of hypophosphatemia and hyperphosphatemia are shown in Figure 20.6 and Figure 20.7, respectively.

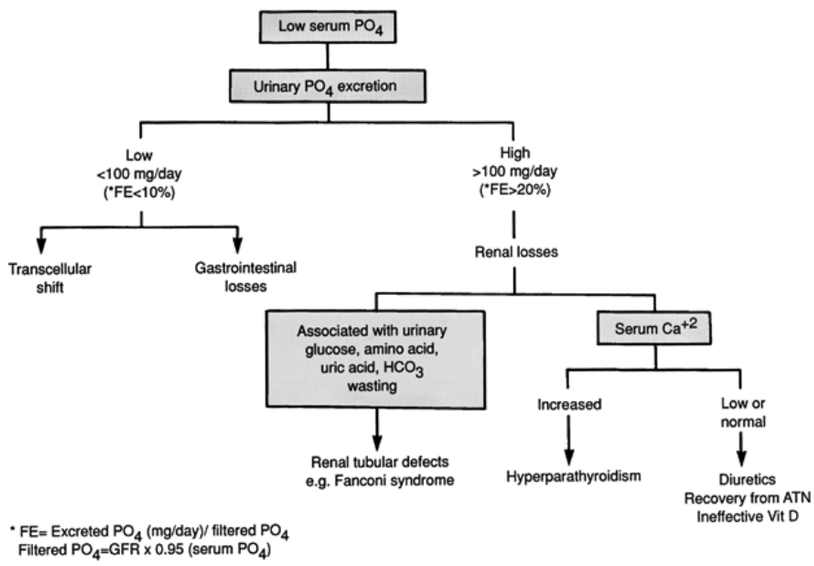


FIGURE 20.6. Evaluation of hypophosphatemia. FE, excreted PO_4 (mg/day/filtered PO_4); Filtered PO_4-GFR × 0.95 (serum PO_4). (From Chernow B. *The pharmacologic approach to the critically ill patient*. 2nd ed. Baltimore: Williams & Wilkins, 1988:629.)

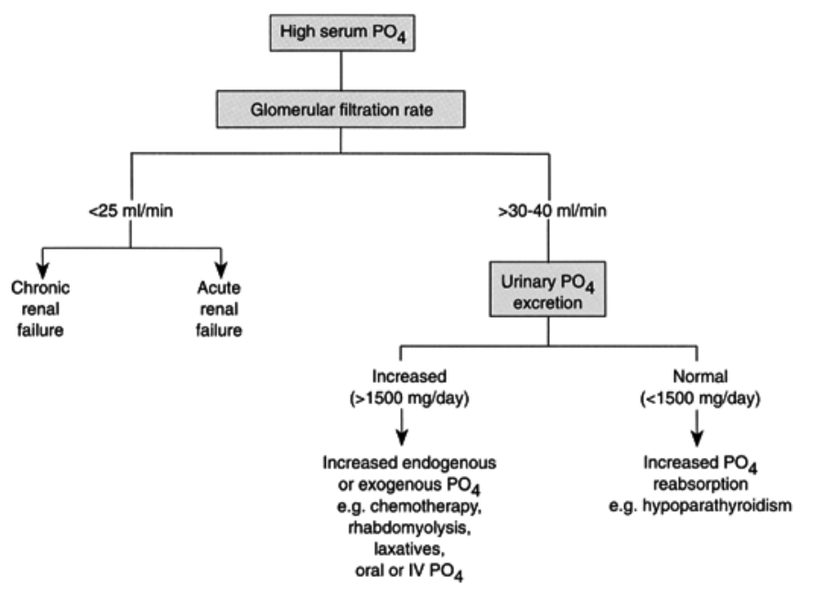


FIGURE 20.7. Evaluation of hyperphosphatemia. (From Chernow B. *The pharmacologic approach to the critically ill patient*. 2nd ed. Baltimore: Williams & Wilkins, 1988:633.)

HYPOPHOSPHATEMIA

Part of "20 - Calcium, Magnesium, and Phosphate"

The incidence of hypophosphatemia in hospitalized patients varies from 2% to 40%, depending on the population studied (89). Severe hypophosphatemia usually results in decreased concentrations of phosphate-containing compounds, such as ATP and membrane phospholipids. These deficiencies in turn are responsible for symptoms of hypophosphatemia, including muscle weakness, respiratory and myocardial insufficiency, and hepatocellular damage (89). Because phosphate is present in nearly all foods, a pure dietary deficiency rarely causes hypophosphatemia. More commonly, hypophosphatemia is caused by renal loss, gastrointestinal malabsorption, or transcellular shifts of phosphate. Mild hypophosphatemia is common in hospitalized patients and usually is not treated, but severe hypophosphatemia (less than 1.0 g/dL or 0.3 mmol/L) requires monitoring and possible replacement therapy. Some causes of hypophosphatemia in hospitalized patients are listed below, by the mechanism of their cause:

1. *Transcellular Shifts*. Because the movement of glucose into cells is accompanied by phosphate, the administration of glucose or insulin will lead to an influx of phosphate into cells that may cause hypophosphatemia. By a complex mechanism involving glycolysis and formation of phosphorylated intermediates in cells, respiratory alkalosis also substantially enhances phosphate uptake by cells (86, 89, 90). Urinary phosphate is usually decreased.
2. *GI losses*. As with many electrolyte disorders, diarrhea and vomiting can diminish GI absorption of phosphate. Because the divalent cations in antacids such as aluminum hydroxide, magnesium hydroxide, or aluminum carbonate bind phosphate and prevent its absorption in the gut, hypophosphatemia often results from chronic excessive use of such antacids (86). Urine phosphate is usually low.
3. *Renal losses*. The causes include: primary hyperparathyroidism, diuretics, hypomagnesemia, or defects in renal tubular absorption of phosphate. Urine phosphate is normal or elevated.

4. *Mixed causes* (86, 90).

   - In diabetic ketoacidosis, the combined effects of acidosis, glycosuria, ketonuria, and insulin therapy deplete phosphate, both by renal loss and cellular uptake. Serum concentrations may decrease rapidly.
   - Acidosis induces mobilization of phosphate from bone and tissues, along with increased renal loss.
   - Alcoholism leads to hypophosphatemia by increasing renal loss and decreasing GI absorption.

   During recovery from thermal burns, phosphate is shifted into cells and secreted into renal tubules, where it is lost in urine.

HYPERPHOSPHATEMIA

Part of "20 - Calcium, Magnesium, and Phosphate"

The most common cause of hyperphosphatemia is decreased phosphate excretion resulting from renal failure. In the early stages of a chronic loss of renal function, the remaining nephrons increase excretion of phosphate. Therefore, severe hyperphosphatemia does not occur until renal disease is well advanced and when the glomerular filtration rate is less than 25 mL/minute. A recent study concluded that a large percentage of longer-term renal dialysis patients have a serum phosphate above 2.1 mmol/L, and that this placed them at an increased risk of death (91). In acute renal failure serum phosphate may rise even more rapidly than urea or creatinine. Consequently, serum phosphate should be monitored in acutely ill patients (88, 89).

Other causes of hyperphosphatemia include increased intake of either phosphate or vitamin D and increased breakdown of cells by severe infections, intensive exercise, or neoplastic diseases. Immature lymphoblasts have about 4 times the phosphate content of mature lymphocytes, so patients with lymphoblastic leukemia may develop hyperphosphatemia. Neonates are especially prone to hyperphosphatemia caused by increased intake because they may not have mature PTH and vitamin D metabolism.

If renal function is normal, a urinary phosphate excretion above 1500 mg/day suggests hyperphosphatemia caused by chemotherapy, rhabdomyolysis, malignant hyperthermia, or excessive administration of phosphate by enema or intravenous route. A urinary phosphate excretion less than 1500 mg/day with normal renal function suggests some form of hypoparathyroidism (89).

Proper Collection and Handling of Samples

Hemolysis will falsely elevate serum phosphate concentrations, and serum should be removed or separated from cells promptly to avoid intracellular loss of phosphate into the plasma or serum phase. Phosphate concentrations will be slightly lower after meals, so the patient should be fasting before the sample is drawn to obtain an accurate baseline level. Intravenous administration of glucose or fructose lowers serum phosphate by promoting the intracellular shift of phosphate.

Methods for Measurement of Serum Phosphate

Since the 1920s, phosphate has been measured by methods in which molybdate reacts with phosphate to form complex molecules of phosphomolybdate. Fiske and Subbarow used a reducing reagent, 1-amino-2-naphthol-4-sulfonic acid (ANS), to form molybdenum blue, a complex polymer with a high molar absorptivity at 660 nm (92). While utilization of the basic reaction of phosphate with molybdate has remained essentially the same through the years, several different reducing agents have been developed. In addition to ANS, agents used have included stannous chloride (with and without hydrazine) and ferrous ammonium sulfate (93).

Simonsen showed that the unreduced phosphomolybdate complex has a high absorptivity at 340 nm, and Daly and Ertingshausen optimized this method for centrifugal analysis (93). Garber and Miller proposed a selected method for phosphate in serum that utilizes protein precipitation with semidine HCl as a reducing agent (94). A dry-slide method for phosphate determination utilizes p-methylaminophenol sulfate as reducing agent while monitoring light reflection at 680 nm (95). Several enzymatic methods for measuring phosphate have been developed (96) but have not been widely used in clinical laboratory practice. Reference ranges for phosphate are shown in Table 20.5.

TABLE 20.5. Reference Ranges for Phosphate[a]

	Serum Phosphate[b]	
	mgp/dL	**mmol/L**
Newborn (1–2 days)	5.5–9.5	1.78–3.07
Infant	4.5–6.5	1.45–2.10
Child	4.5–5.5	1.45–1.78
Adult male	2.3–3.7	0.74–1.20
Adult female	2.8–4.1	0.90–1.32

[a] Serum phosphate concentrations are relatively high at birth, then decline by 3 to 4 days, about the time the parathyroid gland becomes functional. Adults on an average diet should excrete 0.4–1.3 g of phosphorus per day.

[b] Expressed as phosphorus.

REFERENCES

1. Ringer, S. A further contribution regarding the influence of different constituents of blood on contractions of the heart. *J Physiol* 1883;4:29–42.

2. McLean FC, Hastings AB. A biological method for the estimation of calcium ion concentration. *J Biol Chem* 1934;107:337–350.

3. Drop LJ. Ionized calcium, the heart, and hemodynamic function. *Anesth Analg* 1985;64:432–451.

4. Zaloga GP, Chernow B. Calcium metabolism. In: Geelhoed GW, Chernow B, eds. *Endocrine aspects of acute illness*. New York: Churchill Livingstone, 1985:169–204.

5. Gray TA, Patterson CR. The clinical value of ionized calcium assays. *Ann Clin Biochem* 1988;25:210–219.

6. Toffaletti J, Nissenson R, Endres D, et al. Influence of continuous infusion of citrate on responses of immunoreactive PTH, calcium and magnesium components, and other electrolytes in hormal adults during plateletapheresis. *J Clin Endocrinol Metab* 1985;60:874–879.

7. Saha H, Harmoinen A, Pietila K, Morsky P, Pasternack A. Measurement of serum ionized versus total levels of magnesium and calcium in hemodialysis patients. *Clin Nephrol* 1996;46:326–31.

8. Ahren B, Bergenfelz A. Effects of minor increase in serum calcium on the immunoheterogeneity of PTH in healthy subjects and in patients with primary hyperparathyroidism. *Horm Res* 1995;43:294–299.

9. Mundy GR, Guise TA. Hormonal control of calcium homeostases. *Clin Chem* 1999;45:1347–52.

10. Felsenfeld AJ, Machado L, Rodriguez M. The relationship between serum calcitonin and calcium in the hemodialysis patient. *Am J Kid Dis* 1993;21:292–299.

11. Toffaletti J, Gitelman HJ, Savory J. Separation and quantitation of serum constituents associated with calcium by gel filtration. *Clin Chem* 1976;22:1968–1972.

12. Fogh-Andersen N. Ionized calcium analyzer with a built-in pH correction. *Clin Chem* 1981;27:1264—1267.

13. Ladenson JM, Lewis JW, Boyd JC. Failure of total calcium corrected for protein, albumin, and pH to correctly assess free calcium status. *J Clin Endocrinol Metab* 1978;46:986–993.

14. Need AG, Nordin BEC. Rediscovering ionised calcium. *Aust NZ J Med* 1998;28:155–157.

15. Burritt MF, Pierides AM, Offord KP. Comparative studies of total and ionized serum calcium values in normal subjects and patients with renal disorders. *Mayo Clin Proc* 1980;55:606–613.

16. Leunissen KML, van den Berg BW, van Hoof JP. Ionized calcium plays a pivotal role in controlling blood pressure during haemodialysis. *Blood Purif* 1989;7:233–239.

17. Wandrup J. Critical analytical and clinical aspects of ionized calcium in neonates. *Clin Chem* 1989;35:2027–2033.

18. Sheth DP. Hypocalcemic seizures in neonates. *Am J Emerg Med* 1997;15:638–641.

19. Desai TK, Geheb MA, Haupt MT, et al. Hypocalcemia in critically ill patients [Abstract]. *Chest* 1984;86:282.

20. Vadstrup S, Pedersen TE, Weywadt L, Wandrup J. Correlation between severity of septic conditions and circulating levels of ionized calcium. *Intensive Care Med* 1989;15:329–330.

21. Vincent J-L, Jankowski S. Why should ionized calcium be determined in acutely ill patients? *Act Anaesthesiol Scand* 1995;39:Suppl 107, 281–286.

22. Desai TK, Carlson RW, Thill-Baharozian M, Geheb MA. A direct relationship between ionized calcium and arterial pressure among patients in an intensive care unit. *Crit Care Med* 1988;16:578–582.

23. Anast CS, Winnacker JL, Forte LF, Burns TW. Impaired release of parathyroid hormone in magnesium deficiency. *J Clin Endocrinol Metab* 1976;42:707–717.

24. Medalle R, Waterhouse C, Hahn TJ. Vitamin D resistance in magnesium deficiency. *Am J Clin Nutrition* 1976;29:854–858.

25. Ladenson JH, Lewis JW, McDonald JM, Slatopolsky E, Boyd JC. Relationship of free and total calcium in hypercalcemic conditions. *J Clin Endocrinol Metab* 1979;48:393–397.

26. Glendenning P, Gutteridge DH, Retallack RW, Stuckey BGA, Kermode DG, Kent GN. High prevalence of normal total calcium and intact PTH in 60 patients with proven primary hyperparathyroidism: a challenge to current diagnostic criteria. *Aust NZ J Med* 1998;28:173–178.

27. Fleetwood MK, Quinton L, Wolfe J, Maturani D, et al. Rapid PTH assay by simple modifications of Nichols Intact PTH-Parathyroid Hormone assay kit. *Clin Chem* 1996;42:1498.

28. Pandian MR, Morgan CH, Carlton E, Segre GV. Modified immunoradiometric assay of parathyroid hormone-related protein: Clinical application in the differential diagnosis of hypercalcemia. *Clin Chem* 1992;38:282–8.

29. Boyd JC, Lewis JW, Slatopolsky E, Ladenson JH. Parathyrin measured concurrently with free or total calcium in the differential diagnosis of hypercalcemia. *Clin Chem* 1981;27:574–579.

30. Renoe BW, McDonald JM, Ladenson JH. The effects of stasis with and without exercise on free calcium, various cations, and related parameters. *Clin Chem Acta* 1980;103:91–100.

31. Toffaletti J, Abrams B. Effects of in vivo and in vitro production of lactic acid on ionized, protein-bound, and complex-bound calcium in blood. *Clin Chem* 1989;35:935–938.

32. Graham G, Schoen I, Johnson L. The effect of specimen choice, collection, processing, and storage on ionized calcium determinations. In: Moran RF, ed. *Ionized calcium: its determination and clinical usefulness.* Galveston, TX: MVI Publishing, 1986:88–92.

33. Toffaletti J, Blosser N, Kirvan K. Effects of storage temperature and time before centrifugation on ionized calcium in blood collected in plain vacutainers and SST tubes. *Clin Chem* 1984;30:553–556.

34. Moorehead WR, Biggs HG. 2-Amino-2-methyl-1-propanol as the alkalizing agent in an improved continuous-flow cresolphthalein complexone procedure for calcium in serum. *Clin Chem* 1974;20:1458–1460.

35. Toffaletti J. Ionized calcium. In: Pesce AJ, Kaplan LA, eds. *Methods in clinical chemistry.* St. Louis: CV Mosby, 1987:1010–1020.

36. Gambling DR, Birmingham CL, Jenkins, LC. Magnesium and the anaesthetist. *Can J Anaesth* 1988;35:644–654.

37. Altura BM, Altura BT. New perspectives on the role of magnesium in the pathophysiology of the cardiovascular system. I. Clinical aspects. *Magnesium* 1985;4:226–244.

38. Iseri LT, French JH. Magnesium: nature's physiologic calcium blocker [Editorial]. *Am Heart J* 1984;108:188–193.

39. Gums JG. Clinical significance of magnesium: a review. *Drug Intell Clin Pharm* 1987;21:240–246.

40. Turner ET, Osborn JJ, et al. Magnesium and open-heart surgery. *J Thorac Cardiovasc Surg* 1972;645:695–704.

41. Chernow B, Bamberger S, et al. Hypomagnesemia in patients in postoperative intensive care. *Chest* 1989;95:391–397.

42. Zaloga GP, Chernow B. Divalent ions: calcium, magnesium, and phosphorus. In: Chernow B, ed. *The pharmacologic approach to the critically ill patient*, 2nd ed. Baltimore: Williams & Wilkins, 1988:62–627.

43. Toffaletti J, Cooper D, Lobaugh B. The response of parathyroid hormone to specific changes in either ionized calcium, ionized magnesium protein-bound calcium, in humans. *Metabolism* 1991;8:814–818.

44. Elin RJ. Assessment of magnesium status. *Clin Chem* 1987;33:1965–1970.

45. Shils ME. Magnesium in health and disease. *Annu Rev Nutrition* 1988;8:429–460.

46. Berkelhammer C, Bear RA. A clinical approach to common electrolyte problems:hypomagnesemia. *Can Med Assoc J* 1985;132:360–368.

47. Brautbar N, Massry SG. Hypomagnesemia and hypermagnesemia. In: Maxwell MH, Kleeman CR, Narins RG, eds. *Clinical disorders of fluid and electrolyte metabolism*. 4th ed. New York: McGraw Hill, 1987:831–849.

48. Iseri LT, Allen BJ, Brodsky MA. Mg therapy of cardiac arrythmias in critical care medicine. *Magnesium* 1989;8:299–306.

49. White JR, Jr, Campbell RK. Magnesium and diabetes: A review. *Ann Pharmacother* 1993;27:775–780.

50. Salem M, Munoz R, Chernow B. Hypomagnesemia in critical illness. *Crit Care Clin* 1991;7:225–252.

51. Ryzen E. Magnesium homeostasis in critically ill patients. *Magnesium* 1989;8:201–212.

52. Rubeiz GJ, Thill-Baharozian M, Hardie D, Carlson RW. Association of hypomagnesemia and mortality in acutely ill medical patients. *Crit Care Med* 1993;21:203–209.

53. Lum G. Hypomagnesemia in acute and chronic care patient populations. *Am J Clin Pathol* 1992;97:827–830.

54. Broner CW, Stidham GL. Westenkirchner DF, et al. Hypermagnesemia and hypocalcemia as predictors of high mortality in critically ill pediatric patients. *Crit Care Med* 1990;18:921–928.

55. Fiser RT, Torres A, Butch AW, Valentine JL. Ionized magnesium concentrations in critically ill children. *Crit Care Med* 1998;26:2048–2052.

56. Haigney MCP, Berger R, Schulman S, Gerstenblith G, et al. Tissue magnesium levels and the arrhythmic substrate in humans. *Electrophysiol* 1997;8:980–986.

57. Hanline M. Hypomagnesemia causes coronary artery spasm. *JAMA* 1985;253:342.

58. Sheehan JP, Seelig MS. Interactions of Mg and K in the pathogenesis of cardiovascular disease. *Magnesium* 1984;3:301–314.

59. Altura BM, Altura BT. Biochemistry and pathophysiology of CHF: is there a role for Mg? *Magnesium* 1986;5:134–143.

60. Rude R, Manoogian C, Ehrlich L, et al. Mechanism of BP regulation by Mg in man. *Magnesium* 1989;8:266–273.

61. Herzog WR. The use of magnesium in the treatment of patients with acute myocardial infarction. *Coronary Artery Dis* 1996;7:339–340.

62. Seelig MS, Elin RJ, Antman EM. Magnesium in acute myocardial infarction: Still an open question. *Can J Cardiol* 1998;14:745–749.

63. Pearson PJ, Evora PRB, Seccombe JF, Schaff HZ. Hypomagnesemia inhibits nitric oxide release from coronary endothelium: protective role of magnesium infusion after cardiac operations. *Ann Thorac Surg* 1998;65:967–972.

64. Brookes CIO, Fry CH. Ionised magnesium and calcium in plasma from healthy volunteers and patients undergoing cardiopulmonary bypass. *Br Heart J* 1993;69:404–409.

65. Ryan MP. Diuretics and potassium/magnesium depletion: directions for treatment. *Ann J Med* 1987;82(suppl 3A):38–47.

66. Whang R, Oei To, Watanabe A. Frequency of hypomagnesium in hospitalized patients receiving digitalis. *Arch Int Med* 1985;145;655–656.

67. June CH, Thompson CB, Kennedy MS. Profound hypomagnesemia and renal magnesium wasting associated with the use of cyclosporine for marrow transplantation. *Transplantation* 1985;39:620–624.

68. Durlach J, Collery P. Magnesium and potassium in diabetes and carbohydrate metabolism: review of the present status and recent results. *Magnesium* 1984;3:315–323.

69. American Diabetes Association. Magnesium supplementation in the treatment of diabetes. *Diabetes Care* 1992;15:1065–1067.

70. Flink EB. Magnesium deficiency in alcoholism. *Alcohol Clin Exp Res* 1986;10:590–594.

71. Hristova EN, Rehak NN, Cecco S, Ruddel M, Herion D, Eckardt M, Linnoila M, Elin RJ. Serum ionized magnesium in chronic alcoholism: is it really decreased? *Clin Chem* 1997;43:394–399.

72. Agus ZS, Massry SG. *Hypomagnesemia and hypermagnesemia, in Maxwell and Kleeman's clinical disorders of fluid and electrolyte metabolism, 5th ed.* 1994:1104.

73. Revusova V, Zvara U, Karlikova L, Suchane KB, et al. Prognosis of urolithiasis and nephrocalcinosis in hypomagnesemia. *Czech Med* 1985;8:207–213.

74. Levine J, Rapoport A, Mashiah M, Dolev E. Serum and CSF levels of magnesium in acute versus remitted schizophrenic patients. *Neuropsychobiol* 1996;33:169–172.

75. Lampl Y, Geva D, Gilad R, Eshel Y, Ronen L, Sarova-Pinhas I. CSF magnesium level as a prognositc factor in ischaemic stroke. *J Neurol* 1998;245:584–588.

76. Taylor WM. Low serum magnesium concentration in Paget's disease of bone (osteitis deformans). *Ann Clin Biochem* 1985;22:591–595.

77. Wills MR, Sunderman FW, Savory J. Methods for estimation of serum magnesium in clinical laboratories. *Magnesium* 1986;5:317–327.

78. Toffaletti J, Abrams B, Bird C, Schwing M. Clinical validation of an automated thin-film reflectance method for measurement of magnesium in serum and urine. *Magnesium* 1988;7:84–90.

79. Farrell EC Jr. Magnesium. In: Pesce AJ, Kaplan LA, eds. *Methods in clinical chemistry*. St. Louis: CV Mosby, 1987:1021–1026.

80. Elin RJ. Magnesium: The fifth but forgotten electrolyte. *Am J Clin Pathol* 1994;102:616–622.

81. Cecco SA, Hristova EN, Rehak NN, Elin RJ. Clinically important intermethod differences for physiologically abnormal ionized magnesium results. *AJCP* 1997;108:564–569.

82. D'Costa M, Cheng PT. Ultrafiltrable calcium and magnesium in ultrafiltrates of serum prepared with the Amicon MPS-1 System. *Clin Chem* 1983;29:519–522.

83. Hristova EN, Cecco S, Niemela JE, Rehak NN, Elin RJ. Analyzer-dependent differences in results for ionized calcium, ionized magnesium, sodium, and pH. *Clin Chem* 1995;41:1649–1659.

84. Westheimer FM. Why nature chose phosphates. *Science* 1987;235:1173–1178.

85. Ditzel J. Effect of plasma inorganic phosphate on tissue oxygenation during recovery from diabetic keotacidosis. *Adv Exp Med Biol* 1973;37A:163–172.

86. King AL, Sica DA, Miller G, Pierpaoli S. Severe hypophosphatemia in a general hospital population. *South Med J* 1987;80:831–835.

87. Kates DM, Sherrard DJ, Andress DL. Evidence that serum phosphate is independently associated with serum PTH in patients with chronic renal failure. *Am J Kid Diseases* 1997;30:809–13.

88. Slatopolsky E, Rutherford WE, Rosenbaum R, et al. Hyperphosphatemia. *Clin Nephrol* 1977;7:138–146.

89. Chester WL, Zaloga GP, Chernow B. Phosphate problems in the critically-ill patient. In: Geelhoed GW, Chernow B, eds. *Endocrine aspects of acute illness*. New York: Churchill Livingstone, 1985:205–216.

90. Knochel JP. Hypophosphatemia. *Clin Nephrol* 1977;7:131–137.

91. Block GA, Hulbert-Shearon TE, Levin NW, Port FK. Association of serum phosphate and Ca X PO4 product with mortality risk in chronic hemodialysis patients: A national study. *Am J Kid Dis* 1998;31:607–617.

92. Fiske CH, Subbarrow Y. The colorimetric determination of phosphorus. *J Biol Chem* 1925;66:375–400.

93. Farrell EC Jr. Phosphorus. In: Pesce AJ, Kaplan LA, eds. *Methods in clinical chemistry*. St. Louis: CV Mosby, 1987:1038–1042.

94. Garber CC, Miller RC. Revisions of the 1963 semidine HC1 standard method for inorganic phosphorus. *Clin Chem* 1983;29:184–188.

95. Toffaletti J. Automated dry-film method for phosphorus in serum evaluated. *Clin Chem* 1985;31:148–149.

96. Adam A, Boulanger J, Azzouzi M, Ers P. Colorimetric versus enzymatic determination of serum phosphorus. *Clin Chem* 1984;30:1724–1725.

21

Heme Synthesis and Catabolism

György Ábel

Darryl Palmer-Toy

OVERVIEW

Part of "21 - Heme Synthesis and Catabolism"

Scope

Heme is a versatile chemical prosthetic group involved in various essential functions including respiration, detoxification, antimicrobial defense, and second messenger synthesis. Many of the colors of clinical specimens reflect products of heme chemistry: blood, red (heme); urine, yellow (urochrome); bile, yellow-green (bilirubin); and feces, brown (urobilin and mesobilifuscins). Inherited disorders of heme synthesis, collectively known as the porphyrias, are relatively rare. Yet porphyrias often enter the differential diagnosis of various symptoms as a result of the protean manifestations of these disorders. In contrast, inherited or acquired abnormalities of bilirubin excretion frequently are encountered in routine medical practice. This chapter will review disorders of heme synthesis and catabolism, and the role of laboratory tests in the diagnosis of these conditions. Where possible, the molecular basis for the more common biosynthetic defects, associated clinical syndromes, and relevant laboratory assays will be described including diagnostic strategies for the porphyrias and for hyperbilirubinemia. The topic of hemoglobinopathy is discussed elsewhere (Chapter 41).

Heme Biosynthesis

The primary sites of heme synthesis are the bone marrow (about 85% of the total production), and the liver. Figure 21.1 shows the biochemical pathway of heme synthesis, the enzymes involved in the production of each metabolite and the inherited or acquired enzyme deficiencies involved with the particular porphyrias. The first reaction in the synthetic pathway is the condensation of succinyl CoA with glycine catalyzed by the mitochondrial enzyme d-aminolevulinic acid (ALA) synthase, the rate-limiting enzyme of the synthetic pathway. Because this first step of the sequence is rate-limiting, the product ALA and the subsequent intermediary metabolites normally are found only in low, steady-state concentrations in the cells. The concentration of these metabolites in body fluids also is low. Following its synthesis in the mitochondrion, ALA transfers to the cytoplasm. The next step in the pathway is the condensation of two ALA molecules, mediated by ALA dehydratase (porphobilinogen synthase), to form porphobilinogen (PBG). Four PBG molecules then are condensed into hydroxybilane by PBG deaminase (uroporphyrinogen I synthase). Hydroxymethilbilane then is converted to uroporphyrinogen III by uroporphyrinogen III cosynthase. Alternatively, hydroxymethylbilane may convert to uroporphyrinogen I, then to coproporphyrinogen I, using an alternative pathway. Uroporphyrinogen III is converted in the cytosol to coproporphyrinogen III by the fifth enzyme in the pathway, uroporphyrinogen III decarboxylase. This compound then reenters the mitochondrion, where two of the propionic acid groups are decarboxylated and by the enzyme coproporpyrinogen oxidase, resulting in the next metabolite, protoporphyrinogen IX. In the next step, protoporphyrinogen oxidase oxidizes protoporphyrinogen IX by the removal of six hydrogen atoms. The reaction produces porphyrins, which are the oxidized forms of porphyrinogens. In the final step, a Fe^{2+} is incorporated into protoporphyrin IX to form the final product, heme. In the bone marrow, ALA synthase is expressed at higher levels than in the liver where it is repressed by free heme. The enzymes responsible for heme biosynthesis in the bone marrow are regulated in a coordinated fashion. In the liver, ALA synthase is inducible by the same agents that induce the cytochrome P450 enzymes. Because most of the hepatic heme is used in the synthesis of cytochrome P450, the regulation of these enzymes and ALA synthase are coordinated in the liver.

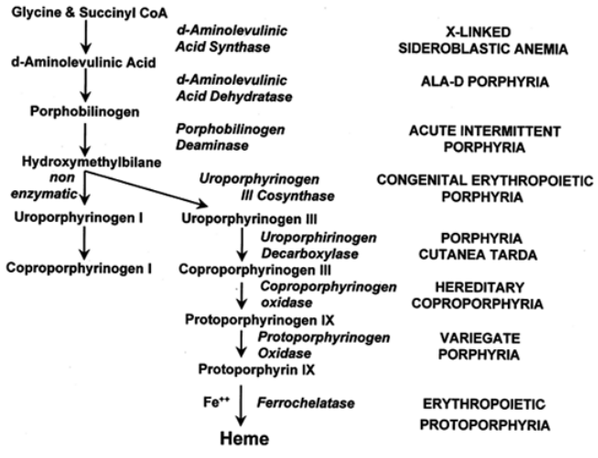


FIGURE 21.1. Heme synthesis: metabolites, enzymes, and the disorders associated with enzyme deficiencies.

THE PORPHYRIAS

Part of "21 - Heme Synthesis and Catabolism"

Porphyrias are diseases resulting from disorders of heme synthesis. The clinical symptoms of different porphyrias significantly overlap, therefore the clinical presentation cannot be used alone to identify the specific disorder. Therefore, the pattern of measured porphyrins and their precursors found in urine, plasma, red blood cells, and feces are used for diagnosis of the particular type of porphyria. Measuring the activity of the enzyme suspected to be deficient and molecular genetic tests are useful adjuncts. The disorders related to the deficiency of each participating enzyme are listed in Figure 21.1. Porphyrias have been classified according to the major clinical manifestations as erythropoietic, hepatic, and hepatoerythropoietic porphyrias (Table 21.1). An alternative classification system divides porphyrias into those with cutaneous or neurological manifestations. Photosensitivity is a key clinical symptom in erythropoietic porphyria, erythropoietic protoporphyria, porphyria cutanea tarda, variegate porphyria, and hepatoerythropoietic

porphyria. Acute intermittent porphyria, hereditary coproporphyria, and variegate porphyria are classified as acute porphyrias because they present with similar abdominal, autonomic, and neuropsychiatric features; whereas porphyria cutanea tarda, erythropoietic protoporphyria, and congenital erythropoietic porphyria are characteristically associated with photosensitivity of the skin, and are classified as “nonacute porphyries.” ALA and PBG, the key metabolites accumulating in acute intermittent porphyria and ALA dehydratase deficiency porphyrias do not produce photosensitization. Consequently, there are no cutaneous symptoms in these disorders (Table 21.2). Most chelated porphyrins, such as heme, are not photosensitizing. In addition, porphyrins chelated to Mg^{2+}, Zn^{2+}, and Cu^{2+} do not show fluorescence and also are not photosensitizing.

TABLE 21.1. CLASSIFICATION OF PORPHYRIAS

ERYTHROPOIETIC
Erythropoietic Porphyria
Erythropoietic Protoporphyria
Erythropoietic Coproporphyria
HEPATIC
Acute Intermittent Porphyria
Variegate Porphyria
Hereditary Coproporphyria
Aminolevulinic Acid Dehydratase Porphyria
Porphyria Cutanea Tarda
HEPATOERYTHROPOIETIC
Hepatoerythropoietic Porphyria
PORPHYRIAS WITH CUTANEOUS MANIFESTATIONS
Erythropoietic Porphyria
Erythropoietic Protoporphyria
Porphyria Cutanea Tarda
Variegate Porphyria
Hepatoerythropoietic Porphyria
PORPHYRIAS WITHOUT CUTANEOUS MANIFESTATIONS
(Aminolevulinic acid and porphobilinogen are not photosensitizing)
Acute Intermittent Porphyria
ALA Dehydratase Deficiency

TABLE 21.2. AMINOLEVULINIC ACID DEHYDRATASE DEFICIENCY PORPHYRIA

Epidemiology:
Very rare, only few patients reported
Autosomal recessive, cDNA showed double heterozygousity was found in cDNA for two independent alleles
Pathophysiology:
Aminolevulinic acid dehydratase deficiency
Clinical presentation:
Symptoms similar to acute intermittent porphyria, no skin lesions
Precipitating factors: stress, ethanol
Laboratory findings:
Erythrocytes: Aminolevulinic acid dehydratase <5% of normal;protoporphyrin elevated
Urine: Aminolevulinic acid, coproporphyrin, uroporphyrin elevated
Differential diagnosis:
Acute intermittent porphyria

ALA Dehydratase Deficiency Porphyria (Porphobilinogen Synthase Deficiency)

ALA synthase is the first and rate-limiting enzyme in the heme biosynthetic pathway. Although this enzyme is not associated with a specific porphyria, the activity of this enzyme is increased in all hepatic porphyrias, leading to an increased excretion of ALA and PBG in the urine.

Homozygous genetic mutations resulting in defects of the second enzyme, ALA dehydratase (ALA-D) manifest in ALA-D

deficiency porphyria, also called “plumboporphyria,” because of the similarity of the finding in this disorder to those observed in lead poisoning. ALA-D deficiency porphyria is a rare disorder with the clinical features of acute porphyrias (Table 21.3). The key laboratory finding is a massive elevation of ALA in the urine with normal urinary PBG excretion. Additional supporting laboratory findings for ALA-D deficiency porphyria are the increased coproporphyrin III in the urine. Laboratory tests for ALA-D deficiency porphyria include urinary ALA, PBG, coproporphyrin isomers, and ALA-D activity in red blood cells.

TABLE 21.3. ACUTE INTERMITTENT PORPHYRIA

Epidemiology:
Autosomal dominant, rarely manifests before puberty, female predominance
Incidence 1.5:100,000, in Lapland 1:1,000
Pathophysiology:
Deficiency of porphobilinogen deaminase
Clinical presentation:
Acute attacks with neurovisceral symptoms, abdominal pain
Skin photosensitivity: absent
Constipation, nausea, muscle weakness – paralysis, bizarre psychotic behavior
Often precipitated by drugs, hormones
Laboratory findings:
Rapid lab tests: Urine darkens to deep red/brown color when exposed to sunlight
Hoesch test or Watson-Schwartz test of urine for porphobilinogen
Urine: Elevated aminolevulinic acid, porphobilinogen, uroporphyrin I
Plasma: Increased porphyrins during attacks
Erythrocytes: Decreased porphobilinogen-deaminase activity, normal porphyrin levels
Feces: Normal or slightly elevated porphyrins
Differential diagnoses:
Variable, may resemble many different diseases, called “little imitator” by Waldenström

Acute Intermittent Porphyria (Porphobilinogen Deaminase Deficiency, Uroporphyrinogen Synthase Deficiency, Hydroxymethylbilane Synthase Deficiency Porphyria, Swedish-Type Porphyria)

The three acute porphyrias are acute intermittent porphyria (AIP), variegate porphyria (VP), and hereditary coproporphyria (HCP), all inherited in an autosomal-dominant manner. Clinically, all of these three disorders present after puberty. Acute intermittent porphyria is a result of four distinct mutants of PBG deaminase. In the most common genotype of AIP, patients have about one half of the normal PBG deaminase. With this genotype, decreased concentrations of PBG deaminase may be found both in patients with clinical disease and in latent carriers of this genetic defect. Most individuals with decreased PBG deaminase activity as a result of this gene defect never develop porphyria. However, they may be at risk if they are exposed to triggering factors, a “second hit.” Attacks are frequently precipitated by a variety of drugs and hormones. Acute attacks of AIP are characterized by abdominal pain, the hallmark symptom of this disorder, constipation, nausea, muscle weakness, paralysis, and bizarre psychotic behavior. Muscle weakness may be severe during a prolonged attack. It is because of the effects of the accumulated metabolites on the peripheral nerves, the nerves in the abdomen, and the central nervous system. The exact mechanism of the symptoms is not well understood. Attacks are more prevalent in women than in men and more often occur during the second half of the menstrual cycle. Attacks can be triggered by oral contraceptives containing synthetic estrogens. Alcohol and various environmental chemicals also are found to have this attack-triggering, porphyrogenic potential. A partial list of porphyrinogenic and “safe” drugs and agents is shown in Table 21.4.

TABLE 21.4. THE “SECOND HIT”: EXAMPLES OF UNSAFE AND SAFE AGENTS AND DRUGS

Unsafe	Safe
Ethanol	Acetaminophen
Anti-epilepsy drugs	Aspirin
Barbiturates	Atropine
Birth control pills	Bromides
Calcium-ch. Blockers	Chloral hydrate
Carbamazepine	Penicillin
Clonazepam	Phenotiazines
Danazol	Narcotic analgesics
Diclofenac	Cimetidine
Ergots	Insulin
Rifampine	Glucocorticoids
Sulfonamides	Streptomycin

The principal laboratory tests in AIP are measurements of ALA and PBG in urine. During acute episodes, the urine may turn dark red/brown from the increased excretion of ALA, PBG, and porphyrins. Excessive PBG in urine is nonenzymatically converted to uroporphyrin I. This process may be slowed by refrigeration and alkalinization of the urine specimen to pH 8–9. The excretion of ALA and PBG usually decreases with clinical improvement. A normal urinary PBG level effectively excludes AIP as a cause for the current symptoms. Most asymptomatic, “latent” heterozygous individuals with PNG deaminase deficiency have normal urinary excretion of ALA and PBG. Therefore, measurement of PBG deaminase in erythrocytes is useful to confirm the diagnosis and screen asymptomatic family members.

Because quantitative tests for ALA and PBG most often are performed in reference laboratories, many hospitals use qualitative screening test methods for PBG. In the Hoesch test, urine is tested with p-dimethylaminobenzaldehyde (Ehrlich's reagent). The specificity of this simple test has been improved in the Watson-Schwartz test with the use of chloroform and n-butanol to remove interfering substances, primarily urobilinogen. Specificity also can be improved by the application of ion-exchange column to remove interfering compounds prior to performing the Ehrlich reaction. Although empirical evidence suggests that the Watson-Schwartz test is suitable for the detection of AIP, the quantitative analysis of a timed urine specimen also is necessary for the definitive diagnosis.

The deficiency of PBG deaminase is detectable in the erythrocytes of most heterozygotes. However, the activity of this enzyme is higher in younger erythrocytes and may increase into the normal range in AIP when erythropoiesis is increased. Thus, a normal result does not necessarily exclude the diagnosis of AIP.

The management of AIP involves high carbohydrate diet, IV glucose and hematin infusions, and the avoidance of the precipitating drugs.

Congenital Erythropoietic Porphyria (Congenital Erythropoietic Porphyria, Congenital Hematoporphyria, Erythropoietic Uroporphyria, Uroporphyrinogen III Cosynthase Deficiency, Hydroxymethylbilane Hydrolase Deficiency, Günther's Disease)

Erythropoietic porphyria (EP) is extremely rare, less than 200 cases have been described (Table 21.5). It is inherited in an autosomal-recessive manner. The enzymatic basis for this disease is the deficiency of uroporphyrinogen III cosynthase. This enzyme catalyzes the synthesis of the asymmetric uroporphyrinogen III molecule. The deficiency of this enzyme leads to the overproduction of uroporphyrinogen I and coproporphyrinogen I, which have no role in heme synthesis. Oxidation of these metabolites to their respective porphyrins, leads to accumulation and excretion of uroporphyrin I (URO I) and coproporphyrin I (COPRO I), beginning during fetal life. The disorder usually is discovered soon after birth because patients excrete unusual red-pigmented urine. Intravascular hemolysis also may occur. Urinary excretion of URO I and COPRO I is extremely high, maybe over 100 times the normal level. Urinary excretion of ALA and PBG is normal. Fecal coproporphyrin also is elevated, with COPRO I as the dominant isomer.

TABLE 21.5. ERYTHROPOIETIC PORPHYRIA

Epidemiology:

Autosomal recessive, very rare (<200 reported cases)

Pathophysiology:

Uroporphyrinogen III cosynthase deficiency

Clinical presentation:

Hemolytic anemia, erythrodontia, splenomegaly, pink-red urine, fluorescent erythrocytes and normoblasts, ulcerations of the eyes

Photosensitivity: marked, early in childhood

Skin reactions: vesicles, bullae, erosions, hypertrichosis, thickened eyebrows and eyelashes, hypermelanosis, skin photosensitivity, later scarring with atrophy, mutilating deformities of hands, ears, face and nose, cicatrizing alopecia

Laboratory findings: Urine: Uroporphyrin I and coproporphyrin I elevated; URO > COPRO

Feces: Coproporphyrin I elevated

Erythrocytes and plasma: Uroporphyrin I and/or zinc protoporphyrin elevated

Plasma: Uroporphyrin I and coproporphyrin I elevated

Enzyme: Uroporphyrinogen III cosynthase decreased, particularly in bone marrow

Differential diagnoses:

Other congenital photodermatoses, hepatic porphyrias

The management of EP includes splenectomy, which may improve hemolysis, the avoidance of exposure to sunlight, and the treatment of secondary skin infections.

Porphyria Cutanea Tarda [Uroporphyrinogen Decarboxylase (UROD) Deficiency Type I: 'Sporadic' Type, and Types II and III: 'Familial' Type Symptomatic Porphyria, Acquired Hepatic Porphyria, Chemical Porphyria]

Porphyria cutanea tarda (PCT) is the most common porphyria, caused by the deficiency of the enzyme uroporphyrinogen decarboxylase (UROD). PCT can be acquired, the onset is often later during life, it is most common in elderly men. The disorder almost occurs in a sporadic manner but cases with familial inheritance also have been described. PCT often is found in association with liver disease, hepatitis C virus infection has been implicated in these disorders in recent years. Alcohol abuse also is among the most common causes. Other diseases, such as uremia, diabetes mellitus, and certain autoimmune disorders also have been found associated with this disorder. For a summary of PCT, see Table 21.6. The activity of UROD is decreased to about half of the normal value. In most cases, this reduced level of activity, however, is sufficient for carrying out the process of heme synthesis. In other cases, the insufficient enzyme activity results in a decreased synthesis of heme and accumulation of the early metabolites of the synthetic pathway. Iron overload and excessive hepatic iron stores often are present. Patients in both sexes with PCT develop cutaneous lesions in areas of the skin exposed to sunlight: the face, dorsa of hands, and forearms, while the legs and feet typically are affected in women.

TABLE 21.6. PORPHYRIA CUTANEA TARDA

Epidemiology:

Essentially an acquired disease, genetic type (mainly autosomal dominant) is also known, the most common porphyria, male predominance

Age of onset is usually third or fourth decade

Hepatitis C virus (HCV) infection and alcohol abuse have been implicated

Pathophysiology: Uroporphyrinogen decarboxylase deficiency

Clinical presentation:

Photosensitivity: Marked, early in childhood, moderate to severe

Skin reactions: Bullae, vesicles, ulcers on light-exposed skin (e.g., on dorsa of hands and feet), increased skin fragility, hyperpigmentation, scarring alopecia

Diabetes mellitus in 25% of patients

Increased liver iron stores and serum iron

Urine may show pink fluorescence

Acute attacks are absent (unlike in variegate porphyria)

Laboratory findings

Urine: Increased uroporphyrin (I > III), coproporphyrin, and 7- carboxylporphyrin III,

Feces: Increased isocoproporphyrins

Erythrocytes: Normal porphyrins

Plasma: Increased 8- and 7-carboxylporphyrins

Enzyme defect: Deficient uroporphyrinogen decarboxylase in liver (type I) and in erythrocytes (type II)

Differential diagnoses: Variegate porphyria, epidermolysis bullosa acquisita

In the acquired form of PCT (type I), the deficiency of UROD occurs both in the hepatocytes and erythrocytes. In the acquired form of the disorder (type II), the enzyme defect is limited to the liver. All forms of PCT manifest with cutaneous symptoms and similar porphyrin profile in urine, feces, and plasma. In PCT, the plasma fluorescence scan exhibits a characteristic porphyrin spectrum that can be described by excessive uroporphyrin and higher carboxylated porphyrins, predominantly 7- and 8-carboxyl porphyrins. Erythrocyte porphyrins are normal, while urinary porphyrins show increased uroporphyrin and heptacarboxyl porphyrin levels.

The possible role of the analysis of UROD activity in PCT is unclear. It generally is accepted that patients with familial PCT have reduced levels of this enzyme in tissues, cultures skin fibroblasts, and erythrocytes. However, most cases are sporadic,

and in the sporadic cases erythrocyte enzyme activities often are normal. Laboratory diagnostic assessment for PCT should include qualitative and quantitative plasma porphyrin analysis, urinary, fecal, and erythrocyte porphyrin measurements.

PCT is considered the most treatable porphyria, its management consists of the avoidance of exposure to direct sunlight, avoidance of ethanol, estrogens, and other precipitating drugs, chemicals such as polycarbonated hydrocarbons, HCl-benzene, and dioxin. Removal of one unit of blood with phlebotomy weekly or biweekly reduces hemoglobin and iron levels. Chloroquine treatment is an alternative option.

Hepatoerythropoietic Porphyria (Hepatoerythrocytic Porphyria)

Hepatoerythropoietic porphyria is a rare disorder associated with a profound deficiency of uroporphyrinogen decarboxylase (Table 21.7). Clinically, the skin manifestations resemble porphyria cutanea tarda and erythropoietic porphyria. The disease typically manifests in childhood, with dark urine and cutaneous symptoms. Plasma of patients with this disorder contain excessive amounts of uroporphyrin, and red blood cells contain excess zinc protoporphyrin. Isocoproporphyrin is the predominant coproporphyrin isomer present in urine and stool. In contrast to porphyria cutanea tarda patients, serum iron is normal in hepatoerythropoietic porphyria. Key laboratory tests are erythrocyte zinc protoporphyrin, urinary porphyrin profile, and urinary coproporphyrin isomers. Uroporphyrinogen decarboxylase activity usually is 5% to 10% of normal. The management of hepatoerythropoietic porphyria includes the avoidance of sunlight, and oral charcoal therapy. Phlebotomy is not indicated.

TABLE 21.7. HEPATOERYTHROPOIETIC PORPHYRIA

Epidemiology:
- Autosomal recessive Extremely rare (<20 cases)
- Early onset, before age 2

Pathophysiology: Uroporphyrinogen decarboxylase deficiency

Clinical presentation:
- Normochromic anemia
- Erythrodontia, pink urine, photosensitivity: marked
- Skin reactions: Vesicles, bullae, erosions, scarring, hyperpigmentation, mutilating scarring deformities of hands, ears, face and nose

Laboratory findings:
- Urine: Elevated uroporphyrinogen I, III and 7-carboxylporphyrine
- Feces: Elevated uroporphyrinogen, coproporphyrin, isocoproporphyrin
- Erythrocytes: Elevated protoporphyrin, usually Zn protoporphyrin
- Enzyme: Markedly decreased uroporphyrinogen decarboxylase
- Serum Fe normal

Differential Diagnoses:
- Erythropoietic porphyria, porphyria cutanea tarda, erythropoietic protoporphyria

TABLE 21.8. HEREDITARY COPROPORPHYRIA

Epidemiology:
- Autosomal dominant
- Age of onset: Any age
- Incidence: Rare (<50 cases reported)

Pathophysiology:
- Deficiency of coproporphyrinogen oxidase

Clinical presentation:
- Neurovisceral symptomatology, similar to AIP
- Precipitating agents: Barbiturates, other drugs
- Photosensitivity: Infrequent
- Skin reactions: Blisters

Laboratory findings:
- Urine: Elevated coproporphyrin III
- Feces: Markedly elevated coproporphyrin III, less elevated protoporphyrin; aminolevulinic acid and porphobilinogen are also elevated during acute attacks
- Enzyme: in fibroblasts, lymphocytes, elevated hepatic ALA-synthase

Differential diagnosis:
- Acute intermittent porphyria, porphyria cutanea tarda

Hereditary Coproporphyria (Idiopathic Coproporphyria, Coproporphyrinogen Oxidase Deficiency)

Hereditary coproporphyria (HCP), an autosomal dominant disorder presents as an acute neurological porphyria, as photosensitivity, or both (Table 21.8). The acute neurological attacks are similar to those in a milder form of AIP. Abdominal pain is the most common symptom. Hepatic function usually is not impaired. The attacks are associated with elevated urine porphobilinogen, and, to a lesser extent, elevated ALA levels. As a result of the enzyme defect, an overproduction of the coproporphyrinogen occurs that inhibits hydroxymethylbilane synthase, resulting in an elevation of porphobilinogen. Coproporphyrin III is elevated in urine and feces. The key laboratory tests are urine porphyrin isomer profile and fecal porphyrin profile. The management of hereditary coproporphyria is similar to the treatment of AIP or VP (carbohydrate rich diet, administration of glucose, hematin, avoidance of precipitating factors).

Variegate Porphyria (South African Type Porphyria, Porphyria Cutanea Tarda Hereditaria, Mixed Porphyria, Protocoproporphyria, Protoporphyrinogen Oxidase Deficiency)

Clinically, VP includes the manifestations of AIP and PCT, which may occur in the same individual (Table 21.9). Most often, patients present with the symptoms of AIP at the first time of discovery. Chronic cutaneous symptoms also are frequently present. Peripheral neuropathy is an important part of the clinical syndrome in many patients. VP is most common among Dutch descendants in South Africa.

TABLE 21.9. VARIEGATE PORPHYRIA

Epidemiology:
- Autosomal dominant
- Age of onset: Usually between 15 and 30 years
- In South Africa: incidence 3/1000, rare elesewhere

Pathophysiology:
- Protoporphyrinogen oxydase deficiency

Clinical presentation:
- Neurovisceral symptomatology similar to AIP
- Photosensitivity and skin lesions: similar to PCT
- Acute attacks similar to AIP precipitated by barbiturates, dapsone, estrogens
- Abdominal pain, nausea, vomiting
- Behavioral changes, seizures, paralysis

Laboratory findings:
- Urine: Elevated aminolevulinic acid, porphobilinogen in acute attacks, normal between attacks
- Feces: Increased protoporphyrin and coproporphyrin (PROTO > COPRO III) and X-porphyrins (ether- acetic acid insoluble porphyrins)
- Enzymatic: Decreased protoporphyrinogen oxydase in fibroblasts, somewhat decreased ferrochelatase; increased hepatic ALA synthetase

Differential diagnosis:
- Porphyria cutanea tarda, hereditary coproporphyria, hepatoerythropoietic porphyria

The primary enzyme defect of protoporphyrinogen oxidase results in excess protoporphyrinogen, which inhibits hydroxymethylbilane synthase giving rise to elevated PBG. Like in AIP, acute attacks are characterized by excretion of large amounts of ALA and PBG, but urinary excretion may occur between attacks as well. Urine coproporphyrin III also is elevated during attacks. In the stool, protoporphyrin, coproporphyrin, and the ether-acid insoluble porphyrins (x-porphyrins) are elevated. The key laboratory tests are urinary PBG and ALA during an attack, fecal porphyrin profile (stool protoporphyrin), and plasma fluorescent scan.

Management of this disorder involves avoidance of precipitating drugs, treatment for neurovisceral symptoms as in AIP, and protection from sunlight.

Erythropoietic Protoporphyria (Erythrohepatic Protoporphyria Protoporphyria, Ferrochelatase Deficiency)

This disease begins early in life and the symptoms vary greatly in severity (Table 21.10). Some individuals may not experience photosensitivity, but most patients are moderately sensitive to sunlight. Symptoms include burning, itching, stinging in light-exposed skin, particularly of the nose, cheeks, and dorsum of the hands. Erythropoietic protoporphyria (EPP) is inherited in an autosomal-dominant manner, and ferrochelatase activity is about half of normal in affected individuals. Excess protoporphyrin is deposited in the skin, resulting in mild photosensitivity starting in childhood. A rare homozygous form of this disorder is accompanied by more serious symptoms. Late onset cases also are known.

TABLE 21.10. ERYTHROPOIETIC PROTOPORPHYRIA

Epidemiology:
Autosomal dominant, variable penetrance
Age of onset: 1–4 years
Pathophysiology:
Ferrochelatase deficiency
Clinical presentation:
Photosensitivity: Mild to severe, onset immediate – within minutes
Pruritus, burning, edema, erythema, urticaria, purpura, rare bullae on nose and hands, waxy scars over nose, hands, erosions, crusts, wrinkling of face, waxy thickening of the knuckles, nose with fine linear scarring
Cholelithiasis, cholestasis, hepatic failure
Anemia: Uncommon
Laboratory findings:
Urine: Normal porphyrins
Feces, erythrocytes and plasma: Elevated protoporphyrin
Erythrocytes show red fluorescence
Enzyme defect: Ferrochelatase deficiency in skin, fibroblasts, liver, and bone marrow
Differential diagnosis:
Congenital erythropoietic porphyria, extreme sunburn, polymorphous light eruption

The laboratory diagnosis of ferrochelatase deficiency can be made by detecting free protoporphyrin in erythrocytes and feces. Plasma protoporphyrins and fecal coproporphyrin also may be increased, and are present in higher concentrations than fecal coproporphyrin. Urine shows normal porphyrin levels and possibly increased coproporphyrin I. The key laboratory tests are free erythrocyte protoporphyrin, plasma porphyrin, plasma fluorescent scan, and fecal porphyrin profile.

In the management of EP, beta-carotene (Solatene) is used for sunburn protection by quenching singlet oxygens and trapping free radicals. Hepatic failure is difficult to manage, liver transplantation has been used successfully.

Erythropoietic Coproporphyria

There are only three published cases of erythropoietic coproporphyria (ECP) known. Photosensitivity associated with this disorder and elevated protoporphyrin and coproporphyrin in red blood cells are the main manifestations of this disorder. Very little is known about this entity.

Lead Poisoning

Lead inhibits two steps in the heme synthetic pathway: ALA dehydratase and ferrochelatase. Consequently, the substrates of these two enzymes, ALA and protoporphyrin accumulate in lead poisoning. In addition, coproporphyrin also is often elevated in urine. Urinary measurement of ALA is a sensitive indicator of lead poisoning. Measurement of protoporphyrins in erythrocytes was until recently an important test for population screening for lead poisoning. Measurement of protoporphyrins, however, is no longer acceptable for the detection of lead poisoning because of its relative insensitivity and lack of specificity.

Secondary Disorders of Porphyrin Metabolism

Various forms of liver disease also can cause increased excretion of porphyrins into the urine, most commonly only coproporphyrin elevations are seen. This also is the case in some patients with hemolytic disease. Rarely, hepatic tumors produce excess porphyrin metabolites.

Protoporphyrins may be elevated because of secondary causes such as disorders involving the system of heme biosynthesis. These disorders include iron deficiency, sideroblastic anemia, secondary polycythemia, and any cause of excessive erythrocyte destruction.

Hereditary tyrosinemia is another disorder related to porphyrin metabolism. Succinylacetone, a metabolite produced during the degradation of tyrosine, inhibits ALA dehydratase, and the patients affected by this disorder develop symptoms resembling AIP.

LABORATORY TESTS FOR PORPHYRIN DISORDERS

Qualitative Methods

If the clinical picture suggests an acute porphyria, the first test should be the measurement of PBG in urine. Increased PBG is the most important biochemical abnormality in urine in AIO, and, during attacks, in VP and HCP. When these disorders are suspected, both PBG and ALA should be measured. These assays help to differentiate AIP, ALA-D porphyria, and lead poisoning. In the latter two, only urine ALA is elevated, while PBG is normal. In cutaneous porphyrias, porphyrins are produced in excessive amounts, VP and HCP both present with photosensitivity and neurovisceral manifestations.

Testing for Porphobilinogen in Urine

The classic method for PBG screening is the Hoesch test in which porphobilinogen is reacted with *p*-dimethyl-aminobenzaldehyde (Ehrlich's reagent) to yield a magenta color at the interface. In the Watson-Schwartz test, the urobilinogen-Ehrlich's reagent complex is removed with butanol into the organic phase, improving the assay's specificity.

Quantitative Measurement of Porphobilinogen

Most quantitative methods employ chromatographic isolation and subsequent reaction with Ehrlich's reagent.

Measurement of Porphyrins in Urine

Because porphyrins exhibit red fluorescence, it often is possible to detect porphyrins in the urine by examination under a Wood's light. Addition of 10% hydrochloric acid enhances fluorescence. Direct spectrophotometric analysis also can be performed. However, these methods are neither sensitive, nor specific. Quantitative methods use solvent extraction from acidified urine followed by fluorimetric or spectrophotometric measurement.

Measurement of d-Aminolevulinic Acid in Urine

Reliable measurement of ALA can be achieved by chromatographic separation from other components, especially from porphobilinogen. The isolated ALA is reacted with acetylacetone to form a pyrrole, which then is reacted with Ehrlich's reagent.

Measurement of Protoporphyrin in Blood

Protoporphyrin can be extracted in the presence of acid, which dissociates zinc from the porphyrin molecule. Thus, the measurement resulting from this method is called “free erythrocyte protoporphyrin.” Direct measurement of protoporphyrin can be achieved from an instrument called a hematofluorometer, which measures zinc-protoporphyrin.

Fecal Porphyrin Analysis

Fecal porphyrin analysis employs repeated acidified organic solvent extraction of porphyrins from the specimen. Then the extracts are centrifuged and prepared for spectrophotometric analysis. Thin-layer chromatography, high-performance liquid chromatography, and capillary electrophoresis also have been used for separation and quantitative analysis of porphyrins.

HEME

Part of "21 - Heme Synthesis and Catabolism"

The role of heme in hemoglobin is discussed in Chapter 41. Other important heme-containing proteins (hemoproteins) include myoglobin, cytochrome P-450, catalase, peroxidase, NO synthetase, Try-pyrrolase, and adenylcyclase.

BILIRUBIN

Part of "21 - Heme Synthesis and Catabolism"

Roughly 1% of a person's red blood cells normally turn over on any given day. This results in the production of 175 to 250 mg of bilirubin daily. Another 75 to 100 mg of bilirubin is derived from ineffective erythropoiesis and the turnover of other hemoproteins. Heme released from these sources is rapidly bound by the serum glycoprotein, hemopexin, and transported to the reticuloendothelial system. If the carrying capacity of hemo-pexin is exceeded, circulating heme binds to albumin to form methemalbumin.

Heme is degraded to biliverdin IX through the microsomal heme oxygenase enzyme complex (HO). This reaction involves cleavage of the α-methene bridge of heme and loss of iron. Also produced is carbon monoxide, long thought to be an incidental side product of heme catabolism, but now recognized as a potent vasodilator. HO is anchored to the endoplasmic reticulum at the cytosolic border. Three isoforms of HO have been identified that are expressed in various tissues. Primarily HO-1 is expressed in the spleen, where most heme is catabolized. HO-1 is a heat-shock protein, inducible by heme and other stimuli: fever, hormones, stress, starvation, and sundry metals and organic compounds. HO-1 is induced in the liver after splenectomy. HO-1 is inhibited by zinc and tin-protoporphyrin. Zinc protoporphorin accumulates when marrow iron stores are low, and this feedback inhibition on heme catabolism may lessen the anemia resulting from iron deficiency.

Biliverdin is highly water-soluble, and would make a logical end product for heme degradation in humans as it serves in birds, reptiles, and amphibians. However, in man, biliverdin is rapidly converted to bilirubin by biliverdin reductase. There is a clear evolutionary trade-off in this extra metabolic step as bilirubin is more hydrophobic than biliverdin. Hydrophobicity allows for transplacental excretion of bilirubin from the fetal to maternal circulation, but bilirubin can selectively deposit in the basal ganglia (kernicterus) if pre- or neonatal hyperbilirubinemia occurs. Both bilirubin and biliverdin are potent free-radical scavengers, but the physiologic significance of this property is unclear. The majority of biliverdin formed in the adult is biliverdin IX_{α}, but biliverdin IX_{β} predominates in the fetus. Four different isoforms of biliverdin reductase, with higher affinities for each

biliverdin IX isomer (biliverdin $IX_{\alpha\text{-}\delta}$), have been identified in the human liver. The hydrophobicity of bilirubin is surprising considering the presence of propionic acid side chain on the B and C rings. However, the cis configuration of the two interring methene bridges leads to intramolecular hydrogen bonding of the acidic groups, rather than solvation. Phototherapy with blue light causes a configurational change in one or both of these methene bridges, leading to more hydrophillic molecules that are excreted readily.

Bilirubin readily binds to albumin, and this complex is transported through the circulation to the liver. Bilirubin is transferred to ligandins within the hepatocyte. Ligandins comprise a family of cytosolic carrier proteins that bind various substrates destined for biochemical modification by the hepatocyte. Ligandin binding prevents the diffusion of these substrates back into the bloodstream. The principal fate of bilirubin is glucuronidation, the conjugation of one or two sugar groups to the propionic acid side chains. Glucoronidation disrupts the intramolecular hydrogen bonds and greatly enhances the solubility of bilirubin. Most bilirubin progresses to the diglucuronide state, in reactions mediated by UDP-glucuronyl transferases (UGTs). Two families of UGTs exist: UGT1, which metabolize bilirubin and phenol; and UGT2, which metabolize steroids. Both families also glucuronidate a number of xenobiotics, and their activities can be modulated by drugs like phenobarbital and hormones like thyroxine and cortisol. Various alleles of the UGTs may predispose individuals to side effects from many common drugs. Characterizing these interactions is the basis for much work in the emerging field of pharmacogenetics.

Conjugated bilirubin is secreted into the bile by the cannicular multispecific organic anion transporter (CMOAT), also known as the multidrug resistance protein (MRP). This protein originally was identified in various tumors during chemotherapy. In the canaliculus, bilirubin mixes with bile salts, phospholipids, and cholesterol to form bile. Bile proceeds first to the gall bladder, where it becomes more concentrated and acidic, and then to the duodenum after mixing with the pancreatic juices. Trace amounts of conjugated bilirubin diffuse out of the hepatocyte, into the circulation, and ultimately into the urine. In contrast, unconjugated bilirubin is protein-bound and does not pass into the urine. Consequently, urine bilirubin reflects the conjugated form and may be elevated in conjugated hyperbilirubinemia.

Urobilinogen

Heme biosynthetic and catabolic pathways are taxonomically widespread, and we are symbiotic in certain aspects of bilirubin metabolism. Many of the heme breakdown products in human bodily fluids and excrement are products of enteric bacterial metabolism. Some bilirubin is deconjugated by bacterial ß-glucuronidases in the terminal ileum and colon, and the unconjugated bilirubin is reabsorbed into the blood stream via the enterohepatic circulation. The remainder of unconjugated bilirubin is reduced to form d-urobilinogen, mesobilirubinogen, and stercobilinogen (collectively known as urobilinogen). A portion of urobilinogen is absorbed via the enterohepatic circulation and finally emerges in the urine. The majority of urobilinogen subsequently is oxidized by bacteria to form the urobilins and excreted in the feces. Urine urobilinogen may be elevated in conditions associated with increased bilirubin production such as hemolytic jaundice.

Laboratory Assays

Bilirubin usually is measured using a red-violet chromogenic reaction with diazotized sulfanilic acid (diazo reaction), on serum, plasma, or urine specimens. The reaction course is biphasic, with a “direct” or rapidly reacting component composed of soluble conjugated bilirubin and bilirubin covalently bound to protein, and an “indirect” or slowly reacting component composed mostly of insoluble, unconjugated bilirubin. Measurements made early in the reaction essentially measure only the direct component, but at least 25% of the conjugated bilirubin is missed. Addition of “accelerators” such as methanol, caffeine, or sodium benzoate enhance the solubility and reaction rate of unconjugated bilirubin, presumably through disruption of intramolecular hydrogen bonds. This allows quantification of the total bilirubin. Indirect bilirubin then can be estimated by subtraction of the direct from the total bilirubin. Typical reference ranges for total and direct bilirubin are 0.3 to 1.2 mg/dL (5 to 20 µM) and <0.3 mg/dL (<5 µM), respectively. Hemoglobin interferes with the reaction, resulting in falsely elevated bilirubin measurements in severe anemia and falsely depressed measurements in hemolyzed specimens. An alternate approach for measuring bilirubin, employs reflectance photometry performed on a patient's skin or on thin films of blood for measurement of either conjugated or unconjugated bilirubin. Bilirubin has an absorption maximum at 455 nm in blood. Unfortunately, oxyhemoglobin also absorbs at this wavelength. However, the oxyhemoglobin component can be estimated from a parallel measurement at 575 nm, so that contribution can be subtracted out. Other substances present in the skin can interfere, and consequently transdermal measurements are only used in newborns. Bilirubin can be measured by HPLC, but this generally is reserved for research purposes.

Urobilinogen is measured routinely as part of dipstick urinalysis through a chromogenic reaction with p-diethylaminobenzaldehyde. Normally, this semiquantitative assessment is calibrated so normal levels of urobilinogen, up to 1 mg/dL of urine, do not register. Urobilinogen is not measured routinely in serum. Formerly, there was some clinical interest in quantifying stool urobilinogen to estimate hemoglobin degradation rates, however, those days are over.

DISORDERS OF BILIRUBIN METABOLISM

Part of "21 - Heme Synthesis and Catabolism"

Disorders of heme catabolism are, for all practical purposes, disorders of bilirubin metabolism. To date, only one case of HO-1 deficiency has been identified, and no human biliverdin reductase mutatations have been described. Although biliverdin levels may become elevated from various hepatic insults, biliverdin itself is excreted readily and apparently is nontoxic. On the other hand, acquired or genetic defects have been identified for most steps in bilirubin uptake, metabolism, and excretion. These defects

lead to hyperbilirubinemia of the unconjugated and/or conjugated type (Table 21.11). When serum bilirubin levels exceed about 2 mg/dL (34 μM), a yellow discoloration of the skin (jaundice) or sclera (icterus) becomes evident. For intrinsically dark-skinned people, inspection of the soles and hands may be necessary to see early jaundice. Only elevated conjugated bilirubin levels will result in patients who report darkening of the urine or yellowing of the urine foam.

TABLE 21.11. DIFFERENTIAL DIAGNOSIS OF JAUNDICE

Unconjugated hyperbilirubinemia
Increased heme catabolism
Hemolytic anemia
Hematoma
Impaired hepatic conjugation
Neonatal jaundice
Gilbert syndrome
Crigler-Najjar syndromes, types I and II
Conjugated hyperbilirubinemia
Impaired hepatic excretion
Hepatocellular disease
Dubin-Johnson syndrome
Rotor syndrome
Posthepatic obstruction

TABLE 21.12. NEONATAL HYPERBILIRUBINEMIA

Epidemiology:
Most, if not all infants
Pathophysiology:
Transient deficiency of UDP glucuronyl transferase – 1
Clinical presentation:
Jaundice, usually more pronounced in upper body
Laboratory findings:
Increased total bilirubin, with initially undetectable direct bilirubin
Transaminases, alkaline phosphatase, albumin, PT and PTT normal
Differential diagnosis:
Hemolytic disease of the newborn

Unconjugated Hyperbilirubinemia

Unconjugated hyperbilirubinemia is diagnosed when total bilirubin is elevated and <20% is direct. This occurs when the amount of unconjugated bilirubin in the blood exceeds the conjugating capacity of the liver: either because the amount of unconjugated bilirubin is especially large, the conjugation capacity of the liver is relatively low, or both.

Large amounts of free heme, and subsequently unconjugated bilirubin may be released into the blood in hemolysis or when hematomas form. As the functional conjugating reserve of the healthy liver is large, total bilirubin levels rarely rise more than two- to fourfold. The diagnosis of hemolysis may be aided by evaluation of the patient's history, complete blood count, and measurement of the serum haptoglobin among other tests. The most common form of unconjugated hyperbilirubinemia is neonatal jaundice (Table 21.12). This transient condition results from inadequate hepatic UGT1 activity, and manifests between days 2 and 5 of life in most newborns. Jaundice at birth suggests other metabolic derangements, as the maternal-fetal circulation protects the newborn from normal heme catabolism. Typically, the level of serum unconjugated bilirubin remains below 5 mg/dL (85 μM), however when levels exceed 20 mg/dL (340 μM) then the developing basal ganglia and other central nervous system structures are susceptible to toxic bilirubin deposition, known as kernicterus. This pigment is grossly evident at autopsy when the brain is sectioned. Brain damage can be severe, and the auditory tracts are particularly susceptible. Hypothyroidism (cretinism) delays the development of UGT1 activity, potentiating kernicterus. Newborns also may have increased hemolysis, particularly if ABO incompatibility exists, which aggravates neonatal jaundice. Certain agents, including chloramphenicol, vitamin K, and perhaps even free fatty acids in mother's milk, may further impair UGT1 activity. Subjective clinical assessment of bilirubin levels frequently is made based on the intensity and distribution of jaundice (which is usually most pronounced over the head and upper body), but these estimates have been shown to be unreliable. Intense visible light phototherapy commonly is employed for infants at risk to covert the unconjugated bilirubin to more soluble forms, as described above. The HO-1 inhibitor, tin-protoporphyrin soon may be used widely to inhibit bilirubin formation in neonatal jaundice.

Three clinically distinct forms of unconjugated hyperbilirubinemia are attributable to mutations in UGT1 (Gilbert's syndrome and Crigler-Najjar syndrome, types I and II). Gilbert's syndrome is defined clinically as the presence of persistent, mild, unconjugated hyperbilirubinemia ranging from 1 to 3.5 mg/dL (17 to 60 μM), and has a prevalence of 5% to 10% (Table 21.13). This level of hyperbilirubinemia is of no physiologic consequence unless the patient is exposed to a heme bolus, as most commonly occurs when the patient has coexisting glucose-6-phophate dehydrogenase deficiency. Psychological or physiologic stress also can exacerbate the jaundice. The severity of neonatal jaundice has been linked to hetero- or homozygosity of common UGT1 alleles that also cause Gilbert syndrome. Often, occult hemolysis must be ruled out in the evaluation of Gilbert syndrome. A helpful diagnostic challenge is to monitor unconjugated bilirubin, which rises after a prolonged fast or caloric restriction in the Gilbert syndrome, but not in the case of occult hemolysis. The uncommon Crigler-Najjar syndrome type II results in higher, but still nontoxic levels of hyperbilirubinemia, ranging from 3. to 20 mg/dL (60 to 340 μM). In both Gilbert syndrome and Crigler-Najjar syndrome type II, there is a relative abundance of monoconjugated bilirubin. UGT1 activity in these patients is inducible by phenobarbital. In contrast, no UGT1 activity is inducible in the rare Crigler-Najjar syndrome type I (see Table 21.14 and Table 21.15). These patients have severe hyperbilirubinemia

over 20 mg/dL (<340 μM), but liver function tests are otherwise normal. These patients generally succumb to kernicterus in childhood. Phototherapy has diminishing value in older children because of their thicker skin, so plasmapheresis or exchange transfusions may be employed to manage these patients.

TABLE 21.13. GILBERT'S SYNDROME

Epidemiology:

~5% of population

Pathophysiology:

Deficiency and or impairment of UDP glucuronyl transferase – 1

Clinical presentation:

Asymptomatic or occasionally with jaundice under stress

Laboratory findings:

Increased total bilirubin < 6 mg/dL, with low direct bilirubin

Transaminase, alkaline phosphatase, albumin, PT and PTT normal

Differential diagnoses:

Acute hepatitis or hepatotoxicity

TABLE 21.14. CRIGLER-NAJJAR SYNDROME TYPE I

Epidemiology:

Rare

Pathophysiology:

Impairment of UDP glucuronyl transferase – 1

Clinical presentation:

Severe jaundice, typically resulting in kernicterus.

Laboratory findings:

Increased total bilirubin above 20 mg/dL, with undetectable direct bilirubin.

Total bilirubin unchanged after phenobarbital ingestion

Transaminase, alkaline phosphatase, albumin, PT and PTT normal.

Differential diagnoses:

Crigler-Najjar syndrome type II

TABLE 21.15. CRIGLER-NAJJAR SYNDROME TYPE II

Epidemiology:

Rare

Pathophysiology:

Deficiency and or impairment of UDP glucuronyl transferase – 1

Clinical presentation:

Mild-moderate jaundice, exacerbated under stress

Laboratory findings:

Increased total bilirubin between 6 and 20 mg/dL, with low to undetectable direct bilirubin

Total bilirubin diminishes after phenobarbital ingestion

Transaminase, alkaline phosphatase, albumin, PT and PTT normal

Differential diagnoses:

Crigler-Najjar type I, Gilbert's syndrome complicated by hemolytic anemia

TABLE 21.16. DUBIN-JOHNSON SYNDROME

Epidemiology:

Rare, except among Sephardic Jews

Pathophysiology:

Impairment of cannicular mixed organic ion transporter

Clinical presentation:

Mild to moderate jaundice, red to brown urine

Laboratory findings:

Increased direct bilirubin ranging from 3–15 mg/dL

Total bilirubin unchanged after phenobarbital ingestion

Urinary coproporphyrin I levels can distinguish heterozygotes from homozygotes

Transaminase, alkaline phosphatase, albumin, PT and PTT normal

Differential diagnosis:

Rotor syndrome

Conjugated Hyperbilirubinemia

Whenever there are defects with hepatic excretion or post-hepatic obstruction, conjugated bilirubin levels generally will rise before unconjugated levels. If more than 50% of the total bilirubin is direct, the condition is termed conjugated hyperbilirubinemia. Clearly, many patients will not fit into either the unconjugated or conjugated hyperbilirubinemia definitions, especially as disease progresses. When conjugated bilirubin is elevated, the urinary bilirubin also may be increased. Hepatocellular diseases and cholestasis also may be accompanied by elevations of alanine and aspartate transaminases and alkaline phosphatase. Relative elevations of the transaminases and alkaline phosphatase are most helpful in distinguishing these etiologies. Typically, hepatocellular disorders produce a disproportionate elevation in transaminases and a slight or modest rise in alkaline phosphate and 5-nucleotidase (hepatocellular pattern) whereas obstructive jaundice (from etiologies such as cholelithiasis, tumors, or other causes) produces a more marked increase in alkaline phosphatase and 5-nucleotidase than in the transaminases (obstructive pattern). Intrahepatic cholestasis usually presents with an obstruction enzyme pattern, but elevated transaminases resembling hepatocellular disease may be observed or in other cases, a mixed enzyme pattern may occur.

Dubin-Johnson syndrome (see Table 21.16) is characterized by persistent, mild, conjugated hyperbilirubinemia ranging from 3 to 15 mg/dL (51 to 255 μM). It is as common as 0.1% in select populations, such as Persian Jews. The condition generally is benign, but stress and oral contraceptive use can exacerbate the jaundice. On autopsy, these patients have black livers resulting from the accumulation of hepatocyte lysosomal pigment, yet other liver function tests are normal. The disorder probably results from mutations in the CMOAT gene. Increasing levels of coproporhyrin I in the urine distinguish hemozygotes from heterozygotes and normals. Another very rare, benign, and unexplained causes of conjugated hyperbilirubinemia is the Rotor syndrome. Lower levels of coproporhyrin I in the urine distinguish Rotor syndrome patients from Dubin-Johnson heterozygotes.

ACKNOWLEDGMENT

Part of "21 - Heme Synthesis and Catabolism"

The authors acknowledge the original contribution by Dr. Thomas P. Koch, entitled “Porphyrins, Bilirubin, and Other Bile Pigments” in the previous edition of this book, on which portions of this chapter are based.

SUGGESTED READINGS

Porphyrias

Bickers DR, Pathak MA, Lim HW. The porphyrias. In: Fitzpatrick TB, Eisen AZ, Wolff K, et al., eds. *Fitzpatrick's Dermatology for General Medicine*. New York: McGraw-Hill, 1993:1854–1893.

Koch TR. Porphyrias, bilirubin, and other bile pigments. In: *McClatchey's Clinical Laboratory Medicine*. 1st ed. New York: Williams & Wilkins, 1994;403–416.

Moore MR. The biochemistry of heme synthesis in porphyria and in porphyrinurias. *Clinics in Dermatol* 1998;16:203–223.

Nuttall KL. Porphyrins and disorders of porphyrin metabolism. In: Burtis CA, Ashwood ER, eds. *Tietsz Textbook of Clinical Chemistry*, 3rd ed. Philadelphia: Saunders, 1999;1711–1745.

Photodermatol Photoimmunol and Photomed. Special reviews on porphyrias. 1998;14:43–73.

Ratnaike S, Blake D. The diagnosis and follow-up of porphyria. *Pathology* 1995;27:142–153.

Special issue on porphyries. *Semin Liver Dis* 1998;18(1):3–101.

Zaider E, Bickers DR. Clinical laboratory methods for diagnosis of porphyrias. *Clinics in Dermatol* 1998;16:277–293.

Bilirubin Metabolism

Deiss A. Destruction of Erythrocytes. in *Wintrobes' Clinical Hematology, 10th ed.*, New York: Williams & Wilkins, 1998:282–287.

Iyanagi T, Emi Y, Ikushiro S-I. Biochemical and molecular aspects of genetic disorders of bilirubin metabolism. *Biochim Biophys Acta* 1998;1407:173–184.

Kamath PS. Clinical approach to the patient with abnormal liver test results. *Mayo Clin Proc* 1996;71:1089–1095.

Yamaguchi T, Nakajima H. Changes in the composition of bilirubin-IX isomers during human prenatal development. *Eur J Biochem* 1995;233:467–72.

22
Toxicology

Tai C. Kwong

SCOPE OF CLINICAL TOXICOLOGY AND THE ROLE OF THE TOXICOLOGY LABORATORY

Part of "22 - Toxicology"

Clinical toxicology in laboratory medicine can be defined as the analysis of drugs in human biological fluids for the purpose of patient care. The two traditional clinical toxicology services provided by clinical laboratories are therapeutic drug monitoring and emergency toxicology. Therapeutic drug monitoring, which is the measurement of serum drug concentration to aid in optimizing drug therapy, is the topic of another chapter. Emergency toxicology is the accumulation of laboratory data to assist in the diagnosis and treatment of poisonings resulting from self-poisoning, accidental ingestion, or exposure.

Because of the complexity of the diagnosis and treatment of poisoning and substance abuse, a team effort involving the clinical staff and laboratory personnel is required. In the emergency department, the clinician first makes an assessment based on history and physical examination, and the patient is treated symptomatically to support vital signs and vital functions. Toxicological data, if available at this point, can influence the diagnosis and subsequent treatment of the patient. If the patient is obtunded or comatose, a drug history may not be available. Drug history, even if available, is frequently inaccurate. Hence, an emergency toxicology screen can validate drug history or identify drugs that are not previously suspected. In those cases where specific treatments are effective, expeditious laboratory identification of the toxic agents is invaluable for prompt institution of treatment: for example, the initiation of N-acetylcysteine therapy for acetaminophen overdose, hemoperfusion for massive barbiturate overdose, or administration of ethanol or 4-methylpyrazol for methanol and ethylene glycol poisoning. A drug screen that fails to identify specific toxic agents still is an important laboratory result because it will influence the clinician to evaluate the differential diagnosis and to explore other diagnostic studies to explain the patient's clinical presentation. The determination of serum drug levels can, in some instances, help to gauge the severity of intoxication and prognosis. Ensuing determinations can be useful in monitoring and evaluating the effectiveness of the treatment procedure.

In the last few years, the widespread abuse of illicit drugs has placed additional demands on the health-care system. Consequently, clinical toxicology laboratories have moved beyond their traditional activities and now are routinely performing urine testing of drugs of abuse for other clinical services such as obstetrics for pregnant drug abusers, pediatrics for their newborns, and drug-dependency treatment programs. Such testing usually focuses on a small panel of the most frequently abused drugs such as cocaine, cannabinoids, and the benzodiazepines, rather than the broad-spectrum screening performed in emergency toxicology (1, 2).

Pregnant patients with a chemical dependency, particularly that of cocaine, are at high risk for obstetric complications (3). Urine testing for drugs of abuse provides evidence for the diagnosis of drug dependency and subsequent monitoring of drug use during the course of the pregnancy. For babies born of drug-dependent mothers, there is a higher risk of congenital malformations and neurobehavioral impairment (4). Therefore, early identification of their addiction not only allows better medical care; it will also permit prompt social evaluation and intervention. Urine drug testing is an objective way to identify these infants. However, because urine collection for infants is difficult and drug concentration is low, urine screening probably underestimates the number of babies at risk. Recent data suggested that testing for drugs in meconium and hair can be alternate approaches (5, 6 and 7).

In many localities, the hospital, upon the demonstration of illicit drug in an infant's urine, may be required to report it to a governmental agency. This may trigger a sequence of events that in extreme cases may result in the removal of the infant from the mother's custody. While the mother's family and social environment are important considerations, the confirmed positive urine drug test result frequently is the overriding medical evidence in this contest of custody. In these instances, the clinical toxicology testing has taken on an additional evidentiary role. Therefore, the laboratory must be prepared to defend its data against challenges.

The emphasis on treatment and rehabilitation of chemical dependency has resulted in increased enrollment in drug treatment programs. Many clinical toxicology laboratories now routinely perform urine testing in support of the diagnosis of these patients' chemical dependency (2). Urine testing also is an adjunct to the counseling and management of these patients by monitoring their progress during rehabilitation. Enrollment in drug treatment programs is not always voluntary. A drug-addicted employee may have enrolled because his employer's substance abuse policy mandates that all employees with a drug problem must be rehabilitated as a condition for continued employment. Successful rehabilitation is evidenced in part by negative drug

test results obtained during and after the treatment program. Therefore, although the toxicology laboratory is performing testing for medical reasons, test results ultimately may be used for administrative purposes, and the potential for legal challenge by dismissed employees is real.

Some clinical toxicology laboratories have extended their expertise in drugs of abuse testing into the new arena of workplace drug testing. This is nonmedical testing, and is a departure from the traditional mission of clinical toxicology. The forensic nature and specific requirements of workplace drug testing are discussed at the end of this chapter.

TOXICOLOGY SERVICE

Part of "22 - Toxicology"

Cost constraints and staffing problems limit the kind of toxicology service a laboratory can provide. The laboratory has to design a drug screen that is cost-effective and one that still can meet the basic needs of clinicians. Furthermore, programmatic issues that need to be considered include what drugs comprise a drug screen, which assays should be qualitative or quantitative, what type of sample is required, when the specimen should be obtained, and how much time is involved in completing a drug screen.

Drug Lists

The ideal laboratory will be one that can analyze for any drug the clinician might encounter, and can do so in the shortest turnaround time. In reality, this goal is difficult to achieve. Methodologies for many drugs are not suitable for a clinical toxicology laboratory, either because they require costly and sophisticated instrumentation, or the turnaround time is too long for them to be clinically useful. Even in laboratories blessed with experienced staff and advanced instrumentation, the financial and personnel requirements for performing timely, comprehensive drug screening are prohibitive. Therefore, a drug screen should be one that is designed to detect a shorter list of drugs, the selection of which is based on the prevalence of these drugs in the population served by the emergency department, the clinical usefulness of their early identification, and the ability of the laboratory to perform the analyses within an acceptable turnaround time. This list should be determined jointly by the emergency department and the laboratory. For example, a very basic clinical toxicology service may include screening tests for ethanol, salicylate, acetaminophen, barbiturates, tricyclic antidepressants, digoxin, cocaine, and opiates. In addition, serum iron, carboxyhemoglobin, and methemoglobin usually are analyzed in the clinical chemistry laboratory. Any additional drugs that have a high prevalence in the locality also should be included. The importance of turnaround time for a drug screen is related directly to the impact the results have on patient management. Therefore, these drug tests are recommended to be available on an immediate (stat) basis. With additional resources and expertise, the laboratory may enlarge its service to include other drug assays: e.g., amphetamines, benzodiazepines, phenothiazines, and sedative-hypnotics. The laboratory should encourage clinicians to inquire about drugs that are not on this list and whether the laboratory or a reference laboratory can detect these drugs.

Analytical Approaches

A drug screen is a compromise among rapid turnaround time, analytical specificity, and sensitivity, and it may be a combination of qualitative and quantitative analyses. An initial screen might be a targeted search for a small number of drugs by a series of tests, each of which is specific for a drug or group of drugs (e.g., spot tests or immunoassays). Obviously, this approach is restricted by the availability of specific methodologies. Moreover, the time required for these individual assays is cumulative, and the turnaround time rapidly deteriorates as the number of drug assays increases. If the number of drugs in a screen is to be expanded, a broad-spectrum screen using a methodology such as thin-layer chromatography or gas chromatography can be employed at the expense of lower sensitivity compared to immunoassays.

Specimen Types And Quantitation

The specimens most frequently submitted for analysis are urine, blood, serum/plasma, and gastric fluid. Drug screens in most clinical laboratories are performed on urine. The advantage of urine is that a large volume can be obtained, allowing analysis of drugs in low concentration. Urine, however, contains metabolites, which may complicate identification. For those drugs that are extensively metabolized, the parent drug may not be present; therefore an assay should be chosen that would identify the metabolites that are present in the greatest concentrations. Urine tests are qualitative, as drug concentrations in urine correlate poorly with clinical effects. Blood (and serum or plasma), although limited by sample volume, is the specimen of choice for quantitative analysis of those drugs for which there is correlation between drug level and toxicity. Quantitative tests usually are more time consuming. If knowledge of a specific level does not influence patient management, quantitation should not be performed. Some tests that provide useful quantitative information include those for acetaminophen, alcohols, salicylate, barbiturates, phonation, theophylline, lithium, digoxin, and iron. Quantitative level, however, can be misleading for drugs for which tolerance can develop. For example, a supposedly lethal barbiturate level for a novice user may be well tolerated by a chronic barbiturate abuser. First gastric aspirate, vomitus, or stomach washings are appropriate specimens if they are obtained soon after ingestion and particularly if pill fragments are noted. Analysis of gastric contents offers no advantage over urine or serum and is not recommended. Moreover, drugs that are absorbed rapidly or that are not taken orally will not be detected.

Accuracy

Although much emphasis has been placed on the rapidity with which a drug screen can be performed and reported, the importance of accuracy needs to be addressed (8, 9 and 10). The analytical methods used in performing a drug screen, including colorimetric assays, immunoassays, thin-layer chromatography (TLC), and gas chromatography (GC) do not necessarily give unequivocal identification of a drug. Some assays are group-specific rather than analyte-specific (e.g., an immunoassay for the opiates

rather than morphine). Some chromatographic methods do not separate isomers (e.g., quinine and quinidine by TLC and GC) or closely related drugs (e.g., the tricyclic antidepressants by GC). Hence there is a need for confirmation of a presumptive positive result obtained by the screening test.

Repeat analysis of the presumptive positive specimen using the same screening test method does not constitute confirmation, although it does serve to reduce random error. Confirmation requires testing a new aliquot of the original specimen using a technique, which is based on a different analytical principle and has greater specificity and at least equal sensitivity to that of the initial test. For example, a presumptive positive result obtained by an immunoassay for opiates cannot be confirmed by a second opiate immunoassay; TLC, which can specifically identify the opiate involved, can confirm it.

While confirmation testing is mandatory in forensic toxicology, in clinical toxicology, where rapid turnaround time to support a diagnosis of a suspected drug overdose is required, it is not always possible to meet these requirements. The clinical laboratories routinely report presumptive positives, but interpretation by the clinician is done within the context of both clinical and other laboratory findings. The laboratory, however, should make every attempt to do confirmatory testing whenever possible. The laboratory report must state accurately whether the positive result is a confirmed positive or an unconfirmed (presumptive) positive so that clinicians can properly evaluate toxicology results.

METHODOLOGIES

Part of "22 - Toxicology"

Spot Tests

Spot tests are among the simplest, most rapid, and inexpensive screening tests. Urine or serum is added directly to the reagents either in a test tube or on a white tile or porcelain dish. The formation of a color product indicates the presence of a specific drug, but more often, a class of drugs (e.g., a positive Folin-Ciocalteu test for phenolic compounds). Proper interpretation of a color test should take into account that a range of color is possible (e.g., phenothiazines give red range, blue, or violet colors with the Forrest reagent). Colors may vary depending on the conditions of the test, the amount of substance present, and the presence of other drugs and extraneous material. A positive test should trigger more definitive testing. Examples of spot tests routinely performed in clinical and forensic laboratories are in Table 22.1 (11).

TABLE 22.1. SPOT TESTS[a]

Drug(s)	Test	Comments
Volatiles (U,S)[b]	Dichromate	Alcohols and aldehydes
Salicylates (U,S)	Trinder's	Positive after therapeutic doses
Acetaminophen (U)	Cresol-ammonia	Positive after therapeutic doses
Phenothiazines (U)	FPN	Color and sensitivity vary with Phenothiazine
Imipramine/desipramine trimipramine (U)	Forrest	Interfered by some phenothiazines
Ethchlorvynol (U,S)	Diphenylamine	Good sensitivity and specificity

[a]From Stevens HM. Colour tests. In Moffet AC, ed. *Clark's isolation and identification of drugs*. 2nd ed. London: The Pharmaceutical Press, 1986: 128–147.

[b] U, urine; S, serum; FPN, fixed pattern noise.

Immunoassays

The immunoassays most frequently used in clinical laboratories are enzyme immunoassays (EMIT), fluorescence polarization immunoassays (FPIA), cloned enzyme donor immunoassay (CEDIA), and agglutination immunoassay (ONLINE). These assays have sensitivity more than adequate for drug overdose situations, and, being homogeneous assays requiring no preliminary extraction required before analysis, they can give test results rapidly. Moreover, they are automated easily.

A drug screen using immunoassays is a combination of individual tests. Although each individual test is fast, the time required for several accumulates. In addition, the spectrum of drugs that can be detected in such a screen is limited by assay kit availability. The immunoassay kits are designed specifically for use with either urine or serum/plasma. When a kit is used with a different matrix (e.g., using a urine assay to analyze blood), the laboratory needs to validate this modification (12).

In recent years, single-use devices, each of which consist of a panel up to seven or eight drugs, have become the mainstays of urine drug screen for many laboratories (13, 14 and 15). They offer the advantage of near-patient testing with results available in minutes, instead of the much longer turnaround time associated with laboratory-based drug screens. Most of these devices have been designed to meet the needs of workplace drug testing. Therefore, they may only consist of assays for the five drug classes (amphetamine and methamphetamine, cannabinoids, cocaine metabolites, codeine and morphine, and phencyclidine) with cutoffs mandated by the federal drug-testing program.

Most immunoassays are not specific for an analyte (16); many will detect a family of closely related compounds with different degrees of cross-reactivity. For example, the opiates assay will detect morphine, codeine, and related narcotics such as hydromorphone, hydrocodone and oxycodone. Therefore, a positive result will require additional testing (confirmation) to differentiate among members of this family. Workplace urine drug testing has been the driving force behind the development of more specific assays. For example, new amphetamine assays are highly specific for amphetamine and methamphetamine with much lower cross-reactivity with other sympathomimetics such as ephedrine. While this approach is intended to reduce the number of presumptive positives needing gas chromatography-mass spectrometry confirmation, the use of such an assay of limited specificity is less useful in clinical toxicology. The users of immunoassays should be familiar with assay specificities for proper interpretation of test results.

Thin-Layer Chromatography

The simplicity, reliability, and low cost of a TLC system are the reasons why this technique is widely used for drug screening and confirmation. Its selectivity of detection is enhanced through the

use of various stationary and mobile phases, and color development reagents (17). For larger laboratories, the ability to perform simultaneous analyses of several specimens is an advantage.

The disadvantage of TLC is lower sensitivity (usually greater than 1 mg/L) compared with GC or high-performance liquid chromatography (HPLC), although the limits of detection are adequate for a large number of commonly encountered drugs in overdose cases. TLC also has lower resolution compared with GC or HPLC, particularly when small size (10 cm) plates are used. Recent developments in high-performance TLC (HPTLC) have demonstrated improvement in both sensitivity and resolution. Despite the simplicity of TLC procedures, proper interpretation of a TLC chromatogram requires highly skilled and experienced personnel. A commercially available TLC system in kit form facilitates a broad-spectrum TLC screen by providing accelerated development and a systematic color detection scheme (TOXI-LAB, ANSYS Diagnostic, Lake Forest, CA)(18).

Gas Chromatography

GC, particularly when it is used with a capillary column, is a powerful tool for separating a mixture of drugs and it is the method of choice for volatile compounds (e.g., anesthetic gases, organic solvent, alcohols). It is less suitable for compounds of low volatility or those with polar functional groups. Careful choice of chemical derivatives can overcome these problems (19). More problematic are compounds that are thermally unstable (e.g., benzodiazepines). Although a wide choice of stationary phases is available, only a few are used commonly in drug analysis such as the roughly equivalently nonpolar dimethyl silicone phases of SE-30, OV-1, and OV-101, and the relatively more polar phenylmethyl silicone, OV-17. A single column of SE-30 (and its equivalents) or OV-17 can separate most of the drugs of interest, and is therefore sufficient for a drug screen. Separation has been advanced further by the use of capillary columns (20, 21). The higher efficiency of these columns results in tall, narrow peaks that improve detection limits. Among the detectors, the flame ionization detector (FID) responds with good sensitivity to nearly all classes of compounds with excellent linearity over a wide range of concentrations. The nitrogen-phosphorus detector (or NPD) has enhanced response to compounds containing nitrogen and phosphorous. Because most drugs contain nitrogen atoms, NPD particularly is useful for drug analysis and has sensitivity up to 100 times that of a FID. The electron capture detector (ECD) has very high sensitivity for compounds with high affinity for electrons such as those containing halogen, nitro, or carbonyl groups. The selectivity of the ECD and its high sensitivity makes it very useful for drugs such as benzodiazepines and phenothiazines, which are present in very low concentrations. The higher selectivity of this detector, however, makes it unsuitable for drug screening.

The use of a mass spectrometer (MS) as a GC detector provides sensitivity and specificity unsurpassed by other detectors. Mass spectrometry in combination with chromatographic separation is the most effective technique for identification of drugs (22, 23). The mode of ionization most widely used is electron impact (EI), which is the only mode available in the low-cost MS equipment. Because fragmentation can be quite extensive with EI, the molecular ion may not be detectable. Therefore, the molecular weight cannot be deduced. In chemical ionization, softer ionization results in less fragmentation, and a prominent molecular ion or quasi-molecular ion (frequently the base peak) allows molecular weight elucidation. Less fragmentation, however, means less structural information for unambiguous identification. Operation of an MS in the CI mode gives even higher sensitivity than EI, but the best performance will require more critical conditions.

Identification of unknowns by MS can be based on the computer matching of the unknown mass spectrum against those in commercially available mass spectral libraries. Alternatively, a laboratory can build its own library of the most frequently encountered drugs. The main disadvantages of MS—the complexity and cost of the instrumentation—have been addressed by the low-cost bench-top instruments that have been designed for routine analysis, multiple-operators, easy maintenance, and good reliability. These instruments have made GC-MS technologically accessible for many clinical laboratories. Still, GC-MS analysis is relatively costly in terms of equipment, personnel, and required time. It is not cost-effective for screening and is more suitable for confirmation testing. In forensic urine drug testing, confirmation by GC-MS analysis is the standard of practice and is mandated in government-regulated programs.

High-Performance Liquid Chromatography

HPLC is similar to GC in its ability to separate complex mixture of drug compounds and a variety of detectors (spectrophotometric, fluorometric, and electrochemical) that can be used to attain high sensitivity (24). HPLC has the advantage over GC in that polar compounds (e.g. drug metabolites) and those that are thermally labile can be analyzed. Until recently, HPLC has found its place in a toxicology laboratory mostly as an instrument for quantitative analysis of targeted groups of drugs (e.g., tricyclic antidepressants, benzodiazepines, barbiturates) rather than for broad-spectrum drug screen. Structural information helpful for identification of unknown compounds eluting from HPLC can be obtained using a photodiode array detector or mass spectrometer detector. An automated liquid chromatography (LC) system equipped with on-line extraction and analytical columns is available commercially with the capacity to identify more than 900 basic drugs and metabolites based on retention and scanning ultraviolet (UV) data (REMENDi HS, BioRad, Hercules, CA)(25, 26).

Recent improvements in interface design have led to the exciting prospect of LC-MS becoming a viable technology in a toxicology analysis, offering the capability to analyze polar compounds (many drug metabolites are) at pg/mL sensitivity level (27, 28)

An example of a TLC-based comprehensive drug screen is presented in Figure 22.1. The urine TLC screen is supplemented by a spot test for salicylate and four immunoassays. TLC does not adequately detect these analytes because these drugs are recovered poorly from urine extraction (salicylate and benzoylecognine) or present in low concentrations without prior hydrolysis of their conjugated metabolites (opiates and benzodiazepines). Urine that tests positive for one of the drugs

often will be processed for confirmatory testing. Quantitation in serum always is necessary for an alcohol; for some other drugs, serum levels may be informative and quantitation will be performed when indicated

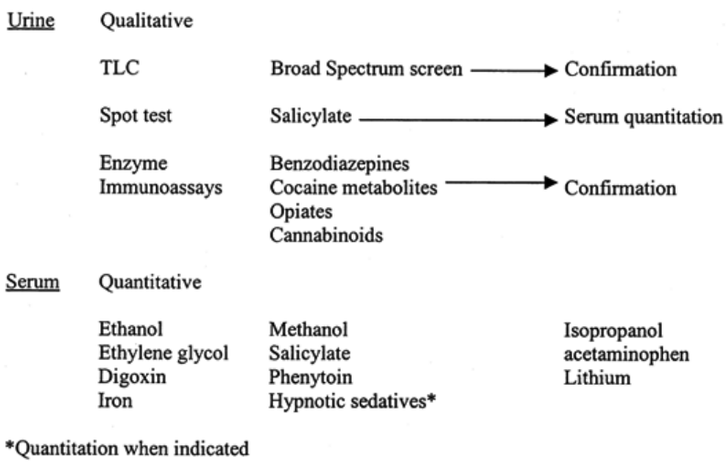


FIGURE 22.1. A thin-layer chromatography based comprehensive drug screen.

CLINICAL TOXICOLOGY OF SELECTED DRUGS

Part of "22 - Toxicology"

Ethanol

Ethanol is the most widely used social drug, and its abuse is one of the most important causes of injury and disease. Recent attempts to combat the problems of driving while under the influence of alcohol (DUI) have placed additional responsibilities on hospitals and clinical laboratories because the results of blood alcohol analyses, although obtained for medical reasons, may be used for legal purposes.

Ethanol usually is consumed in beers (3% to 6% ethanol), wines (10 to 12%), and distilled beverages (40% to 50%). The term “proof” means two times the percent of ethanol by volume. Ethanol is absorbed rapidly, and there is significant intersubject variability in the peak blood-alcohol concentration attained and the time for its achievement because of food intake and physiological variables (29). Ethanol distributes in total body water. Therefore, at equilibrium, ethanol concentration in any tissue or fluid is a function of the water content of that specimen. A plasma/blood ratio of 1.18 (range of 1.10 to 1.35), urine/blood ratio of 1.3, and breath/blood ratio of 2,407 (range of 1,981 to 2,833) have been determined (30, 31).

Ethanol is metabolized mostly by liver alcohol dehydrogenase to acetaldehyde at a rate that follows first-order kinetics at low concentration (less than 20 mg/dL). At higher concentration, metabolism proceeds at zero-order kinetics, which is independent of dose or initial blood-alcohol concentration. The average rate of metabolism is approximately 16 mg/dL/hr for men with significant variability (10 to 25 mg/dL/hr) and may be 10% to 20% lower for women (32). Individuals with high blood-ethanol concentration (greater than 3 g/L) have a rate of metabolism that is higher and more unpredictable (33).

Ethanol is a central nervous system (CNS) depressant. In intolerant subjects, ethanol-induced CNS dysfunction ranges from limited muscular incoordination at blood concentrations less than 50 mg/dL to coma, respiratory failure, and death at 400 mg/dL or higher. It is difficult to interpret blood levels in tolerant subjects; seemingly alert patients with ethanol levels in excess of 600 mg/dL have been reported.

Other diseases such as diabetic ketoacidosis and subdural hematoma can simulate the clinical signs of alcohol intoxication. Blood alcohols are ordered for the investigation of those patients who present with anion gap metabolic acidosis or those who are comatose. Medically, it may be desirable, or necessary, to know the patient's blood alcohol concentration before administrating anesthetics or medications.

Clinical toxicology laboratories are increasingly involved in the collection and analysis of blood for ethanol that subsequently has forensic implications. A set of guidelines has been developed in response to the need for information dealing with various aspects of blood-alcohol analysis (34).

Type of Specimen

Plasma or serum usually is analyzed in clinical laboratories. Most state laws on drinking and driving, however, define alcohol concentration in terms of whole blood. Alcohol concentration in whole blood is lower than in plasma or serum; the frequently used serum/whole blood ratio is 1.18 (34). The analysis of either serum or whole blood yields results with the same clinical significance. However, the difference between serum and whole-blood–ethanol concentration is significant in forensic analysis particularly when a “blood” alcohol result falls in the vicinity of a statutory threshold pertaining to drinking and driving. Conversion of results of alcohol analysis performed on serum or plasma to whole blood using a population mean ratio of 1.18 is discouraged; for legal purposes, it is best to analyze whole blood. Laboratories should identify the specimen type in their reports. Venipuncture sites should be cleansed by nonalcohol-containing disinfectant such as benzalkonium chloride or aqueous povidone-iodine. If analysis will be delayed for a few hours, sodium fluoride at 1% (higher for longer-term storage) is an effective preservative (34).

Urine is not a suitable specimen because blood-alcohol concentration cannot be established with sufficient reliability from

alcohol concentration of a pooled-bladder urine specimen and because of the great variability of the blood:urine ratio.

The use of breath and saliva for ethanol measurement has considerable appeal relative to blood (35, 36). Its advantages are the noninvasive nature of sample collection, the ease of test performance, and rapid turnaround time. Breath-alcohol test is based on Henry's Law, which states that in a closed system under constant temperature, the concentration of a gas above the liquid is proportional to the concentration of the gas dissolved in the liquid. The distribution of the gas in the gas and liquid phases is determined by the partition ratio. In the breath alcohol test, the assumption is that an equilibrium state exists between the alcohol in the blood perfusing the lung and that in the alveolar (breath). If the partition ratio for alcohol is known, blood-alcohol concentration can be determined indirectly by measuring alcohol concentration in breath. In forensic testing, a partition ratio of 2,100 has been adopted and used by measuring devices, although values have been reported to range from 1,981 to 2,833 (31). Careful considerations must be given to developing a quality-assurance program for POCT alcohol tests. Users of clinical breath-alcohol tests can benefit from the collective experience of the forensic community, which has extensive experience in breath-alcohol testing in traffic law enforcement (37).

The advantages of saliva testing are the noninvasive nature of sample collection, ready availability of saliva in sufficient quantity for analysis, and the ease of test performance (36). A device for measuring alcohol in saliva (STC Diagnostics, Bethlehem, PA) is based on enzymatic oxidation of alcohol by alcohol dehydrogenase. Saliva-alcohol results are in good agreement with alcohol concentrations in venous blood and end-expired breath, but insufficient sample volume was found to be a problem with highly intoxicated patients (36).

Analytical Methods

Many methods for alcohol analysis are in use in clinical laboratories (34). They can be grouped into the following categories:

1. Chemical oxidation of ethanol with acid dichromate.

   The reduction of dichromate is proportional to ethanol concentration. This is a nonspecific reaction; other alcohols and paraldehyde (as its metabolite acetaldehyde) also will give positive results. It is used in older breath analyzers.

2. Enzymatic oxidation using alcohol dehydrogenase (ADH).

$$C_2H_5OH + NAD^+ \xrightarrow{ADH} CH_3CHO + NADH + H^+$$

   The extent of interference by methanol and isopropanol depends on the source of ADH and can be minimized by careful choice of assay parameters. Enzymatic methods are simple and the most frequently used methods in clinical laboratories. Laboratories, however, must be prepared to investigate suspected ingestions of alcohols other than ethanol or the co-ingestion of ethanol and another alcohol.

3. Gas chromatography.

   Gas chromatography is the most specific method (38). It yields information on both the identity of the alcohol as well its concentration. For emergency analysis, direct injection of a diluted sample has the advantage that it eliminates the 15- to 30-minute equilibration period required for headspace analysis. The use of an internal standard (e.g., n-propanol) will improve precision.

Concentration Units

Many state laws on drinking and driving define blood-alcohol concentration as percent by weight/volume (% w/v). Clinical laboratories usually report alcohol results in concentration units of milligram alcohol per 100 mL (dL) of blood (or serum). An alcohol concentration of 100 mg/dL is equivalent to 0.1% w/v.

For breath alcohol testing in the clinical setting, the devices should display results in concentration unit of gram of alcohol per 100 mL of blood and not as gram of alcohol per 210 L of breath. Physicians and nurses are accustomed to reviewing medical laboratory test results in concentration units based on 100 mL of blood (or serum).

Methanol and Isopropanol

Methanol and isopropanol are important industrial chemicals that also are available as household items: methanol as a constituent of some antifreeze and windshield washer solutions; isopropanol as a disinfectant (30% to 99.9% solution) or as rubbing alcohol (70% solution). Intoxication with methanol and isopropanol can be from accidental ingestion, industrial exposure, self-poisonings, or as substitutes for ethanol.

Both methanol and isopropanol are absorbed readily following ingestion. They are metabolized by hepatic alcohol dehydrogenase at rates one tenth or less of that for ethanol. Methanol is oxidized to highly toxic formaldehyde and then to formic acid. Formic acid is much more toxic than methanol and accounts for the profound anion gap metabolic acidosis and its ocular toxicity. Formate levels (generally not available in clinical laboratories) are better correlated with poor outcome (39). Methanol toxic symptoms may include inebriation, headache, dizziness, seizure, and coma. Nausea, vomiting, stiff neck, abdominal pain, and malaise also are common complaints. After ingestion, methanol serum level peaks at about 2 hours, but there may be a latent period of about 24 hours when there is a deceiving lack of severe toxic manifestation but during which appropriate treatment is critical (40).

Isopropanol is metabolized to acetone, which accounts for the CNS effects and ketonemia. Toxic manifestations include CNS depression, lethargy, weakness, hypotension, and abdominal pain. Hemorrhagic tracheobronchitis and gastritis are characteristic findings. Clinical laboratory findings include ketonemia, ketonuria, and elevated osmolality. There is no metabolic acidosis. Treatment of isopropanol overdose mainly is supportive and most patients respond well. Patients with very high isopropanol levels (400 to 500 mg/dL) usually are associated with severe hypotension and coma and hemodialysis should be considered (40).

Because the toxicity of methanol is because of ADH-generated toxic metabolites, treatment involves competitive inhibition of ADH with a saturating concentration (100 to 150 mg/dL) of ethanol, the preferred substrate, and allowing the kidneys to eliminate methanol provided adequate urine flow is maintained (40).

Hemodialysis is indicated if methanol level is greater than 25 mg/dL and should be performed until methanol reaches undetectable levels. 4-Methylpyrazol, a specific inhibitor of ADH, is undergoing a clinical trial to assess its efficacy as an antidote for methanol and ethylene glycol poisoning (41).

The method of choice for the identification and measurement of methanol and isopropanol is GC. The GC methods for ethanol generally are also applicable to these two alcohols if acetone is adequately resolved from the alcohols (38). Low methanol levels are seen periodically in alcoholics after binge drinking. Ethanol inhibits methanol metabolism, allowing the accumulation of a small amount of methanol originating from endogeneous production or consumption of fermented beverages. As many as 6% of patients with ethanol levels greater than 100 mg/dL had serum-methanol levels of 4.5 mg/dL or more (42).

The popular enzymatic assay for ethanol will not detect methanol and isopropanol because of the weak enzyme activity when these alcohols are substrates. Toxic serum levels of these osmotically active substances, however, will result in a significant osmolal gap between measured osmolality and calculated serum osmolality, thus allowing the evaluation of an acute situation when specific assays for methanol and isopropanol are not available. A formula to estimate serum osmolality is (43):

[1.86 × Na (mmol/liter) + Glucose (mg/dl)/18 + BUN (mg/dl)/2.8] ÷ 0.93

The calculated osmolality is divided by 0.93 because serum is approximately 93% water. The measured osmolality must be determined by freezing-point depression osmometry (not by using a vapor-pressure osmometer). The expected contribution to measured osmolality of a serum-methanol concentration of 100 mg/dL is 31 mOsm/kg. Ethanol is the most common cause for elevation of serum osmolality. Users of this approach must be aware of its nonspecificity because other alcohols, acetone, and ethylene glycol also can raise serum osmolality if present in sufficient concentration.

Ethylene Glycol

Ethylene glycol is the principal component of antifreeze products and brake fluids. It is absorbed readily in the gastrointestinal tract and is metabolized by oxidization to glycolic acid, glyoxylic acid, and formic acid by alcohol dehydrogenase and glycolic acid oxidase (40). Ethylene glycol itself is only mildly toxic; its metabolites, however, are highly toxic (44). Persistent vomiting, and gradual onset of CNS depression appear within 4 to 8 hours and are accompanied by a large anion gap metabolic acidosis. A common urinalysis finding is the presence of calcium oxalate or hippurate crystals, which can lead to acute renal tubular necrosis developing within 12 hours to 2 days.

Treatment of ethylene-glycol poisoning consists of correction of metabolic acidosis, inhibition of metabolism by ethanol or 4-methylpyrazol, and hemodialysis. Hemodialysis rapidly clears both ethylene glycol and its toxic metabolites from the bloodstream and is recommended for patients who are symptomatic or have blood levels of ethylene glycol greater than 25 mg/dL (40).

There is no easy assay for ethylene glycol, and most clinical laboratories do not provide ethylene glycol analysis. Serum osmolality measurement and the calculation of the osmolal gap can provide a rough approximation of ethylene glycol concentration, but only if ethanol and other alcohols are known to be absent (43).

Early gas chromatographic methods, including those based on direct injection, were plagued by “trailing peaks,” “ghost peaks,” variable recovery, and interference by propylene glycol or 2,3-butandiol (present in serum of some alcoholics) (45). A recent GC procedure has improved precision and accuracy, and also can measure glycolic acid (46).

An enzymatic method is based on the action of glycerol dehydrogenase on ethylene glycol is rapid and easy to perform (47). The enzyme assay will not detect propylene glycol and it is not specific for ethylene glycol; known interfering substances include glycerol and 2,3-butandiol (48). Specimens containing high lactate dehydrogenase activity or lactate level will give false-positive results (49). For laboratories unable to maintain a chromatographic assay for ethylene glycol, the enzyme assay, despite its specificity limitation, is a clinically useful alternative. It has high negative predicative value in eliminating ethylene glycol from the list of etiologies causing unexplained metabolic acidosis.

Acetaminophen

Acetaminophen is an effective analgesic and antipyretic drug that lacks antiinflammatory action. It is available in pure form and also in combination with other drugs such as codeine and propoxyphene. It presents less risk for producing gastrointestinal ulceration and hemorrhage than aspirin and other nonsteroidal antiinflammatory drugs. With the reported link of Reye's syndrome to aspirin use, usage of acetaminophen as an over-the-counter medication has surpassed that of aspirin in recent years. At typical nonprescription doses of 325 to 1,000 mg every 4 hours, acetaminophen is a safe drug. At higher doses, acetaminophen is hepatotoxic, although toxic dosage is variable; liver damage may occur after single doses of 7.5 g or greater in healthy adults or 150 mg/kg in children (50). Individuals on other medications that are hepatotoxic or those who have liver diseases or who are alcoholics, however, will be susceptible to acetaminophen toxic effects at lower doses (51).

Acetaminophen after a therapeutic dose is eliminated mostly as glucuronide or sulfate conjugates. Following overdose, the conjugation pathways are saturated and formation of a highly reactive intermediate (probably N-acetyl-benzoquinoneimine) takes place via the cytochrome P-450 mixed-function oxidase system (52). This metabolite, normally detoxified by endogenous glutathione, reacts with and destroys hepatocytes once glutathione stores are depleted.

Acetaminophen is absorbed rapidly with peak plasma concentration reached within 30 to 120 minutes after therapeutic doses. Delayed peaks may occur with slower gastric emptying following large doses. Clinically, a patient overdosed on acetaminophen may present in four phases (50). In the initial phase, lasting 12 to 24 hours after ingestion, the patient usually exhibits gastrointestinal (GI) irritability, nausea, and vomiting. Some patients may be asymptomatic. During the second phase (24 to 72 hours postingestion) the patient may feel reasonably well while liver function tests prove abnormal. If significant hepatic

necrosis has occurred, the third phase (72 to 96 hours postingestion) is characterized by the sequelae of hepatic necrosis including coagulopathy, jaundice, and encephalopathy. If the patient survives phase 3, complete resolution of hepatic dysfunction will ensue (4 days to 2 weeks).

N-Acetylcysteine is an effective antidote. In the United States, the standard oral regimen consists of a loading dose of 140 mg/kg followed by 17 doses of 70 mg/kg every 4 hours. Treatment is most successful when started within 8 hours of ingestion, and its effectiveness appears to extend to those high-risk patients who are treated as late as 24 hours post ingestion (53). In Canada and Britain, intravenous administration over a 20-hour period is the approved protocol (150 mg/kg loading dose over 15 minutes followed by an additional dose of 50 mg/kg over 4 hours and then 100 mg/kg over 16 hours) (54). A nomogram relating time since ingestion, plasma drug concentration, and risk of hepatotoxicity is used in evaluating the need for *N*-acetylcysteine treatment (Fig. 22.2) (55). The nomogram is for use with acute ingestion only. It cannot be used to assess toxicty resulting from chronic misuse of acetaminophen. Because early treatment is critical to a favorable outcome, and initial plasma-drug concentration is a crucial deciding factor to initiate therapy, prompt and reliable measurement of plasma-acetaminophen level is an important emergency toxicology service. Numerous methods are available for the analysis of acetaminophen. Colorimetric tests such as the cresol-ammonia spot test (11) or the Glynn and Kendal method are fast and sensitive although the Glynn and Kendal method is interfered by salicylates. The colorimetric method based on the prior hydrolysis of acetaminophen and its conjugated metabolites to indophenol is

not recommended because the aforementioned nomogram is based on serum concentration of unconjugatd acetaminophen only (56). HPLC procedures, though simple and rapid, are not as convenient as the FPIA and EMIT methods that are available in most clinical laboratories.

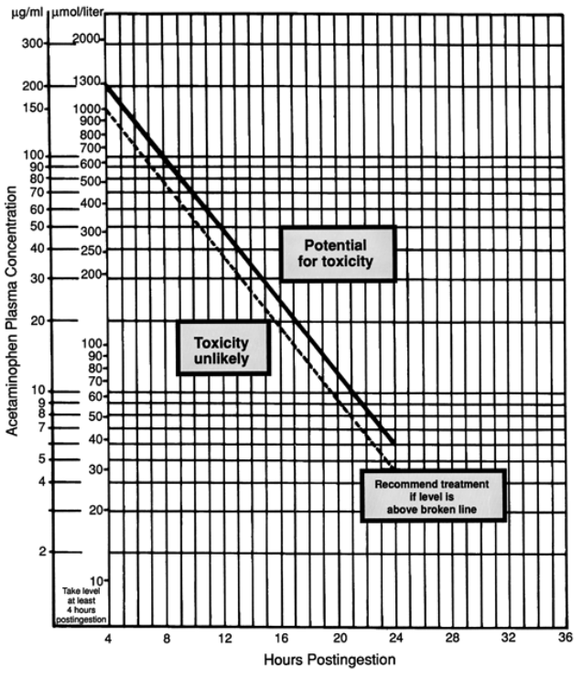


FIGURE 22.2. Nomogram relating acetaminophen plasma concentration, time since ingestion, and risk of toxicity. (Modified from Rumack BH, Matthew H. Acetaminophen poisoning and toxicity. *Pediatrics* 1975;55:871–876.)

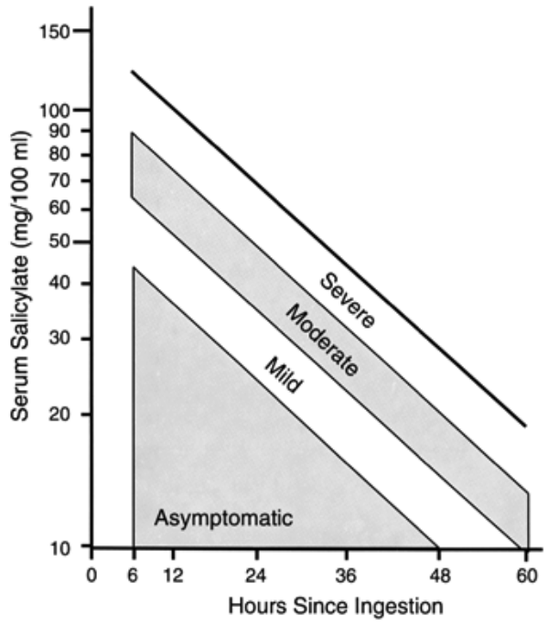


FIGURE 22.3. Nomogram for salicylate poisoning. (From Done AK. Salicylate intoxication. Significance of measurements of salicylate in blood in cases of acute ingestion. *Pediatrics* 1960;26:800–807.)

Salicylates

Salicylate, as one of the least expensive and most widely used drugs, has been the cause of many drug overdose cases, particularly in the very young and the elderly. Salicylate still ranks as a leading cause of childhood poisoning deaths and still is commonly used in self-poisoning by adults (57, 58). Many derivatives of salicylic acid are available commercially; the most important is acetylsalicylic acid (aspirin), which is hydrolyzed rapidly to salicylic acid, and circulates in the blood in the ionized form, salicylate. In serum, salicylate at therapeutic concentration is highly protein bound, and the extent of binding varies with total salicylate concentration. The major metabolic pathways of salicylate conjugation with glycine and with glucuronic acid are saturable. Thus, the half-life of elimination of salicylate as well as the serum level of salicylate will increase disproportionately with increasing dosage (57).

The primary pathophysiological effects of salicylism are complex (59). They include direct stimulation of the respiratory center resulting in hyperventilation, respiratory alkalosis and compensatory excretion of base, uncoupling of oxidative phosphorylation, interference with the Krebs cycle, and accumulation of organic acids leading to metabolic acidosis. In children, respiratory alkalosis is transient and a late-stage dominant metabolic acidosis is common. In adult patients, the most common acid-base disturbance is mixed respiratory alkalosis and metabolic acidosis. Associated with acid-base disturbances are fluid and electrolyte imbalance and dehydration. Other metabolic effects of salicylate toxicity are hyperthermia and impaired glucose metabolism with either hyperglycemia or hypoglycemia. In overdosed patients who are acidemic, the lower blood pH will increase the amount of nonionized salicylate (pKa = 3.0) for transfer into the CNS. Thus, CNS disturbances often accompany those intoxicated patients who are severely acidemic.

The toxic severity after acute ingestion is related to the amount of drug ingested. Ingestion of less than 0.15 g/kg is unlikely to result in toxic symptoms. Mild to moderate toxic reactions can be expected from an ingested dose of 0.15 to 0.39 g/kg and doses more than 0.3 g/kg lead to severe reactions. Ingestion of greater than 0.5 g/kg is potentially lethal. Chronic intoxication or therapeutic overdose is a result of excessive therapeutic administration of salicylate over a period of 12 hours or longer, and zero order kinetics lead to accumulation of salicylate in serum to toxic level (60). Chronic salicylate intoxication is a diagnostic problem particularly among elderly patients because the presenting symptoms often are ascribed to other causes. Thus, the intoxication often goes unrecognized, and as a result, appropriate therapy is delayed. These patients suffer significant morbidity and mortality.

Treatment for salicylate intoxication includes measures to prevent further absorption of salicylate and to correct the metabolic imbalances such as fluid and electrolyte depletion and acid-base disturbances. Alkalization with sodium bicarbonate to enhance the renal elimination of the drug should be considered in adults with serum salicylate level exceeding 50 mg/dL and children with levels greater than 35 mg/dL. If the salicylate level is greater than 100 mg/dL, and the patient is very ill or is unable to eliminate the salicylates, hemodialysis is indicated (59).

The availability of salicylate assay on an emergency basis is critical to confirm the clinical suspicion of acute salicylate intoxication. Diagnosis of chronic salicylate intoxication, particularly in the elderly, is much more difficult without a high degree of suspicion because patients may have become drowsy and confused and are unable to offer a reliable drug history. Therefore, documentation of elevated serum salicylate levels becomes very important in the differential diagnosis (57).

A nomogram (Done's nomogram) (62) has been constructed to facilitate the interpretation of salicylate levels at different intervals after ingestion in order to predict the severity of intoxication (Fig. 22.3). The clinical value of the nomogram is limited because it was based on data from a pediatric population after a single acute ingestion of regular formulation (63). It does not predict the rate of salicylate elimination or future serum salicylate levels and it is not applicable for assessing the severity of chronic salicylate intoxication.

If the ingested salicylate is of an enteric-coated or sustained-released formulation, the absorption of salicylate will be delayed. The diagnosis of salicylate intoxication based on an elevation of serum level on admission can be missed because of delayed absorption (64). If ingestion of enteric-coated or sustained-release salicylate is suspected, the patient must be observed for at least

24 hours, and serum salicylate determination should be repeated. Peak salicylate level may not be attained until 60 to 70 hours postingestion.

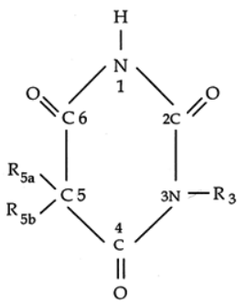
Simple qualitative screening tests such as ferric chloride and Trinder's reagent for urine salicylate are useful for quick confirmation of salicylate overdose if quantitation of serum levels is not available immediately. All positive results should be confirmed using quantitative assays and serum samples (57).

The majority of salicylate assays in use are colorimetric assays based on the reaction of salicylic acid with ferric ion in acid medium to give a purple color complex (56, 57). This simple and rapid reaction is not specific for salicylate; metabolites of salicylate, ketone bodies, catechols such as tyrosine and the catecholamines Ketone bodies and catecholamines frequently are elevated in disease states that are associated with metabolic derangement, such as ketoacidosis and Reye's syndrome. Salicylate concentrations of these patients that are measured by a colorimetric method can be falsely high (65). Colorimetric assays have proven acceptable for the routine clinical use of salicylate levels in the diagnosis of salicylate intoxication. The assay based on Trinder's method (66) is the most commonly used, although newer techniques such as FPIA, HPLC, and enzymatic assay using salicylate hydrolase are available.

A different approach to the measurement of salicylate is the use of the enzyme salicylate hydrolase (EC 1.14.13.1) purified from *Pseudomonas cepacia* (67, 68 and 69). This enzyme, in the presence of reduced nicotinamide adenine dinucleotide (NADH) or reduced nicotinamide adenine dinucleotide phosphate (NADPH), converts salicylate to catechol. Salicylate concentration can be determined by monitoring, photometrically, the consumption of NADH (67) or the formation of catechol (68, 69). Compounds that are known to interfere with the colorimetric assays do not affect the enzymatic methods.

Barbiturates

Barbiturate poisoning either by accident or more often in suicide attempts was a major health problem in the past. It has lessened in recent years because their use as anxiolytics has been replaced by the safer benzodiazepines; medical use of barbiturates now is primarily for treatment of insomnia and convulsive disorders, and as anesthetic and preanesthetic medications. The barbiturates are derivatives of barbituric acid (Fig. 22.4). Pharmacologically, they can be divided broadly into ultra-short (e.g., thiopental), short- to intermediate- (e.g., secobarbital and butalbital), and long-acting groups (e.g., phenobarbital) depending on their duration of action (70). Drug abusers prefer the short- to immediate-acting barbiturates because of the relative rapid onset of action (15 to 40 minutes). Duration of drug action, which may last up to 6 hours, is dependent on lipid solubility. Thus, the short-



Generic (trade) name	R5a	R5a	t1/2(h)	pKa
Methohexital(Brevital)	-CH2-CH=CH2	-C(CH3)-C°CH2-CH3	1-2	7.9
Thiopental(Pentothal)	-CH2-CH3	-CH(CH3)-CH2-CH2-CH3	4-60	7.6
Amobarbital (Amytal)	-CH2-CH3	-CH2-CH2-CH(CH3)2	15-40	7.9
Pentobarbital(Nembutal)	-CH2-CH3	-CH(CH3)-CH2-CH2-CH3	20-30	7.9
Secobarbital(Seconal)	-CH2-CH=CH2	-CH(CH3)-CH2-CH2-CH3	22-29	7.9
Butabarbital(Butisol)	-CH2-CH3	-CH(CH3)-CH2-CH3	34-42	7.9
Butalbital(Fiorinal)	-CH2-CH=CH2	-CH2-CH(CH3)2	35-88	7.6
Phenobarbital(Luminal)	-CH2-CH3	-C6H5	48-144	7.2

R3 = H, except methohexital when R3 = CH3

FIGURE 22.4. Common barbiturates.

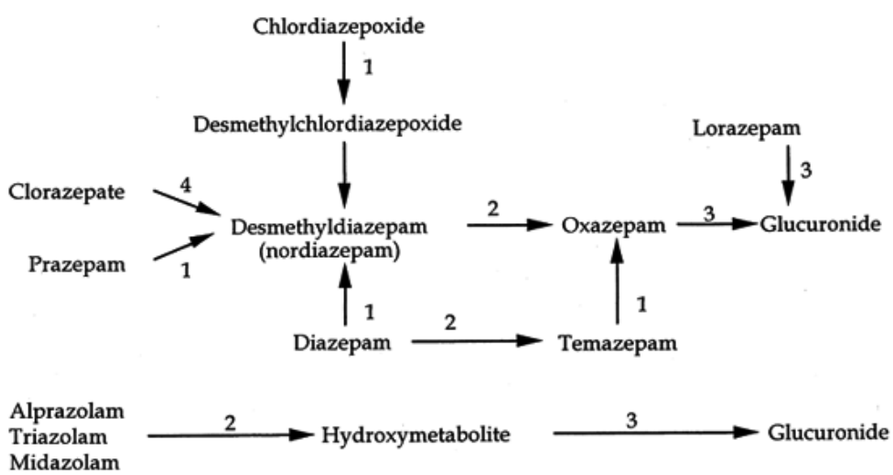


FIGURE 22.5. Major biotransformation pathways of common benzodiazepines. 1, dealkylation; 2, hydroxylation; 3, glucuronidation; 4, decarboxylate.

acting barbiturates have higher lipid solubility, greater potency, and more rapid clearance from the central compartment. In contrast, the long-acting barbiturates have less lipid solubility, lower potency, and much longer half-lives of elimination.

The barbiturates are weak acids with pKa values ranging between 7.2 and 7.9. Hence, at pH 7.4 a barbiturate such as phenobarbital (pK = 7.2) will be about 40% nonionized, whereas secobarbital (pK = 7.9) is 76% nonionized and therefore is more lipid soluble and more rapidly distributed to tissues.

The major actions of the barbiturates are their depressant action on the CNS and cardiovascular system. Acute barbiturate intoxication is characteristically associated with coma and shock; the former must always be differentiated from other forms of coma or CNS injury. Plasma barbiturate levels may be helpful in making a diagnosis, but they are of limited value in predicting the severity of the overdose or the duration of the coma, which are related more closely to brain than plasma barbiturate concentration. In addition, chronic barbiturate users are expected to have higher plasma concentration for any grade of coma because of the tolerance they have developed. Moreover, the depth of coma can be greater than might be expected from plasma concentration if other CNS depressants, such as alcohol, also have been ingested.

Treatment of barbiturate intoxication consists of aggressive support to combat shock and hypoxia in addition to lavage and charcoal treatment. Alkalization of urine to pH 7.5 to 8.0 with sodium bicarbonate will increase the fraction of ionized drug and will enhance excretion in urine. This is helpful for long-acting barbiturates because their principal means of elimination is renal. Alkalization is less effective for short-acting barbiturates because they have higher pKa values, are more highly protein-bound, and they are metabolized primarily by the liver (70).

Immunoassays detect barbiturates as a group and provide rapid results. Barbiturates other than the one used as the assay calibrator may have different cross-reactivities to the assay antibody. For example, an enzyme immunoassay that uses secobarbital as the calibrator will detect the less reactive phenobarbital only at much higher concentrations.

Identification of the specific barbiturate(s) requires chromatographic separation. TLC can effectively separate phenobarbital from other barbiturates, but the short- and intermediate-acting group (amobarbital, butabarbital, butalbital, secobarbital, and pentobarbital)– have similar migration, and identification requires skill and experience (18). Gas chromatography and HPLC methods for identification of the commonly encountered barbiturates are available (71, 72), and mass spectral identification of the barbiturates must be done with care because these drugs are structurally similar and yield similar spectral data (73).

The benzodiazepines are among the most frequently prescribed drugs. They vary in their potency in anxiolytic, hypnotic, muscle relaxant, anticonvulsant, and anesthetic effects. One useful classification of the benzodiazepines is according to their half-lives: long (greater than 24 hour), intermediate-short (5 to 24 hours) and ultra short-acting (less than 5 hours) (Table 22.2). A long-acting benzodiazepine such as diazepam is well suited as an anxiolytic agent, whereas a short-acting benzodiazepine such as triazolam is more appropriate as a hypnotic (70).

TABLE 22.2. COMMON BENZODIAZEPINES

Generic Name	Trade Name	Half-life (hr)
Diazolobenzodiazepines		
Chlordiazepoxide	Librium	6–27
Diazepam	Valium	21–37
Oxazepam	Serax	4–11
Clorazepate	Tranxene	2
Flurazepam	Dalmane	2–3
Lorazepam	Ativan	9–16
Temazepam	Restoril	3–13
Halazepam	Paxipam	14–16
Prazepam	Centrax	1.3
Triazolobenzodiazepines		
Alprazolam	Xanax	10–12
Trialzolam	Halcion	2.6
Midazolam	Versed	1–4

Benzodiazepines also are useful for treating neuromuscular diseases, sleep disorders, seizure, drug withdrawal, and as pre-anesthetic agents. Continuous use has led to tolerance, addiction, dependence, and abuse (74).

All the benzodiazepines are metabolized and cleared by the kidney as glucuronides or sulfates. Many of the metabolites are active, and some benzodiazepines such as clorazepate and flurazepam are considered to be prodrugs (Fig. 22.5). The major biotransformation routes are demethylation (e.g., chlordiazepoxide, temazepam) and hydroxylation (diazepam, nondiazepam).

With the popularity of these drugs, overdose is a frequent occurrence, yet fatalities resulting from benzodiazepines alone are very rare (75). Patients intoxicated with benzodiazepines are sedated and side effects include respiratory depression, weakness, headache, blurred vision, nausea, and diarrhea. Serious intoxication with benzodiazepines usually results from co-ingestion with other CNS depressants (e.g., ethanol). Therefore, in the evaluation of a patient suspected of benzodiazepine intoxication, it is important to investigate if other drugs also have been ingested. Plasma benzodiazepine concentration does not predict the severity of intoxication or outcome. A specific benzodiazepine antagonist, flumazenil, has been reported to effectively reverse benzodiazepine-induced CNS symptoms. Its benefit, however, must be weighed against the risk of precipitating acute benzodiazepine withdrawal, or seizure among patients who also may have taken seizure-causing drugs such as the cyclic antidepressants (76).

Benzodiazepine levels in blood and plasma can be determined by HPLC or GC assays (77, 78). Some benzodiazepines, such as chlordiazepoxide and oxazepam, are heat-labile, and are more suitable for HPLC analysis. GC methods using electron capture or mass spectrometer detection have the best sensitivity capable of detection of triazolo-benzodiazepines at low ng/mL levels, whereas those dependent on FID detection are adequate for drug levels in the high therapeutic or toxic ranges. TLC does not easily detect benzodiazepines without prior acid hydrolysis to form benzophenones, which are strongly fluorescent (79). Specific identification is not possible because different benzodiazepines can form the same benzophenone. Also, some benzodiazepines

(e.g., alprazolam and triazolam) are stable to acid hydrolysis and are not converted to benzophenones.

Commercially available immunoassays for benzodiazepines in urine or serum are designed to detect oxazepam or nordiazepam, metabolites that are common to several benzodiazepines (Fig. 22.5). Immunoassays have sufficient cross-reactivity with many of the newer benzodiazepines to detect therapeutic levels of some (e.g., alprazolam), but only higher levels for others (e.g., triazolam) (78, 80, 81).

Cyclic Antidepressants

The cyclic antidepressants are a major cause of life-threatening drug overdose, and are responsible for more deaths than any other drug classes except the analgesics (58). Drugs in this class of commonly prescribed antidepressants include the tricyclic compounds, such as imipramine, desipramine, trimipramine, amitriptyline, nortriptyline, doxepin, and loxapine, the tetracyclic compounds maprotiline and miaserin, and the bicyclic fluoxetine. Trazadone, a triazolopyridine, and amoxapine, a dibenoxapine, while structurally unrelated to the tricyclic compounds, often are classified or described with them. These cyclic compounds are rapidly absorbed by the GI tract, but absorption may be delayed at higher blood levels as a result of the anticholinergic effect of delayed gastric emptying. The volumes of distribution of these drugs are large (10 to 50 liters/kg) and extent of binding to plasma protein is high (greater than 90%). There is extensive first-pass metabolism, and the major hydroxylated metabolites of the tricyclic and demethylated metabolites of the bi- and tetracyclic compounds are pharmacologically active. The toxic effects of the tricyclic and tetracyclic antidepressants are similar and are mostly related to their anticholinergic and cardiotoxic effects. Anticholinergic effects result in agitation, seizure, coma, and hallucination. Circulatory collapse, serious cardiac arrhythmia, conduction disturbances, and heart block can occur; arrhythmia is the leading cause of death in tricyclic overdoses (82). The toxicity of trazadone differs from that of the tricyclic and tetracyclic antidepressants in that there are no anticholinergic signs or symptoms and the primary manifestations are CNS effects and hypertension. Fluoxetine adverse effects are anxiety, nervousness, and insomnia.

Life-threatening events following a tricyclic overdose usually occur within the first 6 hours of ingestion (83). A patient is at low risk for toxicity if he has not developed the following within 6 hours of ingestion: QRS interval > 100 msec, arrhythmia, altered mental status or seizures, respiratory depression, or hypotension (84).

Tricyclic antidepressant-intoxicated patients who had peak (not admission) serum drug levels greater than 1,000 ng/mL were associated with seizures and ventricular arrhythmias (85). There was no correlation, however, between drug levels and toxic symptoms over a wide range of tricyclic antidepressant levels including those below 1,000 ng/mL and drug levels did not predict severity of symptoms (83). Therefore, measurement of serum tricyclic antidepressant levels, particularly if requested on a “stat” basis, to predict the severity of acute tricyclic antidepressant poisoning is unwarranted.

The tricyclic antidepressants can be detected and identified in urine by TLC or one of the tricylic antidepressant group-specific immunoassays. Because the toxic effects resulting from an overdose of these compounds are so similar, demonstration of the presence of one of these drugs without specific identification often is sufficient for management. Immunoassays (EIA, FPIA) are designed primarily for detection of the major tricyclic antidepressants (imipramine, desipramine, amitriptyline, nortriptyline) but, because of the cross-reactivities of the antibody for other members of the tricyclic antidepressant family, these immunoassays also are capable of detecting other tricyclic antidepressants and metabolites (86). A single-use device designed to test for tricyclic antidepressants and several drugs of abuse has been evaluated for use in the emergency department (13). Although a positive result does not necessarily indicate a toxic concentration of the drug, a negative response will rule out the possibility of an overdose. Quantitative immunoassays (EMIT) for the four major tricyclic antidepressants also are available. These assays require solid-phase extraction from serum or plasma before analysis. Quantitative determination of serum levels is performed best using HPLC (77, 87).

Cyclobenzaprine is prescribed extensively as a muscle relaxant. It is a tricyclic compound, structurally differing from amitriptyline only in having a double bond in the cycloheptane ring. Analytical distinction between the two is difficult—they are indistinguishable in immunoassay cross reactivity and chromatographic

elution on TLC or HPLC. Distinction can be achieved by HPLC using a photo-diode detector to exploit differences in their UV spectra (88) or by careful examination of their different fluorescence characteristics in Stage III of the TOXILAB TLC system (89)

Iron

Iron supplements in various forms of iron salts are readily available and ferrous sulfate, the cheapest and most common iron salt, is involved frequently in overdose. Other iron salts are gluconate, fumarate, succinate, lactate, chloride, ferrocholinate, and glutamate. The elemental iron dose ingested can be calculated from the percentage of elemental iron in the formulation (Table 22.3). Acute iron poisoning is particularly common in the pediatric population, with the majority of the reported exposure occurring in children less than 6 years of age (58).

TABLE 22.3. ELEMENTAL IRON EQUIVALENTS IN IRON SALTS

Salt	% Fe	mg Fe/Tablet[a]
Sulfate	20	65
Gluconate	12	38
Fumarate	33	106
Lactate	19	-
Chloride	28	-
Ferrocholinate	13	-

[a] 325 mg tablet

Fe, iron

In acute iron poisoning, an estimation of the amount of elemental iron is important in assessing potential toxicity. An ingestion of 10 to 20 mg of elemental iron per kilogram of body weight can be associated with toxicity (90). A dose of 20 to 60 mg/kg can pose moderate risk, whereas a dose greater than 60 mg/kg has high risk for serious toxicity.

Five clinical stages of acute iron poisoning have been described. The first stage is the initial postingestion period lasting up to 6 hours. As iron is absorbed in the duodenum and upper small intestine, GI symptoms manifested at this stage are a direct result of the corrosive action of iron; the absence of these symptoms during this period excludes serious injury. During the next stage (latent period) of up to 12 hours, the patient may appear to improve. If the dose ingested is sufficiently large, the patient's condition may progress directly to a third stage (12 to 48 hours postingestion) of systemic toxicity with cardiovascular collapse, seizure, coma, and shock. The fourth stage is hepatic failure 2 to 3 days postingestion. Some patients who survive are still at risk of developing late strictures and some GI tract obstruction (2 to 4 weeks) (90).

Ingestion of iron can be documented by abdominal radiograph, which can reveal radiopaque iron-containing pills prior to dissolution as much as 6 hours or longer after ingestion of adult-strength tablets. Pediatric preparations, which are chewable iron supplements, dissolve rapidly in 30 to 60 minutes and are not always seen (91).

A qualitative test to predict potential for toxicity is the deferoxamine challenge test. A sufficiently large deferoxamine dose (90 mg/kg) is given to bind free iron in plasma and forms the reddish feroxamine complex, which is excreted in urine. The appearance of feroxamine in urine within 4 to 6 hours implies potential toxicity. False-negative results have been reported, and a negative challenge test should not rule out toxicity (92).

Serum iron level usually is determined along with total iron-binding capacity (TIBC). Serum iron levels between 300 and 500 µg/dL are associated with significant GI toxicity. Patients with serum iron levels above 500 µg/dL can exhibit systemic toxicity and shock, and will require chelation therapy. Sampling of blood for determination of serum iron should be on admission and again at 4 to 6 hours postingestion when peak levels are thought to occur.

Theoretically, if the serum iron value is greater than the TIBC, then the free circulating iron is potentially toxic. The use of TIBC is discouraged because of spurious elevation of TIBC following iron poisoning from a laboratory artifact (93).

The management of acute iron poisoning includes removal of residual iron in the GI tract by emesis or lavage and standard support therapy (90). Deferoxamine therapy is used to chelate free serum iron.

Most methods for measuring serum iron levels are based on spectrophotometric measurement of a color complex formed by the chelation of ferrous iron to a dye. Deferoxamine, as a competing chelator, interferes with dye-binding calorimetric assays to give falsely low results (94). Therefore, blood for serum iron levels should be drawn before deferoxamine chelation therapy. Atomic absorption spectroscopy for serum iron levels is not routinely available but is a method that can be used in the presence of deferoxamine.

Cannabinoids

The term cannabinoids denotes a group of more than 60 compounds found in the plant *Cannabis sativa L*. Δ^9-Tetra-hydrocannabinol (THC) is the major psychoactive compound of marijuana.

The effects of THC, when smoked, appear within minutes and seldom last longer than 2 or 3 hours. Oral intake delays the onset of symptoms for 30 minutes to 2 hours but the duration of drug action is longer. The effects of THC are tempered by dose, route of administration, and experience of the user. There usually is a sense of euphoria, an altered perception of time, a keener sense of hearing, and more vivid visual imagery. Both short-term memory and task performance are impaired. Higher doses can induce frank hallucinations, delusions, and paranoid feelings. The most consistent cardiovascular effects are an increase in pulse rate and conjunctival reddening (95).

THC, a lipophilic drug with a large volume of distribution estimated to be about 10 L/kg, is sequested in certain organs such as liver and lung. In man, THC is transformed rapidly first by the hepatic cytochrome P450 enzyme system to 11-hydroxy-Δ^9-THC (11-hydroxy-THC) which then is oxidized by alcohol dehydrogenase to 11-nor-Δ^9-THC-9-carboxylic acid (THCA). 11-hydroxy-THC is a psychoactive metabolite, whereas THCA is devoid of psychoactivity (95).

The primary urinary metabolite is THCA, which exists both as the free acid and glucuronide conjugate. Because of the large tissue storage of THC, continuous reentry of THC from tissues into the central compartment followed by metabolism means

that THCA is excreted into the urine long after a person has stopped using marijuana. The duration of urine samples testing positive is dependent on the dose, metabolism of THC, route and frequency of use, timing of urine collection, quantity of liquid taken prior to specimen collection, and assay cutoff. Generally, infrequent users test positive from 2 to 5 days after each dose when a 20-µg/L cutoff is used, but heavy users have been reported to test positive for over 46 days (96, 97). Thus, a positive urine result indicates only that the subject has been exposed to marijuana hours to weeks before the collection of the urine specimen, and the result cannot be used to estimate the time of exposure. Because of variation in a patient's hydration status, serial tests may see fluctuation of test results between positive and negative. A positive result following a negative one could be a result of residual excretion or dehydration and not reuse of marijuana. Serial determination of urine THCA concentrations normalized to urine creatinine can be helpful in interpretation (98, 99).

Because marijuana is smoked frequently in social situations, it is possible that nonsmokers could inhale a sufficient amount of cannabinoids present in sidestream smoke to produce a positive urine cannabinoids test. Studies have shown that this can occur, but only after exposure to high concentrations of marijuana smoke in a small, unventilated area (100). Such extreme exposure conditions are not encountered in the usual social situations, and with higher cutoffs (e.g., 50 or 100 µg/L) used by many screening assays, the detection of a positive urine from passive inhalation is unlikely (101).

Screening usually is done with immunoassays as they have good sensitivity and they can be automated for batch analysis of large numbers of samples. The popular commercial immunoassays use antibodies to detect THCA, the major urinary metabolite, and many of the other metabolites of THC. These assays, therefore, measure the sum of the immunoreactive THC metabolites. In contrast, the chromatographic assays (TLC, GC, and GC-MS) separate and specifically detect and quantify THCA. The threshold concentrations of these confirmation chromatographic methods usually are set lower than those of the initial tests. For example, in the Federal Workplace Drug Testing Programs, a confirmation threshold of 15 ng/mL is used when the immunoassay cutoff is 50 ng/mL (102).

If a chromatographic assay is used, hydrolysis of the glucuronide metabolites is necessary. TLC has a detection limit as low as 10 µg/L, whereas GC-MS methods, using a deuterated internal standard and selected ion monitoring, can have a limit of quantitation of 5 ng/mL and lower.

Hemp and marijuana both are *Cannabis sativa L*. Hemp, cultivated to produce hemp fiber and seeds for commercial use, generally contains relatively small amount of Δ^{9}-THC. Recently, food products derived from hemp seeds and hemp seed oil have become popular health food items and are easily available on the internet. These products contain varying amount of Δ^{9}-THC (103) and ingestion of these products can lead to positive drug test results for marijuana (104, 105). In one study, volunteers ingested 11 or 22 g of hemp seed oil containing high Δ^{9}-THC concentration (1,500 µg/g of oil). Urine specimens collected were positive for the marijuana test for 6 days with Δ^{9}-THCA concentrations reaching 431 ng/mL by GC-MS analysis. All subjects felt the pyschotropic effects of THC (103).

Opioids

The opioid drugs can be divided into naturally occurring, semi-synthetic, and synthetic groups. The naturally occurring opioids are morphine and codeine, which are derived from the opium poppy, *Papaver somniferum L*. Semi-synthetic opioids include heroin, hydromorphone, oxycodone, and oxymorphone. Synthetic opioids are meperidine, methadone, diphenoxylate, propoxyphene, and the fentanyls. Opioids have similar pharmacological effects and potential for addiction and tolerance; they vary only in their potency for analgesia, duration of action, and extent of abuse (106).

Opioids produce analgesic respiratory depressant, euphoric, and emetic effects. The triad of miosis, coma, and respiratory depression are pathognomonic of opioid poisoning (107). The response to the pure opioid antagonist, naloxone, is both diagnostic and therapeutic for opioid intoxication.

Heroin and methadone are the most frequently abused opioids. In recent years, concomitant use of heroin or methadone with cocaine has been reported. The preferred route of heroin administration is intravenous, although heroin of sufficient purity now is available and is smoked or administered intranasally. Heroin (diacetylmorphine) is deacetylated rapidly to 6-mon-oacetylmorphine (6-MAM) with a half-life of 3 minutes, which then undergoes a second deacetylation step to form morphine (108). Consequently, only 6-MAM and morphine are found in the urine of heroin users (109). Morphine and 6-MAM are equally potent in their opioid effects. Thus, heroin acts as a prodrug for providing the active metabolites, 6-MAM and morphine.

Codeine is rapidly absorbed from an oral dose, with plasma concentration peaking at 1 hour postingestion. It is extensively metabolized to norcodeine, but at least 10% of the dose is transformed to morphine. Most of the dose is excreted in the urine as glucuronide conjugates of codeine, norcodeine, and morphine. Thus, morphine is a metabolite of both heroin and codeine. Consumption of baked goods containing poppy seeds can result in detectable amounts of morphine and codeine in urine (110, 111) because some batches of poppy seeds are contaminated with these opiates. When a urine tests positive for morphine, it becomes important to ascertain whether the morphine comes from codeine in prescription medications, from heroin/morphine abuse, or from poppy-seed ingestion. A set of guidelines has been proposed to help in the interpretation of urine opiates results and the determination of the source of morphine and codeine (111, 112). Test results that can rule out poppy seeds as the sole source for morphine and codeine are: (i) codeine levels greater than 300 ng/mL with a morphine-to-codeine ratio less than 2 (indicative of codeine use); (ii) high levels of morphine (greater than 1 ng/mL) when codeine is undetectable; (iii) total morphine levels exceeding 5,000 ng/mL (indicative of abuse of heroin, morphine, or codeine); (iv) the presence of 6-MAM (a positive indication of heroin use).

To avoid “false-positive” morphine results as a result of poppy-seed ingestion, the Federal Workplace Drug Testing Program has raised the initial and confirmation test thresholds to 2,000 ng/mL of morphine and codeine. In addition, the Program now requires the test for the presence of 6-MAM in urine specimens confirmed to contain morphine (113).

The fentanyls are analgesic-anesthetic drugs that are many times more potent than morphine. Many analogs of fentanyls (3-methyl-, α-methyl-, and parafluoro derivatives) have appeared on the street as “China white” (114). Heroin containing fentanyl is a grave risk to heroin abusers. A standard urine drug screen most likely will fail to detect a fentanyl overdose because fentanyl concentration in urines is very low (µg/L), and the available urine immunoassays for opiates have no reactivity for fentanyls.

Immunoassays are used commonly as preliminary tests for the opiates (morphine and codeine), methadone, and propoxyphene. The synthetic narcotics (dihydrocodone, hydrocodone, hydromorphone, and oxycodone) also are detected by opiates immunoassays, thus necessitating a confirmation test (TLC, HPLC, GC, and GC-MS) that can distinguish between morphine and the other opioids. Conventional TLC has a detection limit of 1 mg/L of free morphine, which makes it suitable for use in overdose cases but not for drug abuse testing.

Cocaine

Cocaine (benzoylmethylecgonine) is an alkaloid extracted from the leaves of the *Erythroxylon coca* plant and purified as the hydrochloride salt (cocaine HCl). It is a powerful CNS stimulant, and in recent years the illicit use of cocaine has increased rapidly.

Cocaine hydrochloride is snorted or administered intravenously. Many cocaine users used to prepare free-base crystals for smoking (“free-basing”) by dissolving cocaine HCl in a solution of baking soda or ammonia, and extracting it with a solvent (ether) that is evaporated, leaving relatively pure cocaine crystals. “Crack” (so named because of the crackling sound made by the crystals when heated) also is a free-base form of cocaine, prepared by precipitation from an alkaline solution. Crack is relatively pure cocaine (80% to 90%), and when heated, is mostly vaporized rather than pyrolyzed. Its low cost and wide availability have worsened the cocaine epidemic.

Cocaine, a powerful CNS stimulant, produces heightened alertness, self-confidence, and an intense feeling of euphoria (rush). These stimulatory effects are followed by depression (crash). It is the positive reinforcement of the rush to escape the crash that leads to chronic cocaine abuse (115). Psychosis, repeated grand mal seizures, and coma are common following acute intoxication. Other clinical manifestations include arrhythmia, myocardial infarction, hypertensive crisis, cerebral vascular accidents, hyperthermia, and respiratory arrest. Spontaneous abortion and *abruptio placentae* are obstetric complications for women who use cocaine during pregnancy (3). Low birth weight and higher risk of congenital malformations, perinatal mortality, and neurobehavioral impairment are common results of a fetus exposed to cocaine in utero (4). Necrosis of the nasal septum from snorting cocaine, lung damage, and pulmonary edema are some of the other complications associated with cocaine use.

Bioavailability of intranasal cocaine is variable (20% to 60%) and appears to be dose dependent. Smoking free-base provides a more effective absorption of cocaine (57% bioavailability) (116), and a rapid peaking in its plasma concentration within 5 minutes. The euphoric effect of smoking cocaine usually lasts for only 20 minutes. Similar results were obtained in subjects given cocaine intravenously. Cocaine is lipid soluble and readily crosses cell membranes and distributes rapidly across the blood-brain barrier and the placenta.

The main routes of metabolism of cocaine are enzymatic and nonenzymatic hydrolysis of the methyl ester giving benzoylecgonine (BE) and enzymatic hydrolysis of the benzoyl group by plasma and liver esterases, yielding ecgonine methyl ester (EME) (117). Further hydrolysis of both of these compounds gives ecgonine. Norcocaine, the product of *N*-demethylation, undergoes similar hydrolysis and is believed to be pharmacologically active, but it is present in very small amounts. The two major metabolites, BE and EME, are excreted into urine in about equal amounts (40% to 50%). Cocaine itself is excreted in very small amounts, less than 1% of the dose in 3 days. Elimination half-lives for BE, EME, and cocaine have been calculated from literature data to be 7.5, 3.6, and 1.5 hours, respectively (118). A pyrolytic product of crack cocaine, anhydroecgonine methyl ester (methylecgonidine, MEG) has been identified and detected in urine, hair, saliva, sweat, and blood of crack smokers, but not in samples obtained form users of cocaine by other routes (119).

A pharmacologically active metabolite, cocaethylene, has been found in cocaine users following concurrent use of cocaine and ethanol (120). Cocaethylene is frequently present in clinical specimens which are positive for benzoylecgonine (121).

Addition of sodium fluoride to inhibit esterase activity and low temperature will stabilize cocaine in blood specimens for long-term storage. The major urine metabolites, BE and EME, are relatively stable.

Immunoassays are popular because they do not require sample pretreatment, and they are fast and easy to perform. Since these assays were designed primarily for use with urine specimens, the target analyte is the major urinary metabolite benzoylecognine (122). Benzoylecognine is not effectively extracted by nonpolar solvents and requires the addition of an alcohol (e.g., ethanol) or chloroform to make the extraction solvent more polar. The resulting extract can be used for various chromatography techniques (TLC, HPLC, GC, and GC-MS). TLC sensitivity is generally at 1 to 2 µg/mL, which is adequate for cocaine overdose cases. Sensitivity has been improved with claimed detection limits of 0.25 µg/mL for use in drug abuse testing (TOXI-LAB). A disadvantage of this TLC procedure is the limited migration of BE away from the origin.

Gas chromatography procedures require derivatization of the carboxylic side group of BE before chromatography (23). Procedures using a flame ionization detector have sensitivity down to 200 µg/mL, which can be enhanced further by the use of a mass-spectrometer detector.

The high specificity and sensitivity of GC/MS assays are utilized routinely for confirmation of BE in urine using deuterated BE as internal standard and selected ion ionization mode.

Ecgonine methyl ester (EME) is a major urinary metabolite of cocaine. Its presence in urine is an indicator of cocaine use is as sensitive and specific as that of BE. The use of EME has the advantage that, unlike BE, it can be easily extracted urine and it has good chromatographic properties in TLC and GC systems (123).

Urine BE levels typically decline rapidly to below the usual 300ng/mL cutoff in 24 to 96 hours since last exposure, but

among long-term, high-dose abusers BE was reportedly detectable after 10 to 22 days (124). Detection times, however, are assay-dependent. In a study of human subjects given 20 mg of cocaine HC1 intravenously, the mean times of detection of the last positive urine specimen (greater than or equal to 300 ng/mL BE) after cocaine administration varied from 16.9 to 52.9 hours, depending on the commercial tests used (TLC, EIA, FPLA, RIA) (122).

Sympathomimetic Drugs

Sympathomimetic drugs mimic the actions of the endogenous neurotransmitters that stimulate the sympathetic nervous system. Sympathomimetic agents include the over-the-counter medications (e.g., ephedrine, pseudoephedrine, and phenylpropanolamine), illegal street drugs (e.g., amphetamines, and methamphetamine, MDMA, and MDA), and herbal preparations (ephedra, Ma Huang). The clinical presentations from these many different agents are the classic sympathomimetic signs and symptoms such as tachycardia, hypertension, diaphoresis, hyperthermia, acute psychosis, and seizures. Poisoning occurs secondary to the use of prescription agents, over-the-counter medications, and illicit drugs. A contemporary scene is the abuse of the designer agents (particularly MDMA or “ecstasy”) at rave dance parties where users have been afflicted with dehydration, hyperthermia, and cardiac arrhythmia (125).

Methamphetamine and amphetamine are strong CNS stimulants, and at low doses produce euphoria, increase alertness, intensified emotions and sense of well being. Because of its ease of manufacturing and ready availability and its longer half-life (10 hours vs. minutes for cocaine), the abuse of methamphetamine has increased in recent years (126). Methamphetamine may be taken orally, intravenously, or by smoking. The smokeable form of methamphetamine has the street name “ice.” Methamphetamine is excreted in the urine largely as the unchanged drug and a small amount (7%) is demethylated to amphetamine. With a pKa of approximately 9.9 for both amphetamine and methamphetamine, elimination of these drugs is highly dependent on urine pH: up to 74% of a dose of amphetamine is excreted unchanged in acidic urine, but only 2% in alkaline urine (30).

The sympathomimetics can be detected by immunoassays. There are many commercially available kits either as reagent packs to be used on automated instruments or single-use devices designed for near-patient testing (15, 16). In choosing which kit to use, one must be familiar with the immunospecificity of the assay and the cutoff used. Most of the products on the market are designed to meet the demand of workplace drug testing for high specificity for methamphetamine and amphetamine, the target drugs of abuse in most testing programs. Assays with low cross reactivity with over-the-counter sympathomimetic amines such as ephedrine or phenylpropanolamine will reduce the need for costly GC-MS confirmation. In the clinical setting, however, it is important to be able to detect other sympathomimetics as well. Emergency department or drug-treatment clinic encounters with designer drugs (MDMA, MDA), prescription (phentermine) or nonprescription (ephedrine, PPA) drugs, and herbal preparations (ephedra) are common.

Specific identification of individual sympathomimetics requires chromatographic assays, typically those based on TLC or GC-MS. In a TLC kit commonly used in many clinical toxicology laboratories, the migration and color characteristics of the sympathomimetics are similar although subtle differences do allow correct identification by expert chromatographers (18). The manufacturer of the kit provides a special (remigration) procedure for detection and differentiation of the common sympathomimetics. The detection limits for the sympathomimetics are in the range of 0.5 to 1.0 (g/mL, which are adequate for testing patients who are suspected of an overdose.

Gas chromatography of the sympathomimetics requires chemical derivatization. The most commonly used with primary and secondary amines include heptafluorobutyric anhydride (HFBA), pentafluoropropionic anhydride (PFPA), trifluoroacetic anhydride (TFAA) and 4-carbethoxyhexfluorobutyryl chloride (4-CB). The methamphetamine artifact has been reported in some specimens when they were used with some of these derivatives (127). The source appeared to be ephedrine, which was in very high concentrations in these specimens. Under the conditions of the assay, a small amount of the ephedrine was converted to methamphetamine. The laboratory should be careful in reporting a positive methamphetamine result in the presence of a high concentration of ephedrine (> 0.1 mg/mL) and when amphetamine, a methamphetamine metabolite, is not detected at ≥200 ng/mL. Various solutions to this problem have been proposed, including lowering the injection port temperature or using a different derivative (128). The most effective solution is the fragmentation of ephedrine and related compounds with periodate treatment (129).

The *d* and *l* enantiomers of methamphetamine have very different CNS stimulant potency. *d*-Methamphetamine is 10 times more powerful as a stimulant, is a controlled substance in the United States, and is a drug of abuse of choice. *l*-methamphetamine has greater peripheral vasoconstrictive action and is used in over-the-counter nasal inhaler (Vicks Inhaler, Procter & Gamble, Cincinnati, OH). All pharmaceutical methamphetamine is the *d* isomer whereas illicitly synthesized methamphetamine can be pure *d*-methamphetamine or racemic methamphetamine. The standard clinical toxicology tests for the amphetamines (TLC, HPLC, or GC-MS) cannot differentiate between the enantiomers. Special chiral derivatizing reagent may be used to convert the isomers into chiral derivatives, which can be separated by standard chromatography (130).

A number of drugs or compounds are metabolized to methamphetamine and/or amphetamine. These include amphetaminil, benzphetamine, clobenzorex, deprenyl, dimethylamphetamine, ethylamphetamine, famprofazone, fencamine, fenethylline, fenproporex, furfenorex, mefenorex, mesocarb, and prenylamine (128). Proper interpretation of positive methamphetamine or amphetamine results must take into account the medication history of the patient.

Date-Rape and Knockout Drugs

The use of drugs to facilitate sexual assault is not new. Ethanol is the most frequently involved drug, being particularly effective when used together with other CNS depressants. Newer drugs

that have been implicated include flunitrazepam and γ-hydroxybutyrate (GHB), although any benzodiazepine is effective, particularly when it is ingested with ethanol (131). These drugs cause the victim to lose consciousness and the ability to make rational decisions. Moreover, perpetrators take advantage of the anterograde amnesia that their victims frequently suffer.

Flunitrazepam (Rohypnol), with street names such as “rooffies,” “rochies,” “rocha,” “rophies,” is not available legally in the United States. It dissolves readily in alcohol and is colorless, odorless, and tasteless. Detection of flunitrazepam in blood or urine is difficult because of the small amount administered (typically 1 to 2 mg) and its extensive metabolism. Documentation of exposure usually is based on the detection of the major metabolite, 7-aminoflunitrazepam. Many of the commercially available immunoassays for benzodiazepines have sufficient cross-reactivity with this metabolite to detect its presence in urine. Definitive confirmation is by GC-MS (132). It is recommended that urine and blood specimens should be collected as soon as possible after the alleged assault and preferably within 24 hours (133).

Gamma-hydroxybutyrate is a naturally occurring product of the metabolism of the neurotransmitter γ-aminobutyric acid (GABA). In the 1980s, it was abused by body builders as a steroid enhancer and was easily available in health-food stores and mail-order outlets until its ban by the Food and Drug Administration in 1990. It continues its popularity as an illicit drug among “rave” party participants for its euphoric and hallucinogenic properties. Gamma-hydroxybutyrate is a CNS depressant and an excessive amount ingested in combination with ethanol can cause profound coma (134, 135). Hence, its implication as a date-rape drug. Gamma butyrolactone (GBL), with a street name of “blue nitro,” and 1,4-butandiol (BD), are converted, *in vivo*, to GHB. In February 2000, GHB, GBL, and BD were added to schedule I of the Control Substance Act.

There is no readily available screening test for GHB. Some laboratories may be able to use their urine organic-acid screen methodologies or specific GC-MS procedures to detect GHB (136, 137).

FORENSIC URINE DRUG TESTING

Part of "22 - Toxicology"

One of the initiatives for achieving a drug-free workplace is to implement urine drug testing programs for job applicants as well as for workers in occupations that are considered critical to public safety and health, e.g., those in nuclear-power plants and the transportation industry. Urinalysis programs remain a controversial approach to dealing with the substance-abuse problem in the workplace. Among the issues being intensely debated are the accuracy of the tests and the reliability of the laboratories conducting them (138, 139).

It is important to appreciate that urine drug testing for the workplace differs from clinical testing in several critical respects (138). For example, workplace drug testing results may be used for administrative purposes such as disciplinary action, dismissal, or denial of employment. Moreover, such test results stand on their own and are interpreted without the usual benefit of the physician-patient relationship or the clinical context in clinical testing. Thus, workplace drug testing must produce definitive test results that are both scientifically and legally defensible, the latter requiring comprehensive documentation to show that the integrity of the specimen has been maintained (chain of custody). This forensic aspect of substance abuse testing usually is unfamiliar to clinical laboratorians.

In addition to standard good laboratory practice, stringent requirements are needed for forensic urine drug testing (FUDT) laboratories. A discussion of these requirements follows.

Laboratory Facility

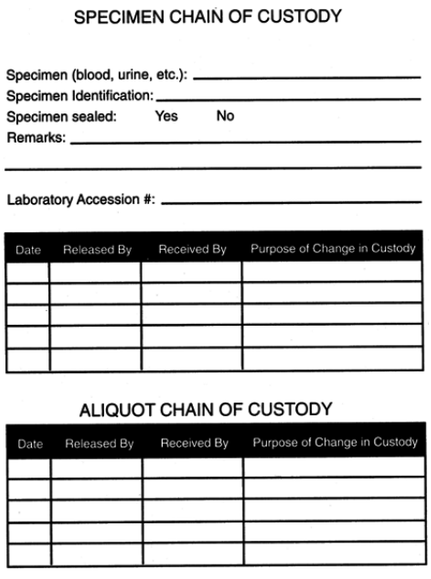
The FUDT laboratory where specimens and records are kept and where analytical work is performed should be separate from areas of clinical activities, and the laboratory should be a limited access area to authorized personnel. The laboratory should have protocols to protect the confidentiality of all records, to ensure the security of computer files relating to FUDT, and to limit computer access to authorized users.

Initial and Confirmation Tests

FUDT is a two-stage process: initial testing followed by confirmatory analysis of the presumptive positive specimens by quantitative GC-MS. Most FUDT programs required the performance of both initial and confirmation tests at the same laboratory site.

Chain-of-Custody Documentation

In FUDT, every urine specimen is treated as evidence. Therefore, there is a need to show proof of integrity of the specimen from specimen collection to its final disposition after analysis is completed. This requires documentation of *what* was done with the specimen (or aliquot), *who* handled the specimen (or aliquot), and *when* it was handled. A sample chain-of-custody form is shown in Figure 22.6.



SPECIMEN CHAIN OF CUSTODY

Specimen (blood, urine, etc.): _____
Specimen Identification: _____
Specimen sealed: Yes No
Remarks: _____
_____

Laboratory Accession #: _____

Date	Released By	Received By	Purpose of Change in Custody

ALIQUOT CHAIN OF CUSTODY

Date	Released By	Received By	Purpose of Change in Custody

FIGURE 22.6. Specimen/aliquot chain-of-custody form.

Because the chain-of-custody procedures are the most challenged aspect of FUDT, they should be an important part of a laboratory's quality assurance program. The chain-of-custody of every specimen must be checked to assure that it has been completed appropriately before certifying results for reporting.

Personnel

The qualifications of a FUDT laboratory director are stringent, requiring training, experience, and expertise in forensic analytical toxicology comparable to that of a diplomat of the American Board of Clinical Chemistry in Toxicological Chemistry or the American Board of Forensic Toxicology.

The forensic aspect of drug testing creates the need for a certifying scientist (CS) that does not exist in clinical laboratories. The CS performs the final review of the entire laboratory process: the chain-of-custody documentation, and standards, blanks and quality-control data of preliminary and confirmation tests before certifying the reports for release.

Accreditation and Licensure

In some states, e.g., New York, FUDT laboratories must meet state regulatory licensure requirements. At the national level there are two certification programs: The College of American Pathologists (CAP) Forensic Urine Drug Testing program and the Department of Health and Human Services (DHHS) National Laboratory Certification Program (NLCP). The federal program is based on a set of guidelines issued by Substance Abuse Mental Health Services Administration (SAMHSA) (102). Certification by this program is required for laboratories engaged in drug testing for the federal government or industries regulated by the Department of Transportation (140). Both programs require acceptable performance in proficiency-testing programs, and laboratories must successfully pass on-site inspection for obtaining and maintaining accreditation or certification.

REFERENCES

1. Kwong TC, Shearer DS. Detection of drug use during pregnancy. *Ob Gyn Clinics N.A.* 1998;25:43–61.

2. Shearer DS, Baciewicz GJ, Kwong TC. Drugs of abuse testing in a psychiatric outpatient service. *Clinics Lab Med* 1998;18:713–726.

3. Chasnoff IJ. Cocaine use in pregnancy: perinatal morbidity and mortality. *Neurotoxicol Teratol* 1987;9:291–293.

4. Gingras JL, Weese-Mayer DE, Hume RFJ, et al. Cocaine and development: mechanisms of fetal toxicity and neonatal consequences of prenatal cocaine exposure. *Early Hum Dev* 1992;31:1–24.

5. Ryan RM, Wagner CL, Schulz J, et al. Meconium analysis for improved identification of infants exposed to cocaine in utero. *J Pediatr* 1994;125:435–440.

6. Rollins DE, Wilkins DG, Gygi SP, et al. Testing for drugs of abuse in hair-experimental observations and indications for future research. *Forensic Sci Rev* 1997;9:23–35.

7. Kwong TC, Ryan RM. Detection of intrauterine illicit drug exposure by newborn drug testing. *Clin Chem* 1997;43:235–242.

8. Osterloh JD. Utility and reliability of emergency toxicologic testing. *Emerg Med Clin North Am* 1990;8:693–723.

9. Brett AS. Implications of discordance between clinical impression and toxicology analysis in drug overdose. *Arch Intern Med* 1988;148:437–441.

10. Kellerman AL, Fihn SD, LoGerto JP, et al. Impact of drug screen in suspected overdose. *Ann Emerg Med* 1987;16:1206–1216.

11. Stevens HM. Colour tests. In: Moffat AC, ed. *Clark's Isolation and Identification of Drugs*. London: The Pharmaceutical Press, 1986:128–147.

12. Simonick TF, Watts VW. Preliminary evaluation of the Abbott TDx for screening of D-methamphetamine in whole blood specimens. *J Anal Toxicol* 1992;16:115–118.

13. Schwartz JG, Hurd ILCJJ. Determination of tricyclic antidepressants for ED analysis. *Am J Emerg Med* 1994;12:513–516.

14. de la Torre R, Domingo-Salvany A, Badia R, et al. Clinical evaluation of the Triage® analytical device for drugs-of-abuse testing. *Clin Chem* 1966;42:1433–1438.

15. Buchan BJ, Walsh JM, Leaverton PE. Evaluation of the accuracy of on-site multi-analyte drug testing devices in the determination of the prevalence of illicit drugs in drivers. *J Forensic Sci* 1998;43:395–399.

16. Liu RH. Evaluation of common immunoassay kits for effective workplace drug testing. In: Liu RH, Goldberger BA, eds. *Handbook of Workplace Drug Testing*. Washington, D.C.: AACC Press, 1995:67–129.

17. Moffat AC. Thin Layer Chromatography. In: Moffat AC, ed. *Clark's Isolation and Identification of Drugs*. London: The Pharmaceutical Press, 1986:160–177.

18. Jarvie DR, Simpson D. Drug screening: evaluation of the Toxi-Lab TLC system. *Ann Clin Biochem* 1986;23:76–84.

19. Poklis A. Gas chromatography. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry* 2nd ed. St. Louis: CV Mosby, 1989:110–125.

20. Watts VW, Simonick TF. Screening of basic drugs in biological samples using dual column capillary chromatography and nitrogen-phosphorous detectors. *J Anal Toxicol* 1986;10:198–204.

21. Drummer OH, Horomidis S, Kourtis S, et al. Capillary gas chromatographic drug screen for use in forensic toxicology. *J Anal Toxicol* 1994;18:134–238.

22. Smith NB. Automated identification by computer of the mass spectra of drugs in urine or serum extracts. *J Anal Toxicol* 1994;18:16–21.

23. Goldberger BA, Cone E. Confirmatory tests for drugs in the workplace by gas chromatography-mass spectrometry. *J Chromatogr* 1994;674:73–86.

24. Bowers LD. Liquid chromatography. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry* 2nd ed. St. Louis: CV Mosby, 1989:94–109.

25. Binder SR, Regalia M, Biaggi-McEachern M, et al. Automated liquid chromatographic analysis of drugs in urine by on-column sample cleanup and isocratic multi-column separation. *J Chromatogr* 1989;473:325–341.

26. Ohtsuji M, Lai JS, Binder SR, et al. Use of REMEDi HS in emergency toxicology for a rapid estimation of drug concentrations in urine, serum, and gastric samples. *J Forensic Sci* 1996;41:881–886.

27. Fitzgerald RL, Rivera JD, Herold DA. Broad spectrum drug identification directly from urine, using liquid chromatography-tandem mass spectrometry. *Clin Chem* 1999;45:1224–1234.

28. Maurer HH. Liquid chromatography-mass spectrometry in forensic and clinical toxicology. *J Chromatogr* 1998;713:3–25.

29. Dubowski KM. Absorption, distribution and elimination of alcohol: highway safety aspects. *J Stud Alcohol* 1985;98–108.

30. Baselt RC, Cravey RH. d-Methamphetamine, l-methamphetamine. *Disposition of toxic drugs and chemicals in man*. 4th ed. Foster City, CA: Chemical Toxicology Institute, 1995:475–479.

31. Hlastala MP. The alcohol breath test: a review. *J Appl Physiol* 1998;84:401–408.

32. Whitfield JB, Starmer GA, Martin NG. Alcohol metabolism in men and women. *Alcohol Clin Exp Res* 1990;14:785–786.

33. Bogusz M, Pach J, Stasko W. Comparative studies on the rate of ethanol elimination in acute poisoning and in controlled conditions. *J Forensic Sci* 1977;22:446–451.

34. Blood alcohol testing in the clinical laboratory; approval guidelines. NCCLS document T/DM6-A. 1997. Wayne, PA, NCCLS.

35. Dubowski KM. *The technology of breath-alcohol analysis*. DHHS Publication No. (ADM) 92-1728. Washington,D.C.: U.S. Department of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism, 1992.

36. Bendsten P, Hultberg J, Carlsson M, et al. Monitoring ethanol exposure in a clinical setting by analysis of blood, breath, saliva, and urine. *Alcohol Clin Exp Res* 1999;23:1446–1451.

37. Dubowski KM. Quality assurance in breath-alcohol analysis. *J Anal Toxicol* 1994;18:306–311.

38. Gadsden RH, Terry CS. Alcohols in biological fluids by gas chromatography (automated head-space method). In: Frings CS, Faulkner WR, eds. *Selected Methods in Emergency Toxicology*. Washington, D.C.: AACC Press, 1986:40–43.

39. Mahieu P, Hassoun A, Lauwerys R. Predictions of methanol intoxication with unfavorable outcome. *Hum Toxicol* 1989;8:135–137.

40. Goldfrank LR, Flomenbaum NE. Toxic alcohols. In: Goldfrank LR, Flomenbaum NE, Lewin NA, et al., eds. *Goldfrank's Toxicologic Emergencies*. 6th ed. Stamford, CT: Appleton & Lange, 1998:1049–1060.

41. Howland MA. 4-Methylpyrazole (Fomepizole). In: Goldfrank LR, Flomenbaum NE, Lewin NA, et al., eds. *Goldfrank's Toxicologic Emergencies* 6th ed. Stamford, CT: Appleton & Lange, 1998:1067–1069.

42. Manlandain H, Cano Y. Serum methanol in the absence of methanol ingestion. *Ann Emerg Med* 1996;28:102–103.

43. Geller RJ, Skyer DA, Herold DA, et al. Serum osmolal gap and ethanol concentration: a simple and accurate formula. *Clin Toxicol* 1986;24:77–84.

44. Hewlett TP, McMartin KE, Lauro AJ, et al. Ethylene glycol poisoning. The value of glycolic acid determinations for diagnosis and treatment. *Clin Toxicol* 1986;24:402.

45. Jones AW, Nisson L, Gladh SA, et al. 2,3-Butandiol in plasma from an alcoholic mistakenly identified as ethylene glycol by gas chromatographic analysis. *Clin Chem* 1991;37:1453–1455.

46. Yao HH, Porter WH. Simultaneous determination of ethylene glycol and its major metabolite, glycolic acid, in serum by gas chromatography. *Clin Chem* 1996;42:292–297.

47. Standefer J, Blackwell W. Enzymatic method for measuring ethylene glycol with a centrifugal analyzer. *Clin Chem* 1991;37:1734–1736.

48. Nisson L, Jones AW. 2,3-Butandiol: a potential interfering substance in the assay of ethylene glycol by an enzymatic method. *Clin Chim Acta* 1992;208:225–229.

49. Eder AF, Dowdy YG, Gardiner JAM, Wolf BA, Shaw LM. Serum lactate and lactate dehydrogenase in high concentrations interfere in enzymatic assay of ethylene glycol. *Clin Chem*. 1996;42:1489–1491.

50. Linden CH, Rumack BH. Acetaminophen overdose. *Emerg Med Clin North Am* 1984;2:103–119.

51. Lauterburg BH, Velez ME. Glutathione deficiency in alcoholics: risk factor for paracetamol hepatoxicity. *Gut* 1988;29:1153–1157.

52. Cocoran GB, Mitchell JR, Vaishnav YN, et al. Evidence that acetaminophen and N-hydroxyacetaminophen form a common arylating intermediate, N-aceyl-p-benzoquinoneimine. *Mol Pharmcol* 1980;18:536–542.

53. Smilkstein MJ, Knapp GL, Kulig KW, et al. Efficacy of oral N-acetylcysteine in the treatment of acetaminophen overdose. *N Engl J Med* 1988;319:1557–1562.

54. Prescott LF, Illingsworth RN, Critchley J. Intravenous N-acetylcysteine: the treatment of choice for paracetamol poisoning. *Br Med J* 1979;2:1097–1100.

55. Rumack BH, Matthew H. Acetaminophen poisoning and toxicity. *Pediatrics* 1975;55:871–876.

56. Stewart MJ, Watson ID. Analytical review in clinical chemistry: methods for the estimation of salicylate and paracetamol in serum, plasma, and urine. *Ann Clin Biochem* 1987;24:552–565.

57. Kwong TC. Salicylate measurement: clinical usefulness and methodology. *CRC Crit Rev Clin Lab Med* 1987;25:137–159.

58. Litovitz TL, Klein-Schwartz W, Caravati EM, et al. 1998 Annual Report of the American Association of Poison Control Centers Toxic Exposure Surveillance System. *Am J Emerg Med* 1999;17:435–487.

59. Flomenbaum NE. Salicylates. In: Goldfrank LR, Flomenbaum NE, Lewin NA, et al., eds. *Goldfrank's Toxicologic Emergencies*. 6th ed. Stamford, CT: Appleton & Lange, 1998:569–591.

60. Proudfoot AT. Toxicity of salicylates. *A J Med* 1983;75:99–103.

61. Anderson RJ, Potts DE, Gabow PA, et al. Unrecognized adult salicylate intoxication. *Ann Int Med* 1976;85:745–748.

62. Done AK. Salicylate intoxication: significance of measurements of salicylate in blood in cases of acute ingestion. *Pediatr* 1960;26:800–806.

63. Dugandzic RM, Tierney MG, Dickinson GE, et al. Evaluation of the validity of the Done nomogram in the management of acute salicylate intoxication. *Ann Emerg Med* 1989;18:1186–1190.

64. Kwong TC, Laczin J, Baum J. Self poisoning with enteric coated aspirin. *Am J Clin Pathol* 1983;80:888–890.

65. Andersen BD, Alexander MS, Ng KJ, Bianchine JR. Aspirin and Reye's disease: a reinterpretation. *Lancet* 1982;1:903s.

66. Trinder P. Rapid determination of salicylate in biological fluids. *Biochem J*. 1957;54:301–303.

67. You K, Bittikofer JA. Quantitation of salicylate in serum by use of salicylate hydrolase. *Clin Chem* 1984;30:1549–1551.

68. Chubb SAP, Campbell RS, Hammond PM, et al. An enzyme mediated, colorimetric method for measurement of salicylate. *Clin Chim Acta* 1986;155:209–220.

69. Morris HC, Overton PD, Ramsey J, et al. Development and validation of an automated, enzyme-mediated colorimetric assay of salicylate in serum. *Clin Chem* 1990;36:131–135.

70. Hobbs WR, Rall TW, Verdoorn TA. Hypnotics and sedatives; ethanol. In: Hardman JG, Limbird LE, Molinoff PB, Roddon RW, eds. *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. 9th ed. New York: McGraw Hill, 1996:361–396.

71. Soo VA, Bergert RJ, Deutsch DG. Screening and quantification of hypnotic sedatives in serum by capillary gas chromatography with a nitrogen-phosphorus detector, and confirmation by capillary gas chromatography-mass spectrometry. *Clin Chem* 1986;32:325–328.

72. Atwell SH, Green VA, Haney WG. Development and evaluation of a method for simultaneous determination of phenobarbital and diphenylhydantoin in plasma by high-pressure liquid chromatography. *J Pharm Sci* 1975;64:806–809.

73. Meatherall R. GC/MS confirmation of barbiturates in blood and urine. *J Forensic Sci* 1997;42:1160–1170.

74. Woods JH, Katz JL, Winger G. Use and abuse of benzodiazepines. *JAMA* 1988;260:3476–3480.

75. Drummer OH, Ranson DL. Sudden death and benzodiazepines. *Am J Forensic Med Pathol* 1996;17:336–342.

76. Winlker E, Almog S, Kriger D, et al. Use of flumazenil in the diagnosis and treatment of patients with coma of unknown etiology. *Crit Care Med* 1993;21:538–542.

77. Mazhar M, Binder SR. Analysis of benzodiazepines and tricyclic antidepressants in serum using a common solid-phase clean-up and a common mobile phase. *J Chromatogr* 1989;497:201–212.

78. Nishikawa T, Ohtani H, Herold DA, et al. Comparison of assay methods for benzodiazepines in urine. *Am J Clin Pathol* 1997;107:345–352.

79. Roets E, Hoogmartens J. Thin layer chromatography of the acid hydrolysis products of nineteen benzodiazepine derivatives. *J Chromatogr* 1980;194:262–269.

80. Fraser AD. Urinary screening for alprazolam, triazolam, and their metabolites with EMIT d.a.u. benzodiazepine metabolite assay. *J Anal Toxicol* 1987;11:263–266.

81. Fraser AD, Bryan W, Isner AF. Urinary screening for alprazolam and its major metabolites by the Abbott ADx and TDx analyzers with confirmation by GC-MS. *J Anal Toxicol* 1991;15:25–29.

82. Weisman RS. Cyclic antidepressants. In: Goldfrank LR, Flomenbaum NE, Lewin NA, Weisman RS, Howland MA, Hoffman RS, eds. *Goldfrank's Toxicologic Emergencies*. 6th ed. Stamford, CT: Appleton & Lange, 1998:925–934.

83. Boehnert MT, Lovejoy FH. Value of the QRS duration versus the serum drug level in predicting seizures and ventricular arrhythmias after an acute overdose of tricyclic antidepressants. *N Engl J Med* 1985;313:474–479.

84. Foulke GE. Identifying toxicity risk early after antidepressant overdose. *Am J Emerg Med* 1995;13:123–126.

85. Spiker DG, Weiss AN, Chang SS, et al. Tricyclic antidepressant overdose: clinical presentation and plasma levels. *Clin Pharmacol Ther* 1975;18:539–546.

86. Orsulak PJ. Therapeutic monitoring of antidepressant drugs: guidelines updated. *Ther Drug Monit* 1989;11:497–507.

87. Kabra PM, Nar NA, Marton LJ. Simultaneous liquid chromatographic analyses of amitriptyline, nortriptyline, imipramine, desipramine, doxepin, and nordoxepin. *Clin Chim Acta* 1981;111:123–132.

88. Demorest DM. Distinguishing cyclobenzaprine from amitriptyline and imipramine by liquid chromatography with UV multi-wavelength or full-spectrum detection. Blood levels of cyclobenzaprine in emergency screens. *J Anal Toxicol* 1987;11:133s

89. Bateh RP. Distinguishing cyclobenzaprine and amitriptyline. *J Anal Toxicol* 1987;11:235

90. Perrone J. Iron. In: Goldfrank LR, Flomenbaum NE, Lewin NA, Weisman RS, Howland MA, Hoffman RS, eds. *Goldfrank's Toxicologic Emergencies*. 6th ed. Stamford, CT: Appleton & Lange; 1998:619–627.

91. Everson GW, Oudjhane K, Young LW, Krenzilock EP. Effectiveness of abdominal radiographs in visualizing chewable iron supplements following overdose. *Am J Emerg Med* 1989;7:459–463.

92. Proudfoot AT, Simpson D, Dyson EH. Management of acute iron poisoning. *Med Toxicol* 1986;1:83–100.

93. Tenenbein M, Yatscoff RW. The total iron binding capacity in iron poisoning. Is it useful? *Am J Dis Child* 1991;145:437–439.

94. Helfer RE, Rodgerson DO. The effects of deferoxamine on determination of serum iron binding capacity. *J Pediatr* 1966;68:804–806.

95. Huestis M. Pharmacology and toxicology of marijuana. *Ther Drug Monit* 1993;14:131–138.

96. Ellis GM, Maun MA, Judson BA, et al. Excretion patterns of cannabinoid metabolites after last use in a group of chronic users. *Clin Pharmacol Ther* 1986;38:572–578.

97. Smith-Kielland A, Skuterud B, Morland J. Urinary excretion of 11-nor-9-carboxy-delta9-tetrahydrocannabinol and cannabinoids in frequent and infrequent drug users. *J Anal Toxicol* 1999;23:323–332.

98. Fraser AD, Worth D. Urinary excretion profiles of 11-nor-9-carboxy-delta9-tetrahydrocannabinol: a delta9-THCCOOH to creatinine ratio study. *J Anal Toxicol* 1999;23:531–534.

99. Huestis M, Cone E. Differentiating new marijuana use from residual drug excretion in occasional marijuana users. *J Anal Toxicol* 1998;22:445–454.

100. Cone E, Johnson RE. Contact highs and urinary cannabinoid excretion after passive exposure to marijuana smoke. *Clin Pharmacol Ther* 1986;40:247–256.

101. Mule SJ, Lomax P, Gross SJ. Active and realistic passive marijuana exposure tested by three immunoassays and GC/MS in urine. *J Anal Toxicol* 1988;12:113–116.

102. Mandatory guidelines for federal workplace drug testing programs. *Federal Register* 1994;59:29908–29981.

103. Lehmann T, Sager F, Brenneisen R. Excretion of cannabinoids in urine after ingestion of cannabis seed oil. *J Anal Toxicol* 1997;21:373–375.

104. Fortner N, Fogerson R, Lindman D, et al. Marijuana-positive urine test results from consumption of hemp seeds in food products. *J Anal Toxicol* 1997;21:476–481.

105. Struempler RE, Nelson G, Urry FM. A positive cannabinoids workplace drug test following the ingestion of commercially available hemp seed oil. *J Anal Toxicol* 1997;21:283–285.

106. Reisine T, Pasternak G. Opioid analgesics and antagonists. In: Hardman JG, Limbird LE, Molinoff PB, Roddon RW, eds. *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. 9th ed. New York: McGraw-Hill, 1996:521–555.

107. Nelson LS. Opioids. In: Goldfrank LR, Flomenbaum NE, Lewin NA, Weisman RS, Howland MA, Hoffman RS, eds. *Goldfrank's Toxicologic Emergencies*. 6th ed. Stamford, CT: Appleton & Lange, 1998:973–995.

108. Inturrisi CE, Max MB, Foley KM, et al. The pharmacokinetics of heroin in patients with chronic pain. *N Engl J Med* 1984;310:1213–1217.

109. Cone E, Welch P, Mitchell JM, et al. Forensic drug testing for opiates: 1. Detection of 6-acetylmorphine in urine as an indicator of recent heroin exposure. *J Anal Toxicol* 1991;15:1–7.

110. Mule SJ, Casella GA. Rendering the "poppy-seed defense" defenseless: identification of 6-monoacetylmorphine in urine by gas chromatography/mass spectroscopy. *Clin Chem* 1988;34:1427–1430.

111. ElSohly MA, Stanford DF. Poppy seed ingestion and opiates urinalysis: a closer look. *J Anal Toxicol* 1990;14:308–310.

112. ElSohly MA, Jones AB. Morphine and codeine in biological fluids: approaches to source differentiation. *Forensic Sci Rev* 1989;1:13–22.

113. Changes to the testing cutoff levels for opiates for Federal workplace drug testing programs. *Federal Register* 1997;62:51118–51120.

114. Henderson GL. Designer drugs: past history and future prospects. *J Forensic Sci* 1988;33:569–575.

115. Gawin GH, Ellingwood EH Jr. Cocaine and other stimulants. *N Engl J Med* 1988;318:1173–1182.

116. Jeffcoat AR, Perez-Reyes M, Hill JM, et al. Cocaine disposition in humans after intravenous injection, nasal insufflation (snorting), or smoking. *Drug Metab Dipos* 1989;17:153–159.

117. Warner A, Norman AB. Mechanisms of cocaine hydrolysis and metabolism in vitro and in vivo: a clarification. *Ther Drug Monit* 2000; 22:266–270.

118. Ambre J, Ruo TI, Nelson J, et al. Urinary excretion of cocaine, benzoylecgonine, and ecgoninemethylester in humans. *J Anal Toxicol* 1988;12:301–306.

119. Kintz P, Sengler C, Cirimele V, et al. Evidence of crack use by anhydroecgonine methylester identification. *Human Exp Toxicol* 1997;16:123–127.

120. Landry MJ. An overview of cocaethylene, an alcohol-derived, psychoactive, cocaine metabolite. *J Psychoactive Drugs* 1992;24:273–276.

121. Bailey DN. Cocaethylene (ethylcocaine) detection during toxicological screening of a university medical center patient population. *J Anal Toxicol* 1995;19:247–250.

122. Cone E, Menchen SL, Paul BD, et al. Validity testing of commercial urine cocaine metabolite assays. 1. Assay detection times, individual excretion patterns, and kinetics after cocaine administration to humans. *J Forensic Sci* 1989;34:15–31.

123. Clark DR, Hajar TM. Detection and confirmation of cocaine use by chromatographic analysis for methylecgonine in urine. *Clin Chem* 1987;33:118–119.

124. Weiss RD, Gawin FH. Protracted elimination of cocaine metabolites in long-term, high-dose cocaine abusers. *A J Med* 1988;85:879–880.

125. Schwartz RH, Miller NS. MDMA (Ecstasy) and the rave: a review. *Pediatrics* 1997;100:705–708.

126. Mackenzie RG, Heischobar B. Methamphetamine. *Pediatr Rev* 1997;18:305–309.

127. Wu AHB, Wong SS, Johnson KG, et al. The conversion of ephedrine to methamphetamine and methamphetamine-like compounds during and prior to gas chromatography/mass spectrometric analysis of CB and HFB derivatives. *Bio Mass Spectrom* 1992;21:278–284.

128. Cody JT. Important issues in testing for amphetamines. In: Liu RH, Goldberger BA, eds. *Handbook of Workplace Drug Testing*. Washington, D.C.: AACC Press, 1995:239–288.

129. ElSohly MA, Stanford DF, Sherman D, et al. A procedure for eliminating interferences from ephedrine and related compounds in the GC/MS analysis of amphetamine and methamphetamine. *J Anal Toxicol* 1992;16:109.

130. Cody JT. Determination of methamphetamine enantiomer ratios in urine by gas chromatography-mass spectrometry. *J Chromatogr* 1992;580:77.

131. ElSohly MA, Salamone SJ. Prevalence of drugs used in cases of alleged sexual assault. *J Anal Toxicol* 1999;23:141–146.

132. ElSohly MA, Feng S, Salamone SJ, et al. GC-MS determination of flunitrazepam and its major metabolite in whole blood and plasma. *J Anal Toxicol* 1999;23:486–489.

133. LeBeau M, Andollo M, Hearn L, et al. Recommendations for toxicological investigations of drug-facilitated sexual assaults. *J Forensic Sci* 1999;44:227–230.

134. Li J, Stokes SA, Woeckner A. A tale of novel intoxication: a review of the effects of gamma-hydroxybutyric acid with recommendations for management. *Annals Emerg Med* 1988;31:729–736.

135. Li J, Stokes SA, Woeckner A. A tale of novel intoxication: seven cases of gamma-hydroxbutyric acid overdose. *Annals Emerg Med* 1998;31:723–728.

136. Couper FJ, Logan BK. Determination of γ-hydroxybutyrate (GHB) in biological specimens by gas chromatography-mass spectrometry. *J Anal Toxicol* 2000;24:1–7.

137. McCusker RR, Paget-Wilkes H, Chronister CW, et al. Analysis of gamma-hydroxybutyrate (GHB) in urine by gas chromatography-mass spectrometry. *J Anal Toxicol* 1999;23:301–305.

138. Kwong TC, Chamberlain RT, Frederick DL, et al. Critical issues in urinalysis of abused substances: report of the Substances-abuse Testing Committee. *Clin Chem* 1988;34:605–632.

139. Schottenfeld RS. Drug and alcohol testing in the workplace: objectives, pitfalls, and guidelines. *Am J Drug Alcohol Abuse* 1989;15:413–427.

140. Part 40-Procedures for transportation workplace drug and alcohol testing programs. *Federal Register* 1994;59:7354–7366.

23

Trace Elements, Vitamins, and Nutrition

Jane M. Yang

Kent B. Lewandrowski

NUTRITION

Part of "23 - Trace Elements, Vitamins, and Nutrition"

The basic food substances required by humans are carbohydrates, fats, and proteins. These macronutrients must be ingested in appropriate amounts to meet the energy needs of the body, which vary with body size, activity, age, and gender. The presence of illness or trauma may augment daily requirements. Proteins and carbohydrates provide about 4 kcal/g. Fats furnish about 9 kcal/g.

In addition to water and energy necessities, there also are requirements for nine essential amino acids (lysine, threonine, leucine, isoleucine, methionine, tryptophan, valine, phenylalanine, histidine), one fatty acid, and 13 vitamins and 15 minerals (Table 23.1). Most of the vitamins and minerals are micronutrients, that is, substances required in daily amounts of less than 100 mg.

TABLE 23.1. MACRONUTRIENTS AND MICRONUTRIENTS

Macronutrients	Carbohydrates
	Fats
	Proteins
Micronutrients	Fat-soluble vitamins
	Vitamins A,D,E, and K
	Water-soluble vitamins
	Vitamin B
	Thiamine (B1), riboflavin (B2), niacin, pyridoxine (B6), biotin, pantothenic acid, folic acid, cobalamin (B12),
	Vitamin C
	Minerals
	Sodium, potassium, calcium, magnesium, chloride, phosphorus
	Trace elements
	Iron, zinc, cobalt, copper, manganese, molybdenum, chromium, selenium, fluoride, nickel

Good nutrition is necessary for normal resistance to illness and for optimal recovery from medical and surgical procedures. The U.S. Recommended Daily Allowances (RDAs) serve as dietary guidelines for amounts of nutrients necessary for the maintenance of good nutrition of almost all “healthy people in the United States.” These amounts may not be adequate in hospitalized patients or other special situations.

DISORDERS OF NUTRITION

Part of "23 - Trace Elements, Vitamins, and Nutrition"

Inappropriate nutrition is the single greatest cause of illness in the world today. The majority of the world's population is subject to some degree of malnutrition in either protein, calories, or essential micronutrients. In developing regions of the world, childhood malnutrition is one of the leading contributors to overall disability and death (1). In contrast, obesity is the predominant nutritional disorder in industrialized nations.

Protein-Energy Malnutrition

Protein-energy malnutrition (PEM) [also called protein-calorie malnutrition (PCM)] occurs with inadequate ingestion of protein and/or calories. In developing countries, PEM predominantly affects infants and children who may manifest characteristics of the classic syndromes, kwashiorkor (protein deficiency) and marasmus (predominant energy deficiency). In the United States and other developed nations, PEM typically occurs secondary to other diseases and affects both adults and children.

Although some patients may have deficiencies of vitamins or minerals, the most common nutritional problem in hospitalized patients is PEM. About 40% of all hospitalized patients have some degree of nutritional deficiency, and as many as 15% suffer from clinically significant PEM. In particular, geriatric patients, alcoholic patients, and those with mental or chronic illnesses are at risk for malnutrition. Causes of malnutrition include decreased oral intake (e.g., anorexia, nausea, dysphagia, pain, gastrointestinal obstruction, poor dentition), increased nutrient losses (e.g., malabsorption, diarrhea, bleeding, protein-losing enteropathy), and increased nutrient requirements (e.g. fever, infection, neoplasm, surgery, trauma, burns).

Kwashiorkor

Kwashiorkor is a syndrome caused by severe protein deficiency despite adequate caloric intake, usually in the form of carbohydrates. In countries where kwashiorkor is endemic, the syndrome often occurs in the child subsequent to weaning.

The syndrome is characterized by growth retardation, muscle wasting, edema, swollen abdomen (from hepatomegaly and ascites), anemia, and hair loss. Dry, hyperpigmented lesions affecting the face, extremities, and perineum (“flaky paint” dermititis) also are characteristic. Children with kwashiorkor typically are lethargic and anorectic. Bradycardia, hypothermia, decreased blood pressure, and cardiac failure also may occur.

In kwashiorkor, body weight may be low or normal as a result of edema or anasarca. Unlike marasmus (see below), subcutaneous fat is present (albeit reduced), and increased fat in the liver (steatosis) produces hepatomegaly. An increased susceptibility to infection presumably results from a defective barrier function in the skin and mucous membranes, impaired granulocyte, humoral, and cell-mediated immune functions, and impaired complement activity. Without treatment, kwashiorkor is fatal, and up to 40% die despite medical intervention.

There are no characteristic biochemical findings of this condition. Serum albumin typically is less than 2.8 g/dL. Prealbumin and retinol-binding protein are similarly diminished. Serum

transferrin may be low, normal, or even slightly elevated if there is coexistent iron deficiency. Electrolyte abnormalities are common especially hypokalemia and hypophoshatemia. Anemia usually is normochromic and normocytic.

Marasmus

Marasmus is a condition that develops from deficient calories in the form of both protein and carbohydrate. These patients exhibit wasting of both muscle and fat. In contrast to kwashiorkor where body fat may be normal, subcutaneous fat is extremely depleted in marasmus-affected children. Kwashiorkor dermititis and edema do not occur in marasmus, and the abdomen is not distended. Affected children typically have a good appetite. Serum albumin and other protein markers often are normal.

Secondary PEM

The presentation of secondary forms of PEM varies greatly depending upon the previous nutritional status of the individual and the chronicity of the associated disease. Chronic disease states may produce a marasmus-like malnutrition. These patients who suffer from chronic conditions have gradual muscle wasting. Serum albumin may be normal or low but rarely less than 2.8 g/dL. More acute illnesses such as sepsis, burns, or trauma tend to produce a more kwashiorkor-like presentation. Edema and ascites may occur and serum albumin may diminish to less than 2.8 g/dL. Recognition and diagnosis of secondary PEM may be difficult. (See Nutritional Assessment).

Obesity

Approximately 25% of Americans are overweight, and about 10% are considered severely obese. Many different systems for classifying body fat and weight exist. Ideal body weight often is provided in terms of body height in prepared tables such as the Metropolitan Life Insurance Company tables. Ideal body weight also may be calculated as 100 pounds plus 5 pounds for every inch above 5 feet in height for women and 106 pounds plus 6 pounds for every inch above 5 feet for men. Adjustments should be made for type of body frame.

A simple but somewhat imprecise definition of obesity is a proportion of fat of more than 20% body weight in men and more than 30% in women (see Nutritional Assessement below for body fat measurements). A useful calculation for classification of obesity is the body mass index (BMI) which is calculated as follows: BMI = weight in kilograms/(height in meters)^2 or weight in pounds/(height in inches)^2 × 703.1 (1a).

BMI	
Grade 0	< 25
Grade I	25–29.9
Grade II	30–40
Grade III	> 40

Risk of health-related disorders increases as the BMI exceeds 25. Obesity may initiate or exacerbate certain conditions such as insulin resistance, diabetes mellitus, hypertension, hypertriglyceridemia, cardiovascular disease, chronic hypoxia and other respiratory problems, sleep apnea, thromboembolism, cancer, cholecystitis, and arthritis.

The obvious treatment for obesity is the reduction of calories. For a diet to be effective, an individual must intake a negative caloric balance. Weight loss also is promoted by increased energy utilization. When obesity reaches a crisis situation, surgical intervention may be warranted.

During the early 1990s, combination therapy for weight loss using the two amphetamine agents, fenfluramine and phenteramine, (known as “fen phen”) became widely popular. Despite its clinical effectiveness and commercial success, reports of valvular heart disease resulting from fenfluramine necessitated its withdrawal from the market in 1997. Phenteramine still is available. Typically, amphetamines and other anorectic drugs are utilized only as a short-term adjunctive therapy to diet and exercise because of their potential toxicity, addictive properties, and abuse potential.

NUTRITIONAL SUPPORT AND ASSESSMENT

Part of "23 - Trace Elements, Vitamins, and Nutrition"

In hospitalized patients, untreated PEM is associated with a three-fold increase in morbidity and mortality. PEM impairs the recovery of hospital patients, increases the risk of infection, and delays wound and tissue repair. The benefits of treating malnutrition are potentially great in terms of improved outcome, reduced morbidity, and reduced length of stay. Furthermore, early intervention may be the key to maximizing benefits in hospitalized patients. Identification of malnourished patients and patients at risk for developing PEM involves both thorough clinical evaluation and measurement of laboratory parameters.

Parenteral Nutrition

Nutritional support of the hospitalized patient may take the form of enteral or parenteral feeding. Patients who cannot be fed enterally may be given nutrients through intravenous routes. Parenteral nutrition may be in the form of peripheral parenteral nutrition (PPN) or total parenteral nutrition (TPN).

Parenteral nutrition is indicated for patients who cannot be fed orally or enterally for more than 5 to 7 days. PPN is delivered

through ordinary peripheral vein catheters and may be preferred for short-term support (less than 2 weeks) because TPN requires central venous catheter insertion and thus carries a higher risk of mechanical complications (pneumothorax, great vessel injury, and brachial plexus injury from catheter insertion) and infection. Only solutions of 1,000 mOsm or less are administered as PPN. High osmolarity solutions cannot be administered through peripheral veins because of the occurrence of phlebitis.

Parenteral solutions are composed of amino acids, dextrose, and lipid emulsions. Electrolytes, vitamins, trace minerals, and medications are added as necessary. Complications of parenteral nutrition include thrombosis, infection (focal catheter wound infections, sepsis), hyperglycemia, hypoglycemia, fluid and electrolyte abnormalities, hepatic abnormalities (elevation of liver function tests, steatosis), biliary disease (cholelithiasis, biliary sludge), micronutrient deficiencies (zinc, iron, vitamin B12), and elevation of BUN.

Nutritional Assessment

Selection of patients for nutritional support and monitoring therapy are based largely on clinical assessment and criteria. The cornerstone of nutritional assessment is physical examination and careful history-taking. Anthropomorphic and traditional laboratory parameters may play an ancillary role. Newer laboratory markers may be useful in identifying early malnutrition and in monitoring the effectiveness of nutritional intervention.

Physical Examination

Physical examination provides a number of clues about the general nutritional status of the patient and specific deficiency states. For example, the loss of subcutaneous fat and muscle indicates general nutritional deficiency while various skin, hair, and mucous membrane changes are characteristic of vitamin deficiencies.

A careful nutritional history and dietary evaluation are important components in the assessment of nutritional status. However, this information is not always reliable. A complete and accurate dietary history is difficult to obtain. Furthermore, direct measurement of consumption over a fixed period does not assess the natural diet nor can it be extrapolated over a long timeframe.

Anthropomorphic Measurements

Anthropomorphic measurements include body weight and other dimensions. A body weight less than 85% of ideal usually indicates a nutritional deficiency. Caliper measurements of skinfold thickness at multiple sites (triceps, biceps, subscapular, and suprailiac areas) are used to estimate body fat and to monitor nutritional status changes. Other measurements include abdominal/limb circumference, body density (water immersion or isotope dilution with ^{40}K), electromagnetic conductivity, bioelectric impedance, and computed tomography (CT) scanning. Disadvantages of anthropomorphic measurements include lack of reliability and reproducibility. Some methods are extremely cumbersome and are therefore impractical for routine use.

Laboratory Parameters

Laboratory parameters may be useful in the evaluation of nutritional status. As our understanding of the importance of adequate nutrition increases, improved chemical markers have been sought for new functions. Nutritional markers are needed to detect malnutrition in its early stages, to identify patients mostly likely to benefit from nutritional intervention, and to monitor the effectiveness of nutritional support. Measures of visceral protein status include the serum proteins, albumin, transferrin, prealbumin, and retinol-binding protein (Table 23.2). These proteins are all synthesized in the liver but vary widely with regard to biologic half-life and clinical utility. The ideal biological marker for nutritional status is sensitive to small increments and losses, responds rapidly to these changes, and is not affected by conditions unrelated to nutritional status. To fulfill these requirements, the marker should have a short half-life, small body pool, a rapid synthetic rate and constant catabolic rate. In addition, the marker should not be an acute phase reactant, and it should not be significantly affected by hydration, liver and renal disease, and other states independent of overall nutritional status. Unfortunately, no marker identified to date can fulfill these requirements.

Serum Albumin

Serum albumin is the traditional biochemical marker for protein nutritional status. Levels greater than 3.5 g/dL are considered normal while less than 3.0 g/dL suggests PEM. Albumin measurements are readily available and strongly correlate with morbidity and mortality in hospitalized patients (2, 3). Patients with serum albumin levels of less than 3.5 g/dL have a fourfold increase in postoperative complications and a sixfold increase in mortality (4). A serum albumin of less than 2 g/dL is associated with a 30-day mortality rate of 62% (5). Malnourished patients also have longer hospital stays and increased numbers of infections and other complications (6).

Albumin has a long biological half-life of about 20 days and a large extravascular body pool (7). Consequently, the serum albumin level responds very slowly to nutritional changes. Albumin is an acceptable marker of long-term nutrition and chronic protein deficiency, but it lacks the rapid responsiveness and sensitivity necessary for short-term monitoring (8). Albumin levels often are depressed only in moderate to severe malnutrition. In addition, albumin is nonspecific. Extravascular stores are mobilized during stress and may falsely elevate albumin levels in malnutrition. Depressed albumin occurs in other conditions including liver and renal disease, intravenous fluid use, the acute phase response, and protein-losing states.

Serum Transferrin

Like albumin, serum transferrin has been used for the general assessment of visceral protein status despite suboptimal sensitivity and specificity. Transferrin levels of less than 100 mg/dL, 100 to 150 mg/dL and 150 to 200 mg/dL have been used to identify severe, moderate, and mild protein malnutrition in patients with normal hepatic and renal function. Transferrin's half-life is

TABLE 23.2. SERUM PROTEINS AS NUTRITIONAL MARKERS

Serum Protein	Half-Life	Reference Ranges	Advantages	Disadvantages
Albumin	19 days	>3.5 g/dL is normal < 3.5 g/dL is associated with delayed healing < 3.0 g/dL suggests PEM < 2.5 g/dL is a strong predictor of increased morbidity and mortality in both hospitalized and ambulatory patients < 2.0 g/dL indicates severe PEM and is associated with a 60% 30-day mortality rate	• Established use in assessment of various groups of patients • Marker for chronic protein deficiency • Easy to perform, inexpensive	• Due to its long half-life of 19 days and large body pool, albumin changes too slowly and does not reflect short-term nutritional changes (e.g., feeding after stress remission). Not suitable for monitoring response to nutritional intervention. • Not specific. Conditions independent of nutritional status may lower albumin e.g., cirrhosis, other liver disease, nephrotic syndromes, other renal disease, use of intravenous fluids, overhydration, burns, protein-losing enteropathies. Elevated in dehydration.
Transferrin	8–10 days	>200 mg/dL normal nutritional state 150–200 mg/dL mild nutritional depletion 100–150 mg/dL moderate < 100 mg/dl severe visceral protein depletion	• Shorter half-life and smaller body pool than serum albumin. Hence, earlier response to protein depletion than serum albumin.	• Limited use in monitoring response to nutritional therapy. Response time better than albumin but not rapid enough. • Not specific. Decreased in cirrhosis, protein-losing states, neoplastic disease, use of high concentration antibiotics (tetracycline, cephalosporins, aminoglycosides). Increased with iron deficiency, pregnancy, use of oral contraceptives, and hepatitis.
Prealbumin (transthyretin)	2 days	< 11 mg/dL signifies inadequate nutritional support Rise of less than 2 mg/dL per week indicates inadequate nutritional support or inadequate response	• Very short half-life and small body pool. Rapid response to diminished nutritional intake and nutritional support. • Sensitive indicator of effectiveness of nutritional support. Daily increments of up to 1 mg/dL indicate good visceral protein anabolic response to nutritional support • May be sensitive enough to allow for early assessment and intervention in patients at risk for malnutrition. • Not significantly affected by hydration status. Less affected by liver disease as other markers. More specific. • High tryptophan content and high essential to non-essential amino acid ratio. More accurate marker for nutritional status.	• Some increase in renal failure (but not to the same extent as seen with RBP) • Decreased in cirrhosis
Retinal-binding protein	12 hours		• Shortest half-life and small body pool. Rapidly responds to changes in nutrition. Good for monitoring short-term nutritional changes.	• Increased and unreliable in renal failure. • Unpredictable changes in liver disease.

shorter (8 days) than that of albumin and therefore, transferrin responds sooner than albumin to nutritional changes but not rapidly enough for use in daily monitoring of nutritional therapy. Serial transferrin levels lack sensitivity and cannot detect critical increments and decrements during the course of nutritional therapy (6). Transferrin levels also are not specific to nutritional status. Cirrhosis and protein-losing states will affect all the serum protein markers to some degree, but transferrin has

poorer specificity because of the fact that neoplastic disease and the use of high concentration antibiotics (tetracycline, cephalosporins, aminoglycosides) decrease levels and iron deficiency, pregnancy, hepatitis, and the use of oral contraceptives increase transferring (7).

Prealbumin

Prealbumin (Transthyretin) has a short half-life of 1.9 days and small body pool, and may be the best serum protein for short-term nutritional evaluation (8). Prealbumin is a sensitive and rapidly responsive nutritional marker used both to detect diminished nutritional intake and to monitor the effectiveness of nutritional intervention. Decrements in prealbumin occur in as little as three days following inadequate nutritional intake. Daily increments of prealbumin (up to 1 mg/dL) indicate a positive nitrogen balance reflecting a good visceral protein anabolic response to nutritional support while increments of less than 2 mg/dL per week signify an inadequate response or inadequate level of nutritional support (7, 9).

Prealbumin has a high essential to nonessential amino-acid ratio and high tryptophan content (Tryptophan plays an important role in protein synthesis initiation). Consequently, prealbumin is thought to be an accurate reflection of nutritional status. Furthermore, prealbumin has good specificity compared to that of other serum proteins. Though prealbumin may be diminished in end-stage liver disease, other conditions such as renal disease, dehydration, and overhydration appear to have lesser effect on prealbumin.

Retinol Binding Protein

Retinol binding protein (RBP) has the shortest half-life (12 hours) and thus may be the most sensitive index of protein status and the most rapid marker to respond to nutritional intervention. However, RBP has inferior specificity compared to that of prealbumin (7, 10). In particular, RBP is elevated considerably in patients with renal dysfunction and cannot be used reliably in these patients.

Other Laboratory Markers

For many clinical situations (e.g., preoperatively), a general nutritional assessment may be made by taking only simple body dimensions such as body weight and height and by measuring one or two of the visceral protein markers in Table 23.2. Less common laboratory measurements are discussed below. The use of these markers often is cumbersome and does not yield any additional clinical information. As such, their application is reserved for selected clinical situations and research.

Nitrogen balance is an older parameter that has been used in monitoring malnourished patients as an indicator of adequate nutritional support and anabolism (7). It is the difference between nitrogen intake and nitrogen excretion. In the healthy, well-nourished adult population, the nitrogen balance should approach zero. During periods of decreased intake, stress, or illness, nitrogen losses may exceed intake causing a negative nitrogen balance. With recovery and proper nutritional support, the nitrogen balance should become positive. The nitrogen balance is calculated as follows:

Nitrogen balance = nitrogen intake - nitrogen loss

Nitrogen intake = 24-hour protein intake (grams)/6.25

Nitrogen loss = 24-hour urinary urea nitrogen/total volume + non-urea nitrogen losses

Ninety percent of nitrogen is excreted renally and urinary nitrogen is comprised mostly of urea (80% to 90%). Direct measurement of total urinary nitrogen is technically difficult. The 24-hour measurement of urinary urea nitrogen usually is the only measurement actually made, and a factor of 2.5 g is used to approximate non-urea losses, that is, non-urea urinary nitrogen (ammonia, creatinine, uric acid, and other nitrogen-containing components) and fecal and skin losses which are small.

Somatic muscle mass is catabolized to preserve visceral protein. Because anthropometric measurements often are not practical, **urinary creatinine** excretion (which reflects protein derived from the somatic compartment) has been used to estimate lean body mass (LBM) as shown below where the average of three consecutive 24-hour urine creatinine measurements is used.

LBM (kg) = 7.38 + 0.02909 × creatinine (mg/day)

The **creatinine height index** (CHI) is the ratio of 24-hour urinary creatinine to ideal creatinine excretion based on height, gender, and age. In most men, CHI values of greater than 7.4 mg/cm and 8.4 mg/cm qualify as moderate and severe malnutrition, respectively. The specificity of the CHI is limited by variation because of renal function, dietary protein, exercise, infections, fever, and other parameters that influence creatinine.

Somatic muscle also may be evaluated by 24-hour urinary excretion of **3-methylhistidine**, which is derived from skeletal muscle actinomyosin. Urinary excretion of 3-methylhistidine increases in muscle catabolism. **Urinary amino acid** patterns may denote chronic protein deficiency. Typically, the total amount of urinary amino acids decreases while relative amounts of branched-chain amino acids are diminished and glycine and serine are augmented. Urinary **hydroxyproline** has been used to assess collagen catabolism and maybe other indicator protein status. These measurements (urinary 3-methylhistidine, amino acids, and hydroxyproline) require special methods and are not readily available.

In contrast, **urinary ketones** and **glucose** are measured easily. Their utility in nutritional assessment is limited to certain clinical situations. The presence of ketones in the urine indicates either an inappropriate use of a TPN solution or a starvation state in which fat stores are mobilized for calories. The ability to develop ketosis after a fast is evidence of fat stores.

Excess glucose in the urine may indicate that spilling of TPN fluid through the glomerulus is taking place. The resulting loss of calories, fluids, and electrolytes may hamper the calculation of other nutritional parameters.

Skin sensitivity testing using the intradermal injection of antigens has been used to demonstrate an impaired immune response in states of severe nutritional deficiency. This test is not specific.

VITAMINS

Vitamins are micronutrients that are not endogenously synthesized or are not made in sufficient amounts to meet physiologic requirements. Many vitamins serve as coenzymes in critical biochemical pathways. The water-soluble vitamins include thiamine (B1), riboflavin (B2), niacin, pyridoxine (B6), biotin, pantothenic acid, folic acid, cobalamin (B12), and vitamin C. Vitamins A, D, E, and K are fat-soluble.

Deficiency and toxicity of vitamins is discussed below and summarized in Table 23.3. Inadequate dietary intake or the use of drugs that destroy vitamins may lead to deficiency states. In the case of fat-soluble vitamins, malabsorption as a result of biliary or pancreatic insufficiency is an additional risk factor. Severe toxicity and fatality following vitamin overdose are rare. Fat-soluble vitamins such as vitamins A and D may cause a syndrome of acute toxicity. Typically, excess water-soluble vitamins are excreted in the urine following acute ingestion. On the other hand, chronic toxicity may develop regardless of the vitamin's water solubility. The laboratory evaluation of each vitamin is summarized in Table 23.4.

TABLE 23.3. VITAMINS

Vitamin	Dietary Source	Deficiency State	Risk Groups for Deficiency	Toxic States
Vitamin A	Meat, fat, eggs, dairy products	Night blindness Complete blindness Skin abnormalities with disordered keratinization Impaired skeletal growth, depressed immune function	Alcoholism, hepatic disease, fat malabsorption, pancreatitis, cystic fibrosis Underdeveloped countries–common dietary deficit	Acute toxicity– headache, drowsiness, blurred vision, blurred vision, abdominal pain, nausea, vomiting, weakness Chronic toxicity–alopecia, multiple bone and skin abnormalities, anemia, hepatotoxicity Increased intracranial pressure, bulging fontanelles in newborns Teratogenicity Carotenemia
Vitamin D	Fish, liver, eggs, butter	Defective mineralization of bone resulting in rickets in children and osteomalacia in adults In rickets, hypocalcemia and resulting tetany, growth retardation with skeletal deformities and bone fractures, listlessness, and muscular weakness	Inadequate nutrition or sunlight exposure in winter at northern latitudes, malabsorption due to biliary or pancreatic insufficiency, liver or renal failure Drugs such as phenytoin, phenobarbital, rifampin that increased degradation	Chronic ingestion may cause hypercalcemia (>11 mg/dL)
Vitamin E (tocopherols)	Vegetable oils, egg yolk, liver, milk	Areflexia, gait disturbance, diminished proprioception	Fat malabsorption, abetalipoproteinemia, cholestatic liver disease, diets with high polyunsaturated fat content	Severe toxicity is rare. Vitamin E antagonizes vitamin K so patients on oral anticoagulation therapy may have excess platelet time prolongation. Premature infants given parenteral vitamin K may develop hepatosplenomegaly, ascites and jaundice.
Vitamin K	Green vegetables	Coagulopathy	Malabsorption, poor dietary intake, sterilization of GI tract from long-term oral antibiotics, newborns	
Vitamin B_1 (thiamine)	Milk, grains, yeast	Wet beriberi, dry beriberi, Wernicke's encephalopathy, Korsakoff's syndrome	Alcoholics, chronic dialysis, glucose administration, diets high in milled rice, ingestion of thiaminases (raw foods)	
Vitamin B_2 (riboflavin)	Liver, kidneys, raw milk, eggs, cheese, fish, green vegetables	Glossitis, cheilosis, dermatitis, neuropathy, and anemia	Isolated deficiency is rare	
Niacin	Whole grains, liver, beef, fruit, vegetables, production from tryptophan	Pellagra (dermatitis, dementia, diarrhea, and death)	Milled corn, carcinoid syndrome, Hartnup disease, isoniazid administration	Overdose causes histamine release and flushing, pruritus, nausea, vomiting, and diarrhea. Chronic overdose may cause hepatitis.
Vitamin B_6 (pyridoxine)	Meat, poultry, fish, yeast, seeds, bran	Dermatitis with cleilosis (dry lips) and glossitis, stomatitis, anemia, depression, and confusion. Infants may present with seizures.	Exclusive ingestion of processed foods. Alcoholism. Use of certain drugs such as isoniazid, hydralazine, penicillamine, cycloserine, and theophylline	Chronic overdose may cause peripheral neuropathy with paresthesias, ataxia.
Vitamin C (ascorbic acid)	Citrus fruits, berries, melons, tomatoes, green peppers, leafy vegetables	Scurvy–arthralgia, fatigue, subcutaneous hemorrhage in the gingiva, extremities, joints and nailbeds (splinter hemorrhages), gingivitis, loose teeth, and hyperkerotic, papular skin lesions.	Diet devoid of fruits and vegetables	Increased oxalate Kidney stones, nephropathy
Biotin	Liver, kidney, egg yolk, milk, fish, nuts	Peritoneal dermatitis, conjunctivitis, alopecia, ataxia Developmental delay	In multiple carboxylase deficiency (MDC) or following raw egg white consumption or parenteral nourishment	

TABLE 23.4. LABORATORY EVALUATION OF VITAMIN STATUS

Vitamin	Laboratory Evaluation of Nutritional Status
Vitamin A	Measurement of serum vitamin A is technically complex and requires protein precipitation and extraction steps. High-pressure liquid chromatography (HPLC) is the preferred method for analysis and is able to separate retinyl esters.[16]
	Serum vitamin A levels do not correlate well with liver stores and typically become depressed only in severe deficiency. Elevated levels indicate hypervitaminosis A. Toxicity may also be determined by measuring the retinyl ester fraction of vitamin A by HPLC. Retinyl esters normally account for 5% of total vitamin A. This proportion is elevated to greater than 30% in toxicity.
	Serum carotene is simpler to determine. Serum is denatured and extracted prior to analysis by spectrophotometry or HPLC.[17] Elevated carotene levels confirm carotenemia. Serum carotene may also be used as a screening test for fat malabsorption. Elevated carotene levels may be used to rule out steatorrhea but lower values are not specific.
Vitamin D	The best test for vitamin D status is serum 25-OH vitamin D_3 level by RIA, nephelometry, or HPLC. A fasting specimen is preferred. Normal levels are approximately 10–60 ng/mL.
	Measurement of the 1,25-OH form of vitamin D may be useful in chronic renal failure. This test is less readily available than 25-OH vitamin D. The reference range for 1,25-OH vitamin D is 15–65 pg/mL.
Vitamin E (tocopherols	Measurements of vitamin E and alpha-tocopherol employ saponification, solvent extraction, and separation by HPLC or GC. Blood levels do not correlate with body storage of the vitamin and are not considered useful in the assessment of vitamin E deficiency or toxicity. The ratio of alpha-tocopherol to total plasma lipid may be a better indicator of vitamin E body stores.[25]
Vitamin K	Direct measurements of vitamin K by chromatographic methods is rarely used. Clinically, prothrombin time is used to assess vitamin K status.
Vitamin B_1 (thiamine)	The most reliable indicator of thiamine status is erythrocyte or whole blood transketolase activity.[32] Levels less than 0.6 U/gram hemoglobin suggest deficiency.
	Deficiency is also defined as an increase in transketolase activity of greater than 15% upon in vitro addition of thiamine diphosphate.
Vitamin B_2 (riboflavin)	Blood and urine riboflavin measurements have little clinical value. The best test is red blood cell glutathione reductase activity before and after addition of FAD. An increase of more than 40% indicates a deficiency state. Less than 20% is normal.
Niacin	Niacin status is evaluated by measurement of urine metabolites, N1-methylnicotinamide (NMN) and its pyridone (pyr).[36] Less than 1.2 mg per day of either corresponds to low intake.
Vitamin B_6 (pyridoxine)	Vitamin B6 may be measured by enzyme assay, radioimmunoassay or HPLC methods. Normal levels of pyridoxine and pyridoxal-5′-phosphate (PLP) are 3–30 mcg/L and 5–50 mcg/L, respectively. Levels less than 1 mcg and 3 mcg indicate deficiency.
	The tryptophan loading test may be performed for deficiency screening (40). Urinary xanthurenic acid is measured before and after oral tryptophan loading. Normal is greater than 50 micromol/24 hours.
	Erythrocyte aspartate aminotransferase (eAST) will be increased following addition of vitamin B6. A large difference between unsupplemented and supplemented eAST indicates deficiency.
Vitamin C (ascorbic acid)	Vitamin C is measured by colorimetric, fluorometric, and HPLC methods. Serum ascorbic acid levels of less than 0.2 mg/dL (11 mol/L) suggest deficiency.
	Vitamin C status may also be assessed with two consecutive 24-hour urine measurements of ascorbic acid; one at baseline, the other, after 2 days of 200 mg oral ascorbic acid. Vitamin C deficiency is defined by a concentration of less than 50 mg ascorbic acid in the second 24-hour urine collection
Biotin	Biotin assay not readily available in most settings. Diagnosis of multiple carboxylase deficiency (MCD) requires enzyme studies.

RIA, radioimmuno assay; GC, gas chromatography; FAD, flavin adenine dinucleotide.

Vitamin A

Vitamin A consists of a group of compounds that have the biological action of retinol. The main vitamin A compounds are retinol (vitamin A_1) and 3-dehydroretinol (vitamin A_2). The alcohol, retinol, is the form that is absorbed and delivered to peripheral tissues bound to retinol-binding protein. The aldehyde, retinal, is required for vision. Retinol combines with fatty acids to form body reserves of vitamin A, which are abundant in the liver. When needed by the body, these retinyl esters are hydrolyzed. Recent studies have elucidated the important role that various retinol-binding proteins play in the regulation of vitamin A absorption, transport, storage, and metabolism (11, 12).

Retinol derivatives (called retinoids) include retinal and the ester, retinoic acid, as well as numerous other synthetic forms, which may or may not possess biologic activity. Liver, fat, eggs, and dairy products are good sources of retinoids. Beta-carotene and other carotenoids are precursors of vitamin A and constitute the yellow-orange pigments found in various vegetables such as carrots.

Vitamin A and carotenoids are not soluble in water. For this reason, they cannot be absorbed directly in the gastrointestinal tract but must be coabsorbed with fats as part of the micelles that form in the intestine. Pancreatic enzymes also are required in order to hydrolyze retinol fatty acid esters. Thus, absorption of vitamin A and utilization of body stores depend upon adequate biliary and pancreatic function.

Vitamin A is essential for vision, epithelial differentiation, and mucus secretion. During dark adaption, retinol in the eye is oxidized to retinal and isomerized allowing its interaction with opsin to form rhodopsin. Rhodopsin is the light-sensitive substance which allows night vision.

Deficiency/Toxicity States

Vitamin A deficiency occurs commonly in developing countries and is a leading cause of blindness in the young (13). Other individuals at risk for vitamin A deficiency include those with fat malabsorption, pancreatic disease, cystic fibrosis, hepatic disease, and alcoholism.

Deficiency of vitamin A causes ocular and dermatologic abnormalities. The initial symptom of vitamin A deficiency is night blindness. With severe deficiency, total blindness may occur. Vitamin A deficiency also may present with skin abnormalities caused by disordered keratinization. Impaired skeletal growth, depressed immune function, and increased morbidity from measles also occur.

Acute vitamin A toxicity may occur following overdoses in excess of 500,000 IU and is characterized by drowsiness, irritability, blurred vision, abdominal pain, anorexia, nausea, vomiting, increased intracranial pressure, intense headache, and muscle weakness (14). After several days, skin desquamation occurs. Chronic toxicity may occur from taking only moderately high amounts over extended periods of time. Alopecia, multiple skin and bone abnormalities, anemia and hepatotoxicity occur with chronic over-supplementation (daily doses in excess of 35,000 IU). In children, as little as 1,500 IU (about four times the recommended daily allowance) may cause chronic toxicity. Doses of 25,000 IU or more during early pregnancy may be teratogenic.

There are over 1,000 synthetic retinoid dervatives. Three major ones, retinol, isotretinoin, and etretinate, may be used in the treatment of psoriasis and acne (15). When used systemically, these retinoid compounds may cause ocular and liver toxicity as well as abnormal serum lipid profiles. They are also teratogens. An excess of carotene from a diet high in carotenoid-containing vegetables such as lettuce and carrots may cause an orange discoloration of the skin, which is not harmful but may be misinterpreted as jaundice. Carotenemia does not lead to hypervitaminosis A probably because of regulation of carotene conversion to vitamin A.

Measurement

Measurement of serum vitamin A is technically complex and requires protein precipitation and extraction steps. The Carr-Price method is a colormetic procedure employing antimony trichloride. High-performance liquid chromatography (HPLC) is a preferred method and is able to separate retinyl esters (16).

Serum vitamin A levels do not correlate well with liver stores and typically become depressed only in severe deficiency. Elevated levels indicate hypervitaminosis A. Toxicity also may be determined by measuring the retinyl ester fraction of vitamin A by HPLC. Retinyl esters normally account for 5% of total vitamin A. This proportion is elevated to greater than 30% in toxicity.

The assay for serum carotene is somewhat simpler to perform. Serum is denatured and extracted prior to analysis by spectrophotometry or HPLC (17). Nonetheless, like vitamin A, serum carotene is not a good indicator of vitamin A status, and its utility is limited. Elevated carotene levels confirm carotenemia. Serum carotene also may be used as a screening test for fat malabsorption. Elevated carotene levels may be used to rule out steatorrhea but lower values are not specific.

Vitamin D

Vitamin D is a fat-soluble vitamin involved in mineral homeostasis. In particular, vitamin D plays an important role in osteoblast function, parathyroid hormone release, and intestinal absorption and renal excretion of calcium (18, 19).

Vitamin D refers to both vitamin D_3 (cholecalciferol) and vitamin D_2 (ergocalciferol). Most of the vitamin D in the body is vitamin D_3 and a much smaller proportion is vitamin D_2. Fish, liver, eggs, and butter are good sources of vitamin D_3. It also may be endogenously synthesized; Ultraviolet (UV) light catalyzes its formation from 7-dehydrocholesterol in the skin. In contrast, vitamin D_2 derives only from plants. Because the amount of vitamin D obtained from food sources and endogenous synthesis often is insufficient, milk has been fortified with vitamin D in most developed countries. Each quart of milk typically contains 10 mcg (or 400 IU) of vitamin D. Vitamin D supplements may consist of vitamin D_3 or vitamin D_2.

Both vitamin D_3 and vitamin D_2 are inactive and undergo hydroxylation to the weakly potent 25-OH vitamin D in the liver and to the most active form of the vitamin, 1,25-OH vitamin D, in the kidneys.

Clinical Syndromes

In vitamin D deficiency, mineralization of bone is defective resulting in rickets in children and osteomalacia in adults. Rickets is characterized by hypocalcemia, tetany, growth retardation with skeletal deformities and bone fractures, listlessness, and muscular weakness.

Vitamin D deficiency in the adult may occur as a result of inadequate nutrition, inadequate sunlight exposure, residence in

northern latitudes, and malabsorption as a result of biliary or pancreatic insufficiency. Because the active form of the vitamin is dependent upon hydroxylation, hepatic and renal disease also may cause manifestations of deficiency. Chronic use of certain drugs such as phenytoin, phenobarbital, and rifampin will increase hepatic metabolism of 25-OH vitamin D.

Excess supplementation may cause acute hypercalcemia, altered mental status, seizures, renal stones, and metastatic calcifications.

Measurement

The major circulating form of vitamin D is 25-OH vitamin D_3, and the best test for assessment of vitamin D status is the serum 25-OH vitamin D_3 level by radioimmunoassay, immunonephelometry, or HPLC methods (20, 21). A fasting specimen is preferred, and normal levels are approximately 10 to 60 ng/mL. Indications for the measurement of the 1,25-OH form of vitamin D are limited to cases of chronic renal failure where 1-hydroxylation is deficient. This test is less readily available than 25-OH vitamin D, and the reference range for 1,25-OH vitamin D is 15 to 65 pg/mL. The best indicator of toxicity is also serum 25-OH vitamin D_3, because 1,25-OH vitamin D levels may not be elevated. Serum calcium is another important measurement in toxicity and typically exceeds 11 mg/dL.

Vitamin E

Vitamin E includes a group of compounds known as tocopherols and tocotrienols. α-Tocopherol is the most active form. Tocopherols are found ubiquitously in foods. Vegetable oils, egg yolk, liver, and milk are particularly good sources of vitamin E (22).

Vitamin E is a fat-soluble vitamin. The daily requirement is closely related to the dietary intake of polyunsaturated fats. Vitamin E is stored in various tissues, especially adipose.

Vitamin E is an antioxidant and as such neutralizes free radicals in red blood cells, neurons, and other tissues (23). In particular, vitamin E is thought to protect polyunsaturated fatty acids in cell membranes.

Clinical Syndromes

Vitamin E deficiency occurs in premature infants and in children and adults with cystic fibrosis, biliary disease, other causes of malabsorption, and abetalipoproteinemia. Erythrocytes have increased fragility resulting in hemolytic anemia. Another feature of vitamin E deficiency is a progressive neurologic disease with ataxia, areflexia, gait disturbance, and diminished proprioception and vibratory sensation (24).

Unlike the fat-soluble vitamins A and D, vitamin E has not been reported to cause severe toxicity in the general population. Excessive intake may cause gastrointestinal upset, lethargy, and headache. Toxicity does occur in patients on oral anticoagulation therapy because vitamin E can potentiate warfarin compounds by antagonizing vitamin K. Premature infants may develop hepatosplenomegaly with ascites, cholestasis, and thrombocytopenia in response to parenteral vitamin E.

Measurement

Methods for the measurement of vitamin E or α-tocopherol are not readily available and typically employ saponification, solvent extraction, and separation by HPLC or gas chromatography (GC). Blood levels do not correlate with body storage of the vitamin and are not considered useful in the assessment of vitamin E deficiency or toxicity. The ratio of α-tocopherol to total plasma lipid may be a better indicator of vitamin E body stores (25).

Vitamin K

There are a number of different vitamin K compounds chemically known as methylnaphthoquinones. Vitamin K_1 (phylloquinone) is found in green vegetables. Intestinal bacteria produce Vitamin K_2 or menaquinone. The contribution of bacterially produced vitamin K_2 to total vitamin K requirement is not known (26, 27).

Vitamin K is required for the posttranslational gamma-carboxylation of glutamic acid residues in prothrombin, clotting factors VII, IX, and X, and inhibitory proteins C and S (28).

Clinical Syndromes

Vitamin K deficiency is characterized by coagulopathy with prolonged prothrombin time and possible hemorrhagic disease. Deficiency occurs with fat malabsorption or biliary obstruction. Parenteral nutrition and prolonged use of broad-spectrum antibiotics also may produce vitamin K deficiency. Newborns are particularly susceptible because of immature liver function and lack of vitamin K synthesis by gastrointestinal bacteria.

Measurement

A variety of direct measurements are available for vitamin K, involving chromatography or spectrophotometry. However, these methods rarely are used except in research applications and it is usually adequate to rely on the measurement of the prothrombin time to assess the vitamin K status (29).

Vitamin B_1 (Thiamine)

Vitamin B_1 is an essential component of the coenzyme, thiamine pyrophosphate. Thiamine diphosphate is the predominant form of the vitamin.

The vitamin is found in most plant and animal tissues, but the best source is unrefined cereal grains, enriched flour, liver, heart, and kidney. Thiamine is removed from grains during the milling process.

Thiamine is essential for the decarboxylation of keto acids and is therefore important in the Krebs cycle. It also is required for transketolase activity which is essential to carbohydrate metabolism. Thiamine also plays an important role in neurons and myelin lipid maintenance.

Clinical Syndromes

Thiamine deficiency occurs in populations that depend on rice as a staple in their diet. The milling process removes the thiamine-rich outer layers of grains. Thiaminases in other food products such as coffee, tea, fish, shellfish, and raw meat also can

destroy thiamine. Loss of thiamine also may occur with dialysis, diarrhea, and diuretic use. Daily requirements for thiamine are increased by carbohydrate intake and diminished by fat.

Body stores of thiamine are limited, and manifestations of deficiency occur in as little as a week on a deficient diet. Thiamine deficiency occurs in alcoholics, in persons with diets of predominantly milled rice and those containing thiaminases, and in certain other settings such as chronic dialysis and refeeding after starvation. Glucose administration may precipitate clinical deficiency in thiamine-depleted patients or worsen the symptoms of Wernicke's encephalopathy.

Thiamine deficiency affects the cardiovascular and central and peripheral nervous systems. Manifestations of deficiency include wet beriberi, dry beriberi, Wernicke's encephalopathy, and Korsakoff's syndrome. Wet beriberi is characterized by severe edema and congestive heart failure. Dry beriberi is a peripheral neuropathy that is caused by degeneration of myelin sheathes. Typically, distal portions of the extremities are more affected than proximal ends. Findings include both sensory (numbness, tingling, with or without pain) and motor changes.

Ophthalmoplegia, horizontal nystagmus, ataxia, dysarthria, and an acute confusional state are features of Wernicke's encephalopathy. Confusion may progress to coma and death. Untreated Wernicke's has a 17% mortality rate (30). Korsakoff's syndrome is characterized by amnesia, learning impairment, and confabulation, and usually occurs in alcoholic patients weeks after an episode of Wernicke's encephalopathy or as Wernicke's subsides following thiamine administration (31). Remote memory typically is preserved in Korsakoff's syndrome.

Measurement

The most reliable indicator of thiamine status is erythrocyte or whole-blood transketolase activity where values less than 0.6 U/g of hemoglobin are indicative of thiamine deficiency (32). Deficiency also may be defined by an increase in transketolase activity of greater than 15% upon *in vitro* addition of thiamine diphosphate (TDP).

Direct measurement of blood thiamine is performed by HPLC and is not available for routine use. Similarly, urinary thiamine is not of clinical value.

Vitamin B_2 (Riboflavin)

Dietary sources of vitamin B_2 include liver, kidney, heart, eggs, and vegetables. Small amounts are found in cereals. The vitamin is present in raw milk but is eliminated by pasteurization (33).

Riboflavin in the coenzymes flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) are important in oxidation-reduction reactions. The precursors FMN and FAD are cofactors for electron transport, liver enzyme action, and corticosteroid production.

Clinical Syndromes

Riboflavin deficiency is characterized by glossitis, cheilosis, dermatitis, neuropathy, and anemia. However, these findings are nonspecific, and the syndrome is almost always found in association with other B vitamin deficiencies. Chronic dialysis, phenothiazines, and cyclic antidepressants may increase riboflavin requirements.

Measurement

Blood and urine measurements of riboflavin have little clinical application. The best test for vitamin B_2 nutritional status is the measurement of red blood cell glutathione reductase activity before and after addition of FAD (34). An increase in activity of less than 20% is normal whereas an increase of more than 40% indicates a deficiency state.

Niacin (Nicotinic Acid)

Niacin refers to nicotinic acid and its derivatives. Dietary sources of niacin include whole grains, beef, liver, fruit, and vegetables. Niacin is not an essential dietary requisite for the reason that it can be formed in the body from the amino acid, tryptophan. The daily dietary requirement for niacin is dependent upon tryptophan intake.

Niacin is a basic part of the coenzymes, nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP), which are utilized in oxidation-reducation reactions throughout the body.

Clinical Syndromes

Niacin deficiency causes the syndrome known as pellagra. Pellagra is characterized by dermatitis, dementia, diarrhea, and death (the four ‘Ds’). This syndrome generally occurs as a result of dietary deficiency. The dermatologic abnormalities are hyperkeratosis, hyperpigmentation, and desquamation. Mental status changes include fatigue, apathy, confusion, amnesia, hallucination, and psychosis. Diarrhea, glossitis, stomatitis, and vaginitis may occur as a result of mucosal inflammation. Pellagra is a progressive disease, and death may ensue.

Diets composed predominantly of milled corn may lead to niacin deficiency because corn is low in tryptophan and the milling process eliminates bioavailable niacin (35). Widespread supplementation of corn and grain products has diminished the occurrence of niacin deficiency as a result of dietary deficit.

Niacin deficiency may occur secondary to other conditions. Carcinoid neoplasms consume an excessive proportion of tryptophan to make serotonin. Hartnup disease is a renal tubular disorder characterized by massive urinary loss of certain amino acids including tryptophan. By depletion of tryptophan, these conditions may precipitate niacin deficiency. Isoniazid administration also may cause a pellagralike syndrome. Isoniazid inhibits pyridoxine, which is required for the conversion of tryptophan to niacin.

Excessive consumption of niacin has been reported to cause flushing, pruritus, nausea, vomiting, and diarrhea from histamine release. Chronic overdosage may cause hepatotoxicity.

Measurement

Niacin status is evaluated by measurement of principal urinary metabolites, in particular, N1-methylnicotinamide (NMN) and

N1-methyl-2-pyridone-5-carboxamide (pyr) (36). These measurements are made using HPLC (37). Urinary excretion of less than 1.2 mg per day of either NMN or pyr corresponds to low niacin intake. Plasma levels of NMN and pyr are significantly less reliable.

Vitamin B_6 (Pyridoxine)

Vitamin B_6 refers to pyridine derivatives including pyridoxine, pyridoxal, and pyridoxic acid. The phosphorylated forms of these compounds serve as cofactors for various enzymes involved in tryptophan metabolism, lipid and carbohydrate metabolism, and the syntheses of neurotransmitters, niacin, and heme.

Vitamin B_6 is ubiquitous in foods. Meat, poultry, fish, yeast, seeds, wheat germ, and bran are particularly rich in vitamin B_6. Smaller amounts are present in milk, eggs, and certain green vegetables. Vitamin B_6 compounds may be destroyed by heat or other food processing.

Clinical Syndromes

Deficiency of vitamin B_6 is rare but has occurred as a consequence of diets limited to processed food. Severe vitamin B_6 deficiency causes dermatitis with cleilosis (dry lips) and glossitis, stomatitis, sideroblastic anemia, depression, and confusion. Infants may present with colic, irritability, and seizures as a result of infant formulas that are overheated.

A significant number of drugs including isoniazid, levodopa, phenelzine, hydralazine, penicillamine, cycloserine, theophylline, oral contraceptives, and ethanol also may precipitate vitamin B_6 deficiency by increasing its loss and/or inhibiting its functions (38). Supplementation of vitamin B_6 is indicated if sideroblastic anemia develops in alcoholics and in patients taking these drugs. Genetic errors of metabolism such as xanthurenic aciduria, primary cystathioninuria, homocystinuria, and glutamate dehydrogenase deficiency also may respond to increased administration of vitamin B_6.

Chronic excessive vitamin B_6 (daily consumption of > 2 g/day) causes neurotoxicity characterized by ataxia, gait disturbance, tingling, and numbness.

Measurement

Vitamin B_6 measurements in biologic fluids are obtained by enzyme assay, radioimmunoassay, or HPLC methods (39). Vitamin B_6 status may be assessed by the measurement of pyridoxine and/or pyridoxal 5′-phosphate (PLP), the major biologically active form of vitamin B_6. Normal levels are 3 to 30 mcg/L and 5 to 50 mcg/L, respectively. Levels less than 1 mcg and 3 mcg, respectively, indicate deficiency. Total vitamin B_6 levels also may be determined and typically include pyridoxine, PLP, pyridoxal, and pyridoxic acid. In vitamin B_6 deficiency, urine pyridoxic acid levels also may be diminished.

Because tryptophan metabolism is dependent upon vitamin B_6, a functional test of vitamin B_6 may be performed by measuring urinary levels of the tryptophan metabolite, xanthurenic acid, following oral administration of 2 g of tryptophan (tryptophan loading test). With adequate vitamin B_6 status, urinary xanthurenic acid should exceed 50 µmole per 24 hours. Less than 50 µmol implies vitamin B_6 deficiency. Xanthurenic acid measurements are made by spectrophotometry, fluorimetry, and HPLC methods (40).

Vitamin B_6 is an essential cofactor of aspartate aminotransferase, and vitamin B_6 nutritional status also may be assessed by measurement of erythrocyte aspartate aminotransferase activity (eAST) before and after the addition of vitamin B_6. A large difference between the unsupplemented and supplemented activities indicates deficiency (41).

Vitamin C (Ascorbic Acid)

Vitamin C (ascorbic acid) is found abundantly in various foods including citrus fruits, berries, melons, tomatoes, green peppers, and leafy vegetables. The daily requirement for vitamin C is met by consumption of even small amounts of fruit or vegetable. Nonetheless, vitamin C is commonly supplemented.

Vitamin C functions as an antioxidant removing free radicals from the body and maintaining metal ions in a reduced state. It also plays an important role in collagen synthesis. In particular, vitamin C serves as a cofactor for protocollagen hydroxylase. Without hydroxylation, collagen cannot form a triple helix and lacks tensile strength. Deficiency of vitamin C most severely affects collagen rich in proline, which is found in blood vessels. Vitamin C also affects tyrosine metabolism and neutrophil chemotaxis, prevents tetrahydrofolate oxidation, and enhances iron absorption and storage.

Clinical Syndromes

Vitamin C deficiency is rare but may result from a diet completely lacking in citrus fruits and leafy vegetables. The most severe form of deficiency is known as scurvy. Scurvy originally was a disease of sailors in the 19th century who did not have access to fresh foods for long periods of time. It is characterized by arthralgia, fatigue, poor wound healing, gingivitis, loose teeth, hyperkerotic, papular skin lesions, and capillary fragility leading to subcutaneous hemorrhage in the gingiva, extremities, joints, and nailbeds (splinter hemorrhages).

Chronic overdosage of vitamin C may be a cause of uric acid and cystine kidney stones. In susceptible individuals, it clearly is associated with oxalate renal calculi. High levels of vitamin C in the urine will cause false-negatives with some conventional methods of dipstick chemistry (glucose, blood, bilirubin, and nitrite) and with quantitative measurements of oxalate.

Measurement

Vitamin C is measured by colorimetric, fluorometric, and HPLC methods. The colorimetric method utilizes Cu (II) to oxidize serum ascorbic acid to dehydroascorbic acid, which reacts with 2,4-dinitrophenylhydrazine to form hydrazone (red color). Hydrazone is quantitated spectrophotometrically. Serum ascorbic acid levels of less than 0.2 mg/dL (11 mol/L) are suggestive of vitamin C deficiency.

Vitamin C status also may be assessed with two consecutive 24-hour urine measurements of ascorbic acid; one at baseline,

the other, after 2 days of 200-mg oral ascorbic acid. Vitamin C deficiency is defined by a concentration of less than 50 mg ascorbic acid in the second 24-hour urine collection.

Folate and vitamin B_{12} are discussed in the chapters in hematology.

Vitamin H (Biotin)

Biotin is found in liver, kidney, pancreas, egg yolk, milk, fish, and nuts. Most dietary biotin exists in a form bound to lysine (biocytin). Biocytin must be hydrolyzed by the intestinal enzyme, biotinidase, for absorption.

Biotin serves as a coenzyme for carboxylases involved in CO_2 transfer, fatty acid synthesis and oxidation, amino acid metabolism, and in cholesterol metabolism and storage.

Clinical Syndromes

Biotin deficiency is rare but may occur in multiple carboxylase deficiency (MCD) or following prolonged raw egg white consumption or parenteral nutrition. Egg whites contain avidin, which binds biotin, prevents its absorption, and may induce biotin deficiency.

Biotin deficiency is characterized by perioral dermatitis, conjunctivitis, alopecia, and ataxia. In addition, developmental delay and more serious neurologic abnormalities may occur in the setting of MCD where biotin deficiency develops as a result of abnormal biotin metabolism. There are two types of congenital disorders leading to MCD, biotinidase deficiency and holocarboxylase synthetase deficiency (42). Organic aciduria occurs as well as hypotonia, seizures, ataxia, feeding difficulty, rash, and alopecia. The age of onset and severity are variable. Biotinidase deficiency responds to oral biotin therapy if introduced before irreversible neurologic damage. Holocarboxylase synthetase deficiency appears to respond only partially to biotin.

Measurements

Definitive diagnosis of MCD requires studies of biotinidase and holocarbosylase synthetase enzymes. Biotin assays are not readily available in most hospital settings. Biotin can be measured by isotope dilution, enzyme immunoassay, and chemiluminescence techniques (43, 44).

TRACE ELEMENTS

Part of "23 - Trace Elements, Vitamins, and Nutrition"

Some minerals are major constituents of the body such as sodium, potassium, chloride, calcium, magnesium, and phosphorus. These elements are required in relatively large quantities. Other minerals are referred to as trace elements because collectively they constitute less than 1% of total body mass. Essential trace elements are iron, zinc, copper, cobalt, manganese, molybdenum, selenium, nickel, chromium, and fluoride. These elements serve as crucial components of many different tissues and enzyme systems. The importance of silicon, tin, and vanadium in humans is not clear. Elements without any known physiologic role are aluminum, arsenic, cadmium, gold, lead, and mercury. Some of these elements such as arsenic, cadmium, lead, and mercury are collectively referred to as “heavy metals.”

Specimen Collection and Contamination

Trace elements are present in biological samples at quantities on the order of micrograms or nanograms per liter. Metal, rubber, wood, paper, skin, dust, and water in the environment, laboratory, and collection apparatus may contain comparable amounts of trace elements. Thus, trace element analysis is susceptible to external contamination from the environment, collection apparatus and containers, and laboratory. Extreme care are must be taken during specimen collection, processing, and analysis. Rubber stoppers and syringe parts must be avoided, and all needles, syringes, and glass containers should be acid-washed or otherwise treated for metal collection. In particular, chromium and manganese are known to leach from metal needles (falsely elevating levels), and selenium and lead adsorb onto glass tubes over time (falsely lowering values) (45, 46). Aluminum levels may be elevated falsely by contamination from rubber stoppers made of aluminum silicate.

The gold standard technique for trace element analysis is to use a plastic catheter and an all-plastic syringe for blood collection. Also suitable is the use of an evacuated Sherwood Monoject trace element blood collection tube [#8881-307006 for serum or #8881-307022 (EDTA) for whole blood] with a Becton Dickinson #5175 20-gauge stainless steel needle or butterfly needles from Terumo or Abbott (43, 44). Earlier reports of normal or pathologic trace element concentrations (prior to 1971) may not be reliable and there are often notable differences in reference ranges among laboratories. These problems may be explained largely by specimen contamination.

Even in the absence of specimen contamination, interpretation of trace element values can be difficult because these substances are stored in various parts of the body, and the serum or urine level may be a poor reflection of actual body stores. It is important, therefore, to select the appropriate body fluid for measurement.

Analytic Methods

As mentioned, extreme care must be taken in the laboratory to avoid sample contamination. The area for trace metal analysis should be isolated from the main clinical laboratory. Solvents must be reagent-grade in purity. Glassware must be cleaned in acid and thoroughly rinsed with purified water.

Analytical sensitivity is crucial to the accurate determination of trace elements, which may be present in quantities as low as nanograms per liter or nanomoles. Atomic absorption spectrophotometry (AAS) is the most popular method for trace metal measurement. The flame AAS mode is simpler to operate. However, the greater sensitivity of electrothermal (flameless) AAS may be required for certain trace elements.

Emission spectrometry, neutron activation analysis, mass spectrometry, and x-ray fluorescence spectrometry are applicable in principle but require expensive apparatus and expertise. Colorimetry is useful in some cases, and anodic stripping voltametry has limited applications (47, 48).

Hair Analysis

Analysis of metals in hair is available through specialized laboratories. Advantages of hair analysis are ease of sample collection, sample stability, and potential length of detection period. Hair grows at a rate of about 1 cm per month (0.4 mm/day) and depending upon the hair length, it may be possible to obtain a history of several months. Furthermore, by fractionating along the length of the hair shaft and analyzing fractions, the chronicity and nature of exposure (bolus versus continuous) may be determined.

To date, the potential benefits of hair testing have been outweighed by its limitations. External contamination of the hair is the most significant problem. Not all substances may be distributed in the body to the hair shaft, and different sections of hair grow at different rates. Pollutants, industrial fumes, sprays, and chemicals that are externally applied may contain elements and are not representative of *in vivo* exposure. Moreover, hair treatments such as bleaching, coloring, and permanent waving may remove metals from the hair shaft or may contain chemicals that interfere with the detection of metals.

Analysis of metals requires solubilizing hair, a solid matrix, and different laboratory methods are not equivalent. Differences in metal measurements also may be intrinsic to the type of hair depending upon hair color, race, sex, age, and body site. Finally, reference ranges for metals in hair have not been well established. The correlation between hair levels and those in vital affected organs is not known. Interpretation is especially troublesome in cases where values are slightly, but not markedly, elevated. In conclusion, there are insufficient data to support the use of hair samples, and caution is advised in interpretation of the results of hair metal analysis.

Deficiency and toxic states for the essential trace elements are discussed below and summarized in Table 23.5. Laboratory evaluation of trace elements is outlined in Table 23.6.

TABLE 23.5. DEFICIENCY SYNDROMES AND TOXICITY CAUSED BY ESSENTIAL TRACE ELEMENTS

Metal	Deficiency Syndrome	Toxic States
Iron	Hypochromic, microcytic anemia	Hemosiderosis, hemochromatosis
Zinc	Juvenile dwarfism (growth retardation, delayed sexual maturation), hypogonadism, hypogeusia, delayed wound healing, depression, lethargy, defective immunity	Zinc oxide fume inhalation may cause metalfume fever (chest pain, cough, headache, nausea, vomiting, chills, fever, weakness). Zinc chloride inhalation may be fatal.
	Acrodermatitis enteropathica	Chronic oversupplementation may cause sideroblastic anemia
Cobalt	See Vitamin B12 deficiency	Ingestion is cardiotoxic. Cardiomyopathy, polycythemia, and goiter are features of beerdrinker's cardiomyopathy. Potential role in uremic cardiomyopathy. Contact with cobalt fume or dust causes dermatitis, rhinitis, conjunctivitis. Chronic inhalation leads to pulmonary fibrosis.
Copper	Premature infants, patients on TPN or zinc supplementation are at risk for deficiency.	Ingestion of copper or copper salts causes hemolytic anemia, gastroenteritis
	Anemia, neutropenia Osteopenia, bone and joint deformities Hypotonia, psychomotor retardation Skin depigmentation Menke's syndrome (hereditary functional copper deficiency)	Wilson's disease (defective biliary elimination with Cu accumulation in various tissues), Kayser-Fleisher rings, hepatitis, cirrhosis, personality and behavioral changes, cognitive decline
Manganese	No well-defined deficiency syndrome. Impaired growth, decreased reproductive function, ataxia, seizures, skeletal deformities may occur.	Acute dust inhalation may cause metal fume fever. Chronic inhalation may cause psychiatric disorders and Parkinson like signs.
Molybdenum	No well-defined deficiency syndrome Mental retardation, seizures, bilateral lens dislocation, sulfaturia, and thiosulfaturia may occur	Fume inhalation may cause anorexia, weakness, fatigue, headache, myalgia, cough, diarrhea
Chromium	Atherosclerosis, impaired glucose tolerance	Hexavalent chromium compounds (chromium trioxide and chromic acid) are toxic. Ingestion may cause hemorrhagic gastroenteritis, hepatic and renal damage, coagulopathy, and hemolysis. Chronic inhalation of dusts and fumes leads to asthma, pneumoconiosis, and lung cancer. Contact causes dermatitis, skin ulceration, and nasal septum perforation.
Selenium	Cardiomyopathy (Keshan disease), osteoarthritis, skeletal muscular degeneration	Abdominal pain, nausea, vomiting, irritability, paresthesia, hyperreflexia Selenious acid ingestion may be rapidly fatal.
Fluoride		Acute ingestion of more than 5–10 mg/kg of fluoride produces vomiting, abdominal pain, gastrointestinal ulceration and hemorrhage, hypocalcemia, hypomagnesemia, tetany, muscle weakness, respiratory depression, coma, arrhythmias, and death. Chronic overexposure (doses as low as 0.1mg/kg per day for 10 to 20 years) produces mottling and skeletal fluorosis (excess bone calcification, tendon and ligament calcification, and joint fusion.) Hydrogen fluoride causes severe corrosive injury to eyes, nose, mouth, and skin. Inhalation leads to sore throat, cough, bronchospasm, chemical pneumonia, and pulmonary edema.
Nickel	None defined	Dermatitis, eczema, asthma Nickel carbonyl causes severe respiratory and neurologic toxicity.

TPN; total parenteral nutrition.

TABLE 23.6. LABORATORY EVALUATION OF ESSENTIAL TRACE ELEMENT STATUS

Metal	Laboratory Evaluation of Nutritional Status
Iron	Serum iron, total iron-binding capacity, ferritin
Zinc	Zinc measurements are made by atomic absorption spectrometry. Colorimetric methods utilizing dithiazone are obsolete.[84]
	Normal serum zinc is 600–1000 mcg/L. Serum levels only roughly correlate with body burden. Levels become depressed only in moderate and severe deficiency, and early or mild zinc deficiency is detectable.[85]
	Fever, estrogen, and stress may lower blood zinc levels. Hemolysis releases red blood cell carbonic anhydrase and falsely elevates zinc concentration. Contact with rubber parts of syringes and collection tubes must be avoided to prevent contamination.
Cobalt	Cobalt can be measured by atomic absorption spectrometry, flameless atomic absorption spectrometry, and neutron-activation analyses. These assays are difficult to perform, and there is no demonstrated clinical application for these measurements.
Copper	Copper may be identified by a colorimetric method using bathocuproine sulfonate.
	Measurements are made by atomic absorption spectrometry. Ceruloplasmin is determined by nephelometric, spectophotometric, and radial immunodiffusion techniques.
	Serum copper measures total copper (free copper and that bound to ceruloplasmin, albumin, and other proteins). Normal levels are 700–1500 mcg/L. Severe toxicity associated with levels in excess of 5000 mcg/L.[86]
	Whole blood levels may correlate better in toxicity. 24-hour urine copper levels are also elevated in toxicity. Normal daily excretion is 15 to 64 mcg.
	Ceruloplasmin is an acute phase reactant. Both serum copper and ceruloplasmin are elevated in inflammatory conditions, estrogen, and pregnancy. Serum copper may be decreased in nephropathy, malabsorption, valproic acid, ACTH, and glucocorticoids.
Manganese	Atomic absorption spectrometry may be used to measure manganese. Other possible techniques include colorimetry, neutron activation analysis, and emission spectrometry.[87] Normal serum levels are 0.4–0.8 ng/mL. Whole blood levels are higher and may be a better measure of body burden (normal 4–12 ng/mL).
	To evaluate toxicity, 24-hour urine concentration may also be useful (normal 1–10 mcg/L). In random urine, manganese content may be reported per gram creatinine. Exposed individuals typically have elevated blood and urine levels. However, if exposure is remote, signs of chronic toxicity may persist after blood and urine levels have normalized.
Molybdenum	There is no generally accepted method for the measurement of molybdenum in biological fluids.
Chromium	Extreme caution must be taken for accurate measurement of chromium. Stainless steel and all metal must be avoided during specimen collection and analysis.
	Laboratory work should be performed under laminar flow conditions. Methods are isotope ratio mass spectroscopy and graphite furnace atomic absorption spectroscopy.
	The normal range for serum chromium is between 0.05 ng/mL and 0.15 ng/mL.[88]
	Higher reference ranges reported in the past are inaccurate. Serum levels exhibit diurnal variation (higher in the morning, lower post-prandially.) Reports of higher reference ranges of up to 18 ng/mL are erroneous and likely due to sample contamination. Serum levels are elevated in acute exposure but are not helpful in the diagnosis of deficiency. Urinary chromium may also be used as a measure of chromium overexposure.
	Chromium deficiency may be demonstrated by improvement in glucose tolerance after chromium supplementation. A modified glucose tolerance test is used in which fasting, 30-, 60-, 90-, and 120-minute samples are measured.
Selenium	Selenium measurements are made by atomic absorption or fluorometry.
	The normal ranges for serum and whole blood are 50–150 mcg/L and 60–240 mcg/L, respectively. Blood is used in the evaluation of both deficiency and toxicity states.
	Serum levels reflect recent intake. Erythrocyte levels are more indicative of remote exposure. Levels may be lowered in acute illness.[89] Hair levels have been used in the diagnosis of deficiency. Normal 0.2–1.4 mcg/gram.
	Levels may be falsely elevated by use of selenium sulfide shampoo.
	24-hour urine is used to monitor exposure. Normal excretion is 15–150 mcg/L.
	Levels greater than 500 mcg/L are most likely associated with toxicity.
	Red blood cell glutathione peroxidase activity is a useful functional test.
Fluoride	Fluoride is measured by ion-specific electrode. Fluoride measurements are useful in the diagnosis of overexposure. Hypocalcemia and hyperkalemia may also occur.
	Normal range for serum fluoride is 10 to 370 mcg/L. Values of up to 450 mcg/L are seen in red blood cells. The half-life of fluoride is 2–9 hours.
	The reference range for unexposed urine is 0.2–1.9 mg/L. The occupational threshold limit is 4–7 mg/L
	Specimen collection containers for fluoride should be plastic since the silicon in glassware may react with fluoride and form volatile compounds that are lost during analysis.
Nickel	Nickel is measured by electrothermal atomic absorption spectrometry.[90] Care must be taken to avoid specimen contamination. Blood nickel levels are not clinically useful unless drawn within the first few minutes or hours of exposure.
	Normal whole blood and serum levels are 3–7 mcg/L and 1–5 mcg/L, respectively.
	Blood nickel is increased following myocardial infarction, cerebral vascular accidents, and burns and decreased in cirrhosis and uremia.
	Urine nickel levels (8-hour or 24-hour urine) are the best measure of severity of toxicity. They may also be used to monitor occupational exposure in workers and chelation therapy.
	Unexposed urine: 2–10 mcg/L
	Exposure threshold: >25 mcg/L
	Moderate exposure: 100–200 mcg/L
	Severe exposure: >500 mcg/L
	Nickel carbonyl is rapidly metabolized to ionic nickel, and urine nickel is the best indicator of toxicity.

Iron

Iron metabolism, deficiency, and overload syndromes are discussed at length in Chapter 41, Red Blood Cell Disorders.

Zinc

Zinc is an essential cofactor of many enzymes including carboxypeptidases, carbonic anhydrase, alkaline phosphatase, ribonucleic acid (RNA) and deoriboxynucleic acid (DNA) polymerase, and alcohol dehydrogenase. Zinc is found in vegetables, cereals, meat, and fish. The daily adult requirement is 12 to 115 mg. Zinc is absorbed in the duodenum and jejunum by a mechanism shared with other elements. Copper, calcium, phosphate, and the presence of various other dietary substances such as fiber and phytate inhibit zinc absorption. For this reason, zinc in vegetables and cereals that is complexed to phytate is not well absorbed. Zinc is transported in the blood bound to albumin and α_2-macroglobulin. The primary route of elimination is biliary. Normally, only small amounts are excreted in urine. This amount may be augmented with chelation therapy.

Clinical Syndromes

Causes or contributing factors of zinc deficiency include low dietary zinc, high intake of phytates, fiber, calcium or other substances that diminished zinc absorption, chronic malabsorptive syndromes, total parenteral nutrition, and penicillamine therapy (from chelation and excretion of zinc).

Zinc deficiency in children is characterized by anemia, growth retardation or dwarfism, and delayed sexual maturation. Affected adults exhibit anemia, hypogonadism, and poor wound healing. Zinc deficiency also appears to cause keratotic skin lesions, depression, lethargy, diarrhea, hypogeusia (impaired taste), and defective macrophage and T-cell functions.

Acrodermatitis enteropathica or Danbolt's disease is another syndrome of zinc deficiency that is caused by defective intestinal absorption. This autosomal-recessive condition is characterized by progressive bullous, pustular skin lesions, diarrhea, and alopecia.

Zinc toxicity occurs most often by inhalation of zinc oxide fumes (metal fume fever) or by ingestion of zinc salts. Zinc is used in metal alloys and paints and in the metal galvanizing process. Zinc chloride is found in wood-preserving products, astringents, moss killer, and fireproofing materials. Zinc supplementation usually is in the form of zinc gluconate.

Zinc oxide fumes are produced when zinc is heated to high temperatures during the welding or smelting process. Inhalation of these fumes causes a syndrome known as metal fume fever (49). Upon exposure, chest pain and cough occur followed in 4 to 8 hours by fever, chills, headache, nausea, vomiting, myalgias, and weakness. Effects usually last about 24 hours and resolve without treatment. In contrast, inhalation of zinc chloride fumes (chemical smoke) causes severe respiratory symptoms and may be fatal. Ingestion of zinc chloride causes caustic injury to the gastrointestinal tract with nausea, vomiting, epigastric pain, and hematemesis. Central nervous system (CNS) depression and renal dysfunction also may occur.

Side effects of zinc gluconate ingestion for treatment of the common cold are nausea, vomiting, and diarrhea. Long-term zinc may lead to copper deficiency, and excessive supplementation on a chronic basis may cause sideroblastic anemia (50).

Cobalt

The only known function of cobalt is as the metal component of vitamin B_{12}. Cobalt also can substitute for zinc in metalloenzymes, but the significance of this is not clear. The recommended daily allowance for vitamin B_{12} contains 0.13 mcg of cobalt. The average diet contains a surplus of cobalt (about 300 mcg). Cobalt is absorbed readily in the gastrointestinal tract. Renal excretion is the major route of elimination.

Clinical Syndromes

Inadequate cobalt causes vitamin B_{12} deficiency (See Chapter 35, HLA: Structure, Function, and Methodologies).

Cobalt is found in steel alloys of diamond-grinding disks, other industrial tools, and jet engines and in pigments in paints, china, and glass. Most commonly, occupational hazard is associated with the dusts or fumes that are produced during use of cobalt-containing tools. Direct contact with these dusts and fumes causes dermatitis, conjuctivitis, and rhinitis, and chronic inhalation results in pulmonary disorders such as allergic asthma,

interstitial pneumonitis and pulmonary fibrosis. Similarly, dusts and fumes from hard metal, a tungsten carbide-cobalt amalgam, are mucosal irritants.

Cobalt ingestion is associated with cardiotoxicity. Formerly, cobalt was used as a foam stabilizer commonly added to beer. Chronic excessive ingestion of cobalt-containing beer leads to polycythemia, goiter, and heart disease characterized by cardiomyopathy, pericardial effusion, and congestive heart failure (“beer drinker's cardiomyopathy”). The toxicity in these cases may have been exacerbated by concomitant nutritional deficiencies and ethanol. Accumulation of cobalt also occurs in uremic patients and may contribute to uremic cardiomyopathy (51).

Copper

Copper is an intrinsic part of a number of metal enzymes including cytochrome C oxidase, lysyl oxidase (connective tissue cross-linking), tyrosinase (melanin synthesis), dopamine-ß-hydroxylase (cathecholamine synthesis), and superoxide dismutase. The daily adult requirement for copper is about 2 mg. The copper content in food is dependent upon the copper content in the local soil. Shellfish and liver are rich sources of copper.

Copper is absorbed readily from the small intestine by a mechanism shared with zinc. Accordingly, there is a reciprocal relationship between zinc and copper absorption. Copper is stored in the liver and released into blood bound to ceruloplasmin. Most of the copper present in serum is bound tightly to ceruloplasmin. The major route of copper elimination is biliary (80%). Less than 5% is excreted in urine.

Clinical Syndromes

Copper deficiency is characterized by hypochromic anemia, neutropenia, bone and joint deformities, decreased skin pigmentation, and neurologic abnormalities including hypotonia and psychomotor retardation. The anemia is responsive only to copper supplementation and occurs because copper is required for iron incorporation into heme. Premature infants are at risk of developing copper deficiency as are patients with chronic diarrhea and malabsorption states and those on long-term total parenteral nutrition (52). Long-term zinc supplementation also may contribute to copper deficiency.

Menke's kinky hair syndrome (trichopoliodystrophy), a sex-linked, recessive disorder, is characterized by a functional copper deficiency. Its features include *pili torti* (kinky hair), retarded growth development, cerebral degeneration, arterial tortuosity, cardiac rupture, and bone and joint deformities. Serum copper and ceruloplasmin are low, and copper therapy does not cure this disease.

An accumulation of copper occurs in Wilson's disease (hepatolenticular degeneration), an autosomal-recessive disorder. Copper absorption is normal while copper excretion into bile is defective. As a result, excess copper deposits into various tissues including liver, brain, cornea, and kidneys. Clinical manifestations include Kayser-Fleischer corneal rings, hepatitis, cirrhosis, intellectual deterioration, behavioral and personality changes, tremor, dysarthria, and gait disturbance. In most cases, ceruloplasmin is decreased to less than 20 mg/dL. Liver copper is increased to greater than 250 mcg/g dry weight, and daily urinary copper excretion is greater than 100 mcg. Serum copper may be low or normal.

Copper is used widely in electrical wires, cooking utensils, and plumbing pipes. It also is found in common alloys such bronze (tin) and brass (zinc), industrial catalysts, furniture polishes, dyes, glass, and ceramics. Copper sulfate is used as an insecticide, herbicide, fungicide, and algicide commonly used in swimming pools (53). Copper smelters are at risk for inhalation of fine copper dust, which causes metal-fume fever similar to that observed with zinc oxide. Anorexia, nausea, and hepatomegaly also may occur. Copper salts are corrosive agents, and ingestion causes hemolytic anemia, nausea, vomiting, diarrhea, emesis, and abdominal pain. Gastrointestinal bleeding, shock, coma, and death also may occur. Hepatic and renal failure are common later sequelae.

Ingestion of excessive copper supplementation or contaminated water or beverages also may cause symptoms of gastroenteristis and hemolytic anemia. Copper leaching from pipes and containers occurs more with acidic and carbonated liquids. Chronic copper exposure leads to hepatic and CNS abnormalities resembling Wilson's disease. Copper toxicity from ingestion of milk and other liquids stored in brass containers is the likely cause of Indian childhood cirrhosis (ICC) (54).

Manganese

Manganese is an essential cofactor of numerous cellular enzymes including mitochondrial superoxide dismutase. Manganese is absorbed through the small intestine. Efficiency of absorption is low. Absorption and elimination appear to be regulated by the amount of dietary manganese. The major route of elimination is biliary with very little excreted in urine.

Clinical Syndromes

Abnormalities that have been associated with low manganese include impaired growth, skeletal deformities, decreased reproductive function, ataxia, and seizures (55, 56). However, these findings occurred in a small number of isolated cases, and a deficiency syndrome has yet to be well defined.

Manganese is used in various metal alloys, batteries, and ceramics. Acute toxicity in the form of metal fume fever (fever, chills, myalgia) may occur following inhalation of manganese fumes. Long-term inhalation may lead to chronic disability characterized by psychiatric and CNS finding (headache, irritability, personality change, hallucinations, memory loss, compulsive laughing or crying) or neurologic signs resembling Parkinson's disease (akinesia, rigidity, tremor). Liver failure may cause manganese neurotoxicity.

Molybdenum

Molybdenum is an essential component of xanthine oxidase, aldehyde oxidase, and sulfite oxidase.

Clinical Syndromes

There is no well-recognized molybdenum deficiency syndrome. Low molybedum may be associated with mental retardation, seizures, bilateral lens dislocation, sulfaturia, and thiosulfaturia.

Molybdenum is used in welding rods. Fume inhalation may cause anorexia, diarrhea, weakness, fatigue, headache, myalgia, and cough. Toxicity also may develop in areas with high molybdenum content in local soil. Hepatosplenomegaly, renal diseases, and inflammatory arthritis may occur (57).

Chromium

Chromium is required for glucose and lipid metabolism (58). Chromium is a crucial component of glucose tolerance factor (GTF), which potentiates insulin action. Meat and whole-grain foods are rich in chromium. Daily intake of chromium should be about 50 to 300 mcg. Chromium is absorbed in the small intestine. Urinary excretion is the major route of elimination.

Clinical Syndromes

The primary feature of chromium deficiency is hyperglycemia or glucose intolerance. Elevated free fatty acids, accerelated atherosclerosis, peripheral neuropathy, and encephalopathy also may occur in chronic chromium deficiency. Long-term parenteral nutrition is a known cause of deficiency. However, more extensive documentation of chromium deficiency cases has been hindered by the inability to make sensitive and specific chromium measurements. Chromium deficiency may exist frequently in adults with glucose intolerance or insulin resistance diabetes as well as in individuals with general malnutrition. Deficiency is suggested by the improvement of glucose tolerance following chromium supplementation.

Chromium is found in various metal alloys, pigments, dyes, and preservatives. Toxic chromium compounds are used in steel making, electroplating, photography processing, and leather tanning industries. Hexavalent chromium compounds (chromium trioxide and chromic acid) are the most toxic (59, 60). Ingestion of these corrosive agents may cause hemorrhagic gastroenteritis, hepatic and renal damage, coagulopathy, and hemolysis. Chronic inhalation of dusts and fumes may cause asthma, pneumoconiosis, and lung cancer. Contact leads to dermatitis, skin ulceration, and nasal septum perforation. Bivalent and trivalent forms such as chromic oxide and chromic sulfate are relatively insoluble and nontoxic.

Selenium

Selenium is an important antioxidant in the body. It is a component of the enzyme, glutathione peroxidase (61). This enzyme contains four atoms of selenium per molecule and functions to prevent oxidative damage to cells by quenching peroxides. Selenium also is a part of the enzyme, thyroxine-5′-deiodinase, and as such is involved in thyroid hormone metabolism.

Selenium is the least abundant and potentially the most toxic of all the essential elements (62). Seafood, meats, and cereals contain selenium, and daily intake is dependent upon the local soil content of selenium. The recommended daily allowance is 50 to 200 mcg. Selenium is absorbed from the gastrointestinal tract and is excreted readily into urine. The site and mechanism of its absorption are not well known.

Clinical Syndromes

Deficiency of selenium has been linked to cardiomyopathy and skeletal muscular degeneration.

Deficiency occurs in patients on long-term parenteral nutrition and endemically in areas with low selenium content in soil. Keschan disease is a form of dilated cardiomyopathy that affects mostly children in China and has been associated with selenium deficiency (63). Decreased selenium also has been linked to endemic osteoarthritis (Kashin-Beck disease) and skeletal muscle weakness and pain (64).

Selenium salts (sodium selenite, sodium selenate, iron selenide) are used in the electronic, semiconductor, steel, glass, ceramic, and pigment industries (65). Selenium is used also as a dietary supplement (elemental selenium) and as a component of antidandruff shampoo (selenium sulfide). Elemental selenium and selenium sulfide are relatively low in toxicity. Acute overdose of selenium or its salts may cause nausea, vomiting, abdominal pain, fatigue, and irritability. Chronic daily exposure to more than 3 mg of selenium may cause hyperreflexia, paresthesias, alopecia, white striae on the nails, metallic taste, and garlic odor on the breath. Exposure to selenium dusts produces respiratory-tract irritation, epistaxis, and cough.

Selenious acid (H_2SeO_3) is the most toxic form of selenium (66). It is used as a gun bluing agent. As little as 15 mL is potentially fatal. Ingestion causes gastrointestinal burns and erosion, hypersalivation, vomiting, diarrhea, muscle spasm, pulmonary edema, seizure, stupor, and respiratory depression.

Fluoride

Fluoride is absorbed in the small intestine and is incorporated into bone and teeth. Fluoride increases bone density and calcification and prevents tooth decay. Fluoride supplementation of public drinking water for the prevention of dental caries is limited to 1.5 mg per liter. Daily intake of fluoride is dependent upon consumption of fluoridated water and foods prepared with that water.

Clinical Syndromes

Fluoride compounds are widely used as insecticides and rodenticides. Hydrofluoric acid (aqueous hydrogen fluoride) is used to remove rust, to etch glass, and to manufacture semiconductor chips.

Fluoride in pharmacologic doses (15 to 60 mg per day) stimulates osteoblasts and increases trabecular bone mass and has been used to treat post-menopausal osteoporosis (67). Toothpaste usually contains no more than 1 mg fluoride per gram, and the maximum allowable amount of fluoride in a single tube is 260 mg. Approximately 0.05 mg/kg per day is optimal for prevention of tooth decay.

Acute ingestion of more than 5 to 10 mg/kg of fluoride may produce vomiting, abdominal pain, gastrointestinal ulceration and hemorrhage, hypocalcemia, hypomagnesemia, tetany, muscle weakness, respiratory depression, coma, arrhythmias, and death. Contact with hydrogen fluoride can cause severe corrosive injury to eyes, nose, mouth, and skin. Inhalation of the gas leads

to sore throat, cough, bronchospasm, chemical pneumonia, and pulmonary edema. Chronic fluoride overexposure produces discoloration of tooth enamel (mottling) and skeletal fluorosis (excess bone calcification, tendon and ligament calcification, and joint fusion.) Fluorosis occurs with chronic doses as low as 0.1 mg/kg/day for 10 to 20 years.

Nickel

The role of nickel as an essential micronutrient has not been well established. Daily dietary intake of nickel is about 0.3 to 0.6 mg. Nickel is absorbed poorly through the gastrointestinal tract. It is distributed mostly to the kidneys, liver, and lungs. Tissue accumulation is minimal. The primary route of elimination is renal.

Clinical Syndromes

Nickel is used in metal plating, electronic devices, and various metal alloys. Stainless-steel contains about 15% nickel. Nickel is emitted into the environment largely as a result of fuel oil combustion and waste and sewage treatments.

Nickel is a skin, mucous membrane, and respiratory tract irritant. Contact dermatitis, eczema, urticaria, and asthma occur in sensitized individuals. Most stainless-steel and costume jewelry allergies are likely because of nickel. Toxicity of metallic nickel and nickel alloys is low.

Nickel carbonyl is a very toxic form of nickel formed by contact of nickel with carbon monoxide (68). Inhalation exposure occurs in the rubber, electronic parts, battery, and electroplating industries. Immediate symptoms are headache, nausea, dyspnea, dizziness, and chest pain followed in 1 to 5 days by severe pulmonary symptoms (cyanosis, pneumonitis), arrhythmias, weakness, muscle pain, and convulsion.

Metallic nickel and many nickel compounds are probable or known human carcinogens. In particular, nickel sub-sulfide, which is found in refinery dust, causes lung and nasal carcinomas.

Iodine is discussed in Chapter 16, Endocrine Function and Carbohydrates.

NONESSENTIAL TRACE ELEMENTS

Part of "23 - Trace Elements, Vitamins, and Nutrition"

Sources and toxic effects of non-essential trace elements are discussed and presented in Table 23.7.

TABLE 23.7. TOXIC EFFECTS OF NONESSENTIAL TRACE ELEMENTS

Metal	Toxic Effects
Aluminium	Dialysis encephalopathy
	Pulmonary fibrosis following dust inhalation
Arsenic	Organic forms are not toxic.
	Arsine gas inhalation causes hemolysis, abdominal pain, and renal failure
	Acute inorganic arsenic ingestion causes hemorrhagic gastroenteritis, hypotension, shock, death, cold, clammy skin, muscle cramps, delirium, seizure, coma with delayed peripheral neuropathy and Aldrich-Mees' lines.
	Chronic toxicity may cause malaise, anorexia, anemia, gastroenteritis, hepatitis, skin lesions (hyperpigmentation, hyperkeratosis), Aldrich-Mees' lines, and peripheral neuropathy.
	Carcinogenic.
Cadmium	Inhalation produces a chemical pneumonitis.
	Chronic exposure to cadmium causes pulmonary fibrosis, renal failure, and osteomalacia.
Gold	Allergic dermatitis, stomatitis, alopecia, hypersensitivity pneumonitis, drug-induced lupus, bone marrow suppression, hepatitis, colitis, proteinuria, encephalopathy, and neuropathy
Lead	Acute toxicity is characterized by abdominal pain, nausea, vomiting, hepatotoxicity, hemolysis, renal failure, convulsions, and coma.
	Subacute or chronic exposure causes malaise, irritability, anorexia, anemia with basophilic stippling, constipation or diarrhea, other gastrointestinal complaints, renal toxicity, hypertension, and central nervous system manifestations including ataxia, incoordination, headache, diminished concentration, and acute encephalopathy with increased intracranial pressure.
	Chronic exposure in children is associated with low intelligence, decreased growth, diminished hearing, hyperactivity, and aggressive behavior.
	Permanent mental retardation may occur at high lead levels.
Mercury	Elemental mercury inhalation causes gingivitis, tremor, and neuropsychiatric changes (irritability, angry fits, mood lability, memory loss, depression, anxiety), corrosive bronchitis, and acrodynia.
	Inorganic mercury especially mercuric chloride causes hemorrhagic gastroenteritis with esophageal erosions, hematemesis, severe abdominal cramps and renal failure.
	Organic mercury especially methylmercury causes neurotoxicity including memory loss, personality changes, ataxia, paresthesias, motor weakness, dysarthria, hearing loss, and constricted visual fields. Quadraplegia, coma, and death may occur in severe to xicity. Also teratogenic.

Aluminum

Aluminum is the most abundant metal and the third most abundant element (after oxygen and silicon) in the earth's crust. It is found in all human tissues but has no demonstrated biological role. The average daily oral intake ranges from 10 to 100 mg. Food alone usually accounts for no more than 2 to 40 mg of aluminum per day. Use of aluminum food packaging, pans, and utensils can increase markedly aluminum intake. Individuals taking aluminum hydroxide antacids or phosphate-binding therapy may ingest 300 mg or more per day. Only 2% to 6% of ingested aluminum is absorbed (69). Renal excretion is the major route of elimination.

Patients with chronic renal failure and workers with aluminum in the workplace are at risk for developing aluminum toxicity. In renal failure, patients are not able to eliminate aluminum from aluminum-containing agents that are administered in order to lower their serum phosphate (70). Previously, those patients undergoing dialysis were at added risk because of aluminum in the dialysate. A progressive encephalopathy usually occurs after months or years of exposure to aluminum. Common manifestations include speech disturbance (stuttering or stammering), diminished attention and memory, disorientation, visual and auditory hallucinations, paranoia, and suicidal ideation. Microcytic anemia and neuromuscular abnormalities such as tremor, twitching, myoclonus, and immobility also may occur. Excessive amounts of aluminum also may cause osteomalacia and hypercalcemia because aluminum blocks the incorporation of calcium into osteoid. Occupational toxicity is most often a result of inhalation of aluminum dusts, which causes respiratory-tract irritation and leads to pulmonary fibrosis. Aluminum welders and smelters are at particular risk.

High levels of aluminum are found in the neurofibrillary tangles in the hippocampus and cerebral cortex in Alzheimer's disease. However, the causal relationship of aluminum and Alzheimer's disease has not been proven.

Measurement

Measurement of aluminum in body fluids is performed using atomic absorption spectrometry.

Because aluminum is ubiquitous, extreme caution is necessary to avoid specimen contamination in the laboratory.

Blood levels are used widely to monitor dialysis patients (71). Normal nontoxic blood levels are less than 15 mcg/L. Levels greater than 100 mcg/L are considered elevated and may indicate toxicity, but blood levels do not correlate with total body burden or severity of clinical toxicity.

Very little aluminum is excreted in urine and the urinary elimination half-life of aluminum is highly variable. Nevertheless, aluminum measurement in 24-hour urine may be used to monitor occupational exposures and chelation therapy. Hair measurement is not useful.

Arsenic

Arsenic is the second most common heavy metal (after lead) to cause significant poisoning. The toxicity of arsenic depends upon the form of arsenic (72). The most frequent source of arsenic poisoning is pesticides that contain inorganic arsenic compounds. **Organic arsenic** forms are found in seafood and include arsenobetain and arsenocholin. These forms are not toxic. Their half-lives are 4 to 6 hours, and they are completely cleared in urine in 1 to 2 days. **Metallic** arsenic also is nontoxic.

Arsine gas (AsH_3) is the most toxic form of arsenic. It is formed in metal smelting and refining and in the microelectronics industry when metals containing arsenic react with acid. Exposure to as little as 0.05 ppm is considered toxic, and 25 to 50 ppm may be fatal. Arsine gas is a colorless gas with a garlic-like odor. It is well absorbed by inhalation but is noncorrosive and does not cause respiratory symptoms. There is a latent period of up to 24 hours followed by abdominal pain and massive intravascular hemolysis. Secondary renal failure typically occurs in a couple of days.

Inorganic forms of arsenic are toxic. Arsenic trioxide (As_2O_3) is the most common commercial form of arsenic. It and other trivalent forms of arsenic (arsenite) react with free sulfhydryl groups and thereby inhibit important enzymes. Pentavalent arsenic (arseniate) as found in arsenic acid is somewhat less toxic than arsenite because of lesser solubility and is capable of substituting for phosphate and inhibiting mitochondrial oxidative phosphorylation.

Both trivalent and pentavalent forms of arsenic are found in pesticides, insecticides, herbicides, wood preservatives, and certain pigments used in glass manufacturing. Unintentional exposures occur as a result of gardening, boating, and ingestion of water or foods retaining pesticide residues. Smoking of cigarettes of tobacco sprayed with arsenic pesticides also may cause toxicity. Arsenic also may be used to cause intentional poisoning. Because of their high potency and ability to cause rapid death, arsenic compounds are used not infrequently for suicidal and homicidal purposes.

Contact with inorganic arsenic causes skin and mucous membrane irritation. Ingestion usually causes immediate symptoms of hemorrhagic gastroenteritis including nausea, vomiting, hematemesis, and rice water or bloody diarrhea. A garlic odor on the breath is characteristic. Hypotension, shock, and death may ensue. Other features include cold, clammy skin, delirium, encephalopathy, seizure, and lactic acidosis.

Delayed manifestations of acute arsenic poisoning may occur several weeks to months following exposure and include alopecia, hepatotoxicity, and neurologic abnormalities. The delayed peripheral neuropathy involves mostly the lower extremities and is characterized by distal paresthesias, ascending motor weakness, and paralysis. Aldrich-Mees lines, which are transverse striae on nails, occur at about 5 weeks after the exposure.

Chronic arsenic poisoning usually occurs from inhalation of arsenic compounds in the workplace or from ingestion of contaminated food or water (73, 74). Affected individuals may present with nonspecific complaints of fatigue, weakness, irritability, malaise, and anorexia. There may also be a garlic odor on breath, delirium, encephalopathy, seizures, Aldrich—Mees lines, neutropenia, anemia with basophilic stippling, gastroenteritis, hepatomegaly, chronic hepatitis, and cirrhosis. Characteristic skin lesions (pattern of hyperpigmentation and hypopigmentation with hyperkeratosis) typically develop only after years of exposure. Chronic arsenic also may cause a progressive peripheral neuropathy that is characterized by paresthesias or sensory loss with weakness, tremor, fasciculations, and loss of reflexes that occurs

first in the soles of the feet. Debilitating muscle weakness and atrophy may develop. Hoarseness, conjunctivitis, dermatitis, and perforation of the nasal septum may be present as a result of contact with arsenic dusts and fumes.

Arsenic also is a human **carcinogen**. A high incidence of lung cancer and hepatic angiosarcoma in the metal-smelting industry has been attributed to inhalation of dust from arsenic compounds such as arsenic trioxide and sulfide. Skin cancer has been associated with the chronic use of Fowler's solution (1% potassium arsenite) which was used in the first half of the 1900s as a topical agent to treat various skin disorders. Arsenic also may be linked to renal and bladder cancers.

Measurement

Arsenic is measured by atomic absorption spectrometry. An older colorimetric method utilizes silver diethyldithiocarbamate.

The short half-life of arsenic (about 2 to 6 hours) limits its detection period in blood. Blood levels are elevated for only 2 to 4 hours immediately following exposure (75) and are useful only in the first few hours of acute toxicity. Blood levels decline rapidly despite ongoing toxicity and cannot be used reliably in the assessment of chronic toxicity. Normal blood levels are less than 70 mcg/L. Fatal arsenic overdoses have been associated with blood levels ranging from 600 to 3,300 mcg/L (76). However, blood levels are highly variable and do not correlate well with the severity of toxicity.

A 24-hour urine is the most reliable specimen for assessment of arsenic poisoning. In normal, unexposed individuals, arsenic should be less than 50 mcg/L. Levels greater than 200 mcg/L indicate toxicity. Those between 50 and 200 mcg/L are considered borderline. Toxic forms of arsenic should cause elevated urine levels for about 6 to 20 days following exposure. Ingestion of seafood (nontoxic organic arsenic) also will elevate urine levels (200 to 1,700 mcg/L), and abstinence from seafood for 3 days prior to collection is recommended.

Hair measurement often is useful in the investigation of chronic arsenic poisoning because hair may afford a longer detection period. However, external contamination is an important consideration. About 30 hours are required after ingestion for arsenic deposition into hair to take place. Hair arsenic levels are normally less than 1 mcg/g. In chronic poisoning, these levels are elevated to 1-5 mcg/gram or more.

Most screening tests measure total arsenic. Speciation may be necessary to distinguish toxic inorganic arsenic compounds and nontoxic organoarsenicals. Typically, arsenic species are resolved by liquid chromatography or similar technique prior to their individual measurement by atomic absorption. It is important to remember when interpreting results that inorganic arsenic is converted in the body to organic metabolites, monomethylarsonic acid (MMA) and dimethylarsonic acid (DMA), consecutively. Following recent exposure to inorganic arsenic, there will be some MMA and DMA in the urine, and the amount of MMA exceeds that of DMA. In cases of chronic toxicity, DMA is greater than MMA. These differences can sometimes be used to date the exposure.

Cadmium

Cadmium has no physiologic role. Cadmium is poorly absorbed from the gastrointestinal tract. Cadmium has a very long biologic half-life of 10-30 years and is the heavy metal most likely to accumulate in the body. As such, the toxicity of cadmium is of significant public health concern.

Environmental exposures to cadmium in the general population occur through contaminated air, food and water. Cadmium occupational risk is high in nickel cadmium battery manufacturing, metal welding and smelting, and electroplating industries. Cadmium is also used as a pigment and stabilizer in paints and plastics.

Cadmium inhibits enzymes with critical sulfhydryl groups. It also interferes with zinc absorption and the function of zinc metalloenzymes. Cadmium accumulates in the kidneys and effects damage to the renal tubules. This nephrotoxicity is exacerbated in zinc deficiency.

Exposures to cadmium occur by ingestion or inhalation of fumes and dust. Acute inhalation is many times more toxic than ingestion. Inhalation produces a chemical pneumonitis characterized by cough, dyspnea, and chest pain. Pulmonary edema and death from respiratory failure may occur. Ingestion may cause abdominal cramping and diarrhea.

Chronic exposure to cadmium causes pulmonary fibrosis, renal failure, and osteomalacia. In the 1960s, a metals processing plant in Japan released cadmium into the local water which contaminated food products and subsequently caused severe cadmium poisoning called the Itai-Itai or “ouch ouch” disease. Prominent features were osteomalacia, skeletal deformities, and severe leg and back pain probably resulting from calcium and phosphorus disturbances secondary to renal damage (77).

Elevated cadmium has also been linked to hypertension (78). Furthermore, cadmium is a known carcinogen, and cadmium smelters have an increased incidence of lung cancer. Unlike lead and arsenic, cadmium is not a well-established cause of peripheral neuropathy. There is only one known reference that makes this association (79).

Measurements

Because of cadmium's high renal and pulmonary toxicity, exposed workers should be monitored periodically by cadmium measurements, urine β_2-microglobulin, and pulmonary function studies.

Electrothermal atomic absorption spectrometry for the measurement of cadmium in biologic samples has replaced an older dithizone colorimetric method. Special caution to avoid contamination is necessary because of the minute amounts of cadmium in normal blood and urine. Cadmium is highly bound to red cells, and the normal range for whole blood cadmium is 0.5 to 1.0 mcg/L. Cigarette smokers have higher levels, 1.4 to 4.0 mcg/L. Whole-blood measurements greater than 5 mcg/L are potentially toxic. Whole-blood cadmium levels will fall to normal months after last exposure while urine cadmium will remain elevated for longer periods.

For urine measurements, a 24-hour specimen is preferred. Urine cadmium is normally below 1 mcg/L. Values may be expressed

in micrograms per gram creatinine. Levels greater than 5 mcg/g creatinine are associated with irreversible renal damage. Urine β_2-microglobulin is elevated in cadmium-induced renal damage and levels correlate with duration and extent of cadmium exposure. Hair measurements also may be elevated in chronic toxicity. Normal levels are less than 1.0 mcg/gram.

Like other metals, cadmium in the body binds to the protein, metallothionein. Metallothionein appears to have a protective effect against cadmium toxicity. Elevated levels of urine metallothionein appear to correlate with cadmium accumulation. However, this test lacks specificity for cadmium since copper and zinc also will elevate levels.

Gold

Oral methotrexate therapy has largely replace gold therapy in the treatment of rheumatoid arthritis. Gold sodium thiomalate and gold sodium thioglucose both have been employed as a treatment for rheumatoid arthritis. These substances inhibit lysosomal enzymes and thereby reduce inflammation in the joints.

Gold has been associated with allergic dermatitis, stomatitis, alopecia, hypersensitivity pneumonitis, drug-induced lupus, bone-marrow suppression, hepatitis, colitis, proteinuria, encephalopathy, and neuropathy. Severe toxicity occurs in approximately 3% of cases. Because of these concerns, patients receiving gold therapy are monitored by platelet count, urinary protein, and serum aspartate aminotransferase. Gold toxicity does not appear to be dose-dependent. Furthermore, there is no correlation between serum or urine measurements of gold and toxicity or efficacy of therapy (80). Therefore, measurement of gold is almost never performed.

Lead

Lead is the most common heavy metal to cause poisoning and is of particular concern in children. Paints manufactured prior to 1977 contained significant amounts of lead, and buildings covered with these paints continue to be a major source of lead exposure. Toxicity occurs from lead ingested in the form of paint chips or inhaled as dust, debris, or fumes. At risk for lead toxicity are workers in certain industries and children who repeatedly ingest paint chips or lead-contaminated debris. Lead exposure also occurs from consuming plants grown in lead-contaminated soil and water from lead pipes, ceramics, and crystal containers. Other sources of lead are moonshine, certain herbal medications, and older gasoline products. In the 1990s, the Environmental Protection Agency oversaw the removal of tetraethyl lead from gasoline in the United States. Industries associated with significant lead exposure include lead battery manufacturing, ore smelting and refining, torching, soldering, building demolition and renovation, ammunitions and shipbuilding. In children, exposure occurs most commonly by ingestion of lead paint chips, dust, or soil. Industrial exposure generally is because of the inhalation of fumes or dusts.

Lead interferes with membrane integrity, neurotransmitter function, heme synthesis, and mitochondrial oxidative phosphorylation. In particular, lead inhibits the activity of the enzymes, aminolevulinic acid dehydrase, which converts aminolevulinic acid (ALA) to porphobilinogen, and coproporphyrinogen decarboxylase, which converts coproporphyrinogen to protoporphyrin. As a result, urinary ALA and coproporphyrin are elevated. Lead also interferes with incorporation of iron into the tetrapyrrole ring of hemoglobin.

Manifestations

Acute toxicity may be caused by pica, the ingestion of nonfood substances, and is characterized by abdominal pain, nausea, vomiting, hepatotoxicity, hemolysis, renal failure, convulsions, and coma. More commonly, toxicity is subacute or chronic in nature. Features include malaise, irritability, anorexia, anemia with basophilic stippling, constipation or diarrhea, other gastrointestinal complaints, renal toxicity, hypertension, and CNS manifestations including ataxia, incoordination, headache, diminished concentration, and acute encephalopathy with increased intracranial pressure. A peripheral neuropathy characterized by distal motor weakness (called lead palsy) occurs more often in adults than in children and does not appear to be dose-related. Hearing impairment, lead lines (blue lines in gingiva), and a metallic taste are also characteristic of lead toxicity.

The developing CNS is susceptible particularly to lead toxicity. Studies in children have shown that chronic low-level exposure to lead (whole blood levels over 10 mcg/dL) is associated with low intelligence, decreased growth and diminished hearing in children (81). Developmental delay, hyperactivity, and aggressive behavior also are linked to lead, and permanent mental retardation may occur at higher lead levels.

Measurements

Whole blood lead level is the best indicator of lead exposure. Plasma and serum are not appropriate. To avoid contamination, blood should be collected using lead-free syringe and tubes, and venipuncture technique is preferred over capillary blood sampling by fingerstick. Normal whole blood lead levels are less than 20 mcg/dL (1 mcmol/L) in adult urban populations and less than 10 mcg/dL (0.5 mcmol/L) in rural settings. In adults, levels over 40 mcg/dL(2 mcmol/L) are considered toxic, and overt intoxication (acute encephalopathy, lead colic, nephropathy) occurs at levels greater than 100 mcg/dL. All patients with blood level concentrations of 45 mcg/dL (2.2 mcmol/L) or higher should be chelated with cacium EDTA, dimercaprol (BAL), or dimercapto succinic acid (DMSA).

Toxic blood levels in children are significantly lower. During the 1980s, levels in excess of 25 mcg/dL (1.2 mcmol/L) were considered toxic. In 1991, the threshold level was revised downward to 10 mcg/dL (0.48 mcmol/L) (48). Current CDC recommendations for children are listed in Table 23.8.

TABLE 23.8. CENTERS FOR DISEASE CONTROL GUIDELINES FOR EVALUATION OF LEAD POISONING IN CHILDREN (1991)

Class	Blood Lead in mcg/dL	Concentration in mcmol/L	Comments
I	< 10	< 0.5	Not considered to be lead poisoning
II	10–14	0.5–0.7	Many children (or a large proportion of children) with blood lead levels in this range should trigger community-wide childhood lead poisoning prevention activities. Children in this range should be rescreened frequently.
III	15–19	0.7–0.97	Child should receive nutritional and educational interventions and more frequent screening. If the blood lead level persists in this range, environmental investigation and intervention should be done.
III	20–44	1–2.1	Child should receive environmental evaluation, remediation and a medical evaluation. Such a child may need pharmacologic treatment of lead poisoning.
IV	45–69	2.2–3.3	The child will need both medical and environmental intervention, including chelation therapy.
V	> 70	> 3.3	The child should be considered a medical emergency. Medical and environmental intervention must begin immediately.

From Olsen KR, ed. *Poisoning and drug overdose*, 2nd ed. Norwalk, CT: Appleton & Lange, 1994:210–213.

Urinary lead excretion is variable, poor and lacks correlation with CNS effects. Urine levels also will decrease more rapidly than blood levels following untreated exposures. Urine lead measurements are most useful for monitoring chelation therapy. A 24-hour urine collection is indicated and must be done using a lead-free container (e.g., a polyethylene bottle that has been washed with hydrochloric acid). An EDTA challenge test is the measurement of urinary lead excretion (8-hour or 24-hour) following a dose of EDTA (82). This test may be used in children

with levels between 25 and 44 mcg/L; those with high urinary lead levels would be additional candidates for chelation.

Lead measurement is taken using atomic absorption spectrometry or anodic stripping voltametry. Graphite tube flameless atomic absorption methods also have been developed. Measurements should be made without delay because over time, lead in standard solutions and samples will adsorb onto the walls of glass tubes.

In lead toxicity, elevations occur in free erythrocyte protoporphyrin, zinc protoporphyrin, urinary coproporphyrin, and urinary ALA levels as a result of impaired heme synthesis. None of these biochemical markers are specific for lead toxicity. Blood protoporphyrin levels become elevated several weeks following significant lead exposure and also are elevated in iron deficiency. Patients with inborn errors of metabolism also will have elevated urinary coproporphyrin and ALA excretion.

Mercury

Mercury has no known biological function but is a well-recognized toxin. Mercury is found in elemental, organic, and inorganic forms (83). Elemental mercury is used in thermometers, dental amalgams, paints, and chemical industries. It is a liquid that is vaporized readily. Following inhalation, mercury passes through alveoli and diffuses into tissues, notably the CNS, and causes a triad of gingivitis, tremor, and neuropsychiatric changes (irritability, angry fits, mood lability, memory loss, depression, anxiety). Acute inhalation also may cause a corrosive bronchitis. Chronic vapor exposure may cause acrodynia, an idiosyncratic reaction characterized by pain and pink discoloration of the extremities, sweating, anorexia, and insomnia. In contrast, ingestion of elemental mercury does not cause poisoning because gastrointestinal absorption is very poor. A common scenario for poisoning is ironing of clothes that are contaminated with liquid-mercury droplets. Heat from the iron causes mercury vaporization. The presence of multiple dental amalgams has been associated with elevated mercury levels presumably as a result of vaporizing small amounts of mercury during mastication. However, their potential to cause clinical toxicity is controversial.

Inorganic mercury occurs in mercurous (Hg^{+}) and mercuric (Hg^{++}) forms. Formerly, inorganic forms of mercury were used commonly in disinfectants ($HgCl_2$), calomel laxative (Hg_2Cl_2), and hat manufacturing ($HgNO_2$). Exposure more often is by ingestion than through inhalation or dermal exposure. Ingestion of inorganic mercury salts, especially mercuric chloride, can cause hemorrhagic gastroenteritis with esophageal erosions, hematemesis, and severe abdominal cramps. Shock and cardiovascular collapse may ensue. Nephrotic syndrome and renal failure also occur.

Serious mercury toxicity is most commonly caused by organic mercury. Organic mercury compounds are well absorbed gastrointestinally and dermally and are distributed widely in the body. Organic mercury exists in short-chain alkyl (methyl, ethyl), aryl (phenyl) and long-chain forms. Long-chain and aryl mercury compounds have a labile carbon-mercury bond and undergo conversion to inorganic mercury in the body. Short-chain organic mercury compounds are found in contaminated fish, fungicides, and antiseptic products. Bacteria and marine organisms convert inorganic mercury in pollutants to methylmercury, and ingestion of methylmercury-contaminated fish caused an epidemic in Japan during the 1950s known as Minamata disease. Short-chain organic compounds, especially methylmercury, cause neurotoxicity including memory loss, personality changes, ataxia, paresthesias, motor weakness, dysarthria, hearing loss, and constricted visual fields. Quadraplegia, coma, and death may occur in severe toxicity. Methylmercury also is teratogenic and causes mental retardation and cerebral palsy. In contrast to elemental and inorganic mercury, which are primarily renally

excreted, organic mercury is eliminated primarily in feces with less than 10% excreted in urine.

Measurement

Mercury analysis is performed by flameless atomic absorption spectrometry. Whole-blood mercury levels are used to confirm exposure to mercury. Levels less than 2 mcg/dL are considered normal. Levels greater than 35 mcg/dL are considered toxic. Short-chain mercury compounds have high affinity for red blood cells (RBC:plasma ratio of 5:1 to 20:1), which may be used to differentiate them from other mercury forms. Speciation into organic and nonorganic fractions also may be performed to determine the type of mercury exposure.

Urine levels are particularly useful in monitoring chelation therapy. Urine should be collected as a 24-hour collection in an acid-washed container. The urine level also will remain elevated for a longer period of time than blood following exposure. In inorganic mercury poisoning, urine levels of greater than 100 mcg/L are considered elevated. In methylmercury intoxication, urine is not a good index of total body burden because less than 10% is excreted in urine.

REFERENCES

1. Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: global burden of disease study. *Lancet* 1997;349:1436–1442.

1a. F. Xavier Pi-Sunyer. Obesity. In: Bennett JC, Plum F, eds. *Cecil Textbook of Medicine*, 20th ed. Philadelphia: W.B. Saunders Company, 1996;1161.

2. Bistrian BR, Blackburn GL, Hallowell E, et al. Protein status of general surgical patients. *JAMA* 1974;230:858–860.

3. Bistrian BR, Blackburn GL, Vitale J, et al. Prevalence of malnutrition in general medical patients. *JAMA* 1976;235:1567–1570.

4. Seltzer MH, Bastidas JA, Cooper DM, et al. Instant nutritional assessment. *JPEN J Parenter Enteral Nutr* 1979;3:157–159.

5. Reinhardt GF, Myscofski JW, Wilkens DB, et al. Incidence and mortality of hypoalbuminemic patients in hospitalized veterans. *JPEN J Parenter Enteral Nutr* 1998;4:357–359.

6. Bernstein LH, Shaw-Stiffel TA, Schorow M, et al. Financial implications of malnutrition. *Clin Lab Med* 1993;13:491–507.

7. Spiekerman, AM. Proteins used in nutritional assessment. *Clin Lab Med* 1993;13:353–369.

8. Gofferje H. Prealbumin and retinol-binding protein—highly sensitive parameters for the nutritional state in respect of protein. *Med Lab* 1978;5:38–44.

9. Bernstein LH, Leukhardt-Fairfield CJ, Pleban W, et al. Usefulness of data on albumin and prealbumin concentrations in determining effectiveness of nutritional support. *Clin Chem* 1989;35:271–274.

10. Forse RA, Shizgal HM. Serum albumin and nutritional status. *JPEN J Parenter Enteral Nutr* 1980;4:450–454.

11. Hendriks HF. Retinoid (vitamin A) metabolism. *Eur J Clin Nutr* 1996;50(suppl 3):S2–S6.

12. Ross AC. Overview of retinoid metabolism. *J Nutr* 1993;123(suppl 2):346–350.

13. Olson JA. Hypovitaminosis A: contemporary scientific issues. *J Nutr* 1994;124(suppl 8):1461S–1466S.

14. Bendich A. Safety of vitamin A. *Am J Clin Nutr* 1989;49:358–371.

15. Millikan LE. An update on common skin diseases. Acne, psoriasis, contact dermatitis, and warts. *Postgrad Med* 1992;91:96–98, 101–104, 107–110 passim.

16. McClean SW, Ruddel ME, Gross EG, et al. Liquid-chromatographic assay for retinol (vitamin A) and retinol analogs in therapeutic trials. *Clin Chem* 1982;28:693–696.

17. Matthews-Roth MM, Stampfer MJ. Some factors affecting determination of carotenoids in serum. *Clin Chem* 1984;30:459–461.

18. Audran M, Kumar R. The physiology and pathophysiology of vitamin D. *Mayo Clin Proc* 1985;60:851–866.

19. Doppelt SH. Vitamin D, rickets, and osteomalacia. *Orthop Clin North Am* 1984;15:671–686.

20. Haughton MA, Mason RS. Immunoephelometric assay of vitamin D-binding protein. *Clin Chem* 1992;38:1796–1801.

21. Hollis BW, Napoli JL. Improved radioimmunoassay for vitamin D and its use in assessing vitamin D status. *Clin Chem* 1985;31:1815–1819.

22. McLaughlin PJ, Weihrauch JL. Vitamin E content of foods. *J Am Diet Assoc* 1979;75:647–665.

23. Rock CL, Jacob RA, Bowen PE. Update on biological characteristics of the antioxidant micronutrients: vitamin C, vitamin E, and the carotenoids. *J Am Diet Assoc* 1997;96:693–702.

24. Tanyel MC, Mancano LD. Neurologic findings in vitamin E deficiency. *Am Fam Physician* 1997;55:197–201.

25. Thurnham DI, Davies JA, Crump BJ, et al. The use of different lipids to express serum tocopherol: lipid ratios for the measurement of vitamin E status. *Ann Clin Biochem* 1986;23(pt 5):514–520.

26. Lipsky JJ. Nutritional sources of vitamin K. *Mayo Clin Proc* 1994;69:462–466.

27. Conly JM, Stein K. The production of menaquinones (vitamin K2) by intestinal bacteria and their role in maintaining coagulation homeostasis. *Prog Food Nutr Sci* 1992;16:307–343.

28. Suttie JW. Vitamin K and human nutrition. *J Am Diet Assoc* 1992;92:585–590.

29. Widdershoven J, van Munster P, De Abreu R, et al. Four methods compared for measuring des-carboxy-prothrombin (PIVKA-II). *Clin Chem* 1987;33:2074–2078.

30. Lindberg MC, Oyler RA. Wernicke's encephalopathy. *Am Fam Physician* 1990;41:1205–1209.

31. Homewood J, Bond NW. Thiamin deficiency and Korsakoff's syndrome: failure to find memory impairments following nonalcoholic Wernicke's encephalopathy. *Alcohol* 1999;19:75–84.

32. Baines M, Davies G. The evaluation of erythrocyte thiamin diphosphate as an indicator of thiamin status in man, and its comparison with erythrocyte transketolase activity measurements. *Ann Clin Biochem* 1988;25(pt 6):698–705.

33. Rivilin RS. Riboflavin. *Adv Exp Med Biol* 1986;206:349–355.

34. Prasad AP, Lakshmi AV, Bamji MS. Interpretation of erythrocyte glutathione reductase activation test values for assessing riboflavin status. *Eur J Clin Nutr* 1992;46:753–758.

35. Carpenter KJ. The relationship of pellagra to corn and the low availability of niacin in cereals. *Experientia Suppl* 1983;44:197–222.

36. Jacob RA, Swendseid ME, McKee RW, et al. Biochemical markers for assessment of niacin status in young men: urinary and blood levels of niacin metabolites. *J Nutr* 1989;119:591–598.

37. Terry RC, Simon M. Determination of niacin metabolites 1-methyl-5-carboxylamide-2-pyridone and N-1-methylnicotinamide in urine by high-performance liquid chromatography. *J Chromatogr* 1982;232:261–274.

38. Bhagavan HN, Brin M. Drug-vitamin B6 interaction. *Curr Concepts Nutr* 1983;12:1–12.

39. Edwards P, Liu PK, Rose GA. A simple liquid-chromatographic method for measuring vitamin B6 compounds in plasma. *Clin Chem* 1989;35:241–245.

40. Liu M, Wang GR, Liu TZ, et al. Improved fluorometric quantification of urinary xanthurenic acid. *Clin Chem* 1996;42:397–401.

41. Vermaak WJ, Barnard HC, van Dalen EM, et al. Correlation between pyridoxal-5′-phosphate levels and the percentage activation of aspartate aminotransferase enzyme in haemolysate and plasma during in vitro incubation studies with different B6 vitamers. *Enzyme* 1986;35:215–224.

42. Baumgartner ER, Suormala T. Multiple carboxylase deficiency: inherited and acquired disorders of biotin metabolism. *Int J Vitam Nutr Res* 1997;67:377–384.

43. Bayer EA, Ben-Hur H, Wilchek M. A sensitive enzyme assay for biotin, avidin, and streptavidin. *Anal Biochem* 1986;154:367–370.

44. Williams EJ, Campbell AK. A homogeneous assay for biotin based on chemiluminescence energy transfer. *Anal Biochem* 1986;155:249–255.

45. Moody JR, Lindstrom RM. Selection and cleaning of plastic containers for storage of trace element samples. *Anal Chem* 1977;49:2264–2267.

46. Moyer TP, Mussmann GV, Nixon DE. Blood collection device for trace and ultra-trace metal specimens evaluated. *Clin Chem* 1991;37:709–714.

47. Delves HT. Assessment of trace element status. *Clin Endocrinol Metab* 1985;14:725–761.

48. Gibson RS. Assessment of trace element status in humans. *Prog Food Nutri Sci* 1989;13:67–111.

49. Sandstead H. Assessment of Zinc Nutriture. *J Lab Clin Med* 1991;118:299–300.

50. King JC. Assessment of Zinc Status. *J Nutr* 1990;120(suppl 11): 1474–1479.

51. Turnlund JR, Keen CL, Smith RG. Copper status and urinary and salivary copper in young men at three levels of dietary copper. *Am J Clin Nutr* 1990;51:658–664.

52. Baruthio F, Guillard O, Arnaud J, et al. Determination of manganese in biological materials by electrothermal absorption spectrometry: a review. *Clin Chem* 1988;34:227–234.

53. Chappuis P, Poupon J, Deschamps JF, et al. Physiologic chromium determination in serum by zeeman graphite furnace atomic absorption spectrometry. *Biol Trace Elem Res* 1992;32:85–91.

54. Hawker FH, Stewart PM, Snitch PJ. Effects of acute illness on selenium homeostasis. *Crit Care Med* 1990;18:442–446.

55. Ong Cn, Chua LH, Lee Bl, et al. Electrothermal atomic absorption spectrometric determination of cadmium and nickel in urine. *J Anal Toxicol* 1990;14;28–33.

56. Blanc PD. Metal fume fever. In: Olson KR, ed. *Poisoning and Drug Overdose*, 2nd ed. Norwalk, CT: Appleton & Lange, 1994;214–215.

57. Brown EF, Greist A, Tricot G. Excessive zinc ingestion. A reversible cause of sideroblastic anemia and bone marrow depression. *JAMA* 1990;164:1441–1443.

58. Lins LE, Pehrsson K. *Lancet* 1976;1:1191–1192.

59. Danks DM. Copper deficiency in humans. *Annu Rev Nutr* 1988;8:235–257.

60. Lamont DL, Duflou JA. Copper sulfate; not a harmless chemical. *Am J Forensic Med Pathol* 1988;9:226–227.

61. Muhlendahl K, Lange H. Copper and childhood cirrhosis. *Lancet* 1994;344:1515–1516.

62. Carl GF, Keen CL, Gallagher BB, et al. *Neurology* 1986;36:1584–1587.

63. Hall AJ, Margetts BM, Barker DJ, et al. *Paediatri Perinat Epidemiol* 1989;3:131–136.

64. Rajagopalan KV. Molybdenum: an essential trace element in human nutrition. *Ann Rev Nutr* 1988;8:401–427.

65. Anderson RA. Chromium metabolism and its role in disease processes in man. *Clin Physiol Biochem* 1986;4:31–41.

66. Ellenhorn MJ, ed. *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2nd ed. Baltimore: Williams and Wilkins, 1997;1551–1552.

67. Baruthio F. Toxic effects of chromium and its compounds. *Biol Trace Elem Res* 1992;32:145–153.

68. Shamberger RJ. Selenium metabolism and function. *Clin Physiol Biochem* 1986;4:42–49.

69. Lockitch G. Selenium: clinical significance and analytical concepts. *CRC Crit Rev Clin Lab Sci* 1989;27:483–541.

70. Locklick F, Taylor GP, Wong LTK, et al. Cardiomyopathy associated with nonendemic selenium deficiency in a Caucasian adolescent. *Am J Clin Nutr* 1990;52:572–577.

71. Levander OA. A global view of human selenium nutrition. *Ann Rev Nutr* 1987;7:227–250.

72. Ellenhorn MJ, ed. *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2nd ed. Baltimore: Williams and Wilkins, 1997;1607–1608.

73. Carter RF. Acute selenium poisoning. *Med J Aust* 1966;1:525–538.

74. Riggs BL, Hodgson SF, O'Fallon Wm, et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 1990;232:802–809.

75. Zhicheng S. Acute nickel carbonyl poisoning. *Br J Ind Med* 1986;43:422–434.

76. Fleming LW, Prescott A, Steward WK, et al. Bioavailability of aluminum. *Lancet* 1981;9:433.

77. Van De Vyver Fl, Silva FJE, D'Haese PC, et al. Aluminum toxicity in dialysis patients. *Contrib Nephrol* 1987;55:198–220.

78. Chazan JA, Lew NL, Lowrie EG. Increased serum aluminum. An independent risk factor for mortality in patients undergoing long-term hemodialysis. *Arch Intern Med* 1991;151:319–320.

79. Ellenhorn MJ, ed. *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2nd ed. Baltimore: Williams and Wilkins, 1997;1538–1543.

80. Hindmarsh JT, McCurdy RF. Clinical and environmental aspects of arsenic toxicity. *CRC Crit Rev Clin Lab Sci* 1986;23:315–347.

81. Lee-Feldstein A. A comparison of several measures of exposure to arsenic. Matched case-control study of copper smelter workers. *Am J Epidemiol* 1988;129:112–124.

82. Ellenhorn MJ, ed. *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2nd ed. Baltimore: Williams and Wilkins, 1997;1542.

83. Rehling CJ. Poison residues in human tissues. *Prog Chem Toxicol* 1967;3:363–386.

84. Tsuchiya K. Causation of ouch-ouch disease (itai-itai byo)—an introductory review. Keio. *J Med* 1969;18:195–211.

85. Thind GS, Fischer GM. Plasma cadmium and zinc in human hypertension. *Clin Sci Mol Med* 1976;51:483–486.

86. Viaene MK, Roels HA, Leenders J, et al. Cadmium: a possible etiological factor in peripheral polyneuropathy. *Neurotoxicology* 1999;20:7–16.

87. Heath MJ, Gillett GT, Swannell AJ, et al. Patterns of gold levels in urine, serum, and saliva in patients with rheumatoid arthritis undergoing therapy. *Ann Rheum Dis* 1987;46:827–828.

88. Marecek J, Shapiro IM, Burke A, et al. Low level exposure in childhood influences neuropsychologic performance. *Arch Environ Health* 1983;38;355–359.

89. *Preventing lead poisoning in young children*, Washington D.C.: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, 1991;3.

90. Markowitz ME, Rosen JF. Need for the lead mobilization test in children with lead poisoning. *J Pediatr* 1991;119:305–310.

91. Olson, KR. ed. *Poisoning & Drug Overdose*, 2nd ed. Norwalk, CT: Appleton & Lange, 1994;210–213.

24
Inborn Metabolic Errors

Tina M. Cowan

Miriam G. Blitzer

INTRODUCTION

Part of "24 - Inborn Metabolic Errors"

Human metabolism consists of a large number of interrelated and coordinated biochemical reactions, each of which is governed by the activity of one or more enzymes. Many of these reactions are involved in the degradation of proteins, lipids, and carbohydrates, which are derived either from the ingestion of food or from the normal turnover of endogenous compounds. An *inborn error of metabolism* is defined classically as a genetic defect in which reduced or absent enzyme activity leads to a specific block in a metabolic (usually catabolic) pathway. Such a block can lead to abnormal accumulation of substrate (and often its secondary metabolites) and/or a depletion of downstream intermediates, leading in turn to clinical pathology. Inborn errors typically are categorized by the metabolic pathway that is impaired, and include disorders of amino acid, organic acid, fatty acid, and carbohydrate metabolism. Clinically, most inborn errors of metabolism present with symptoms that could be compatible with more than one category of metabolic disease (e.g., inborn errors of amino acid and organic acid metabolism), or with more than one disease within a single category (e.g., the mucopolysaccharidoses Hunter and Hurler syndromes). A patient's presenting symptoms also can be suggestive of a nongenetic disease (e.g., sepsis, Reye syndrome). Therefore, metabolic disorders must be considered in the differential diagnosis of a wide variety of conditions, and their precise diagnosis is a prerequisite for optimal treatment and effective genetic counseling, including a discussion of the expected prognosis, recurrence risks, availability of heterozygote testing, and prenatal diagnosis.

In addition to the overlap in clinical presentation of many metabolic disorders, a large degree of biochemical heterogeneity exists. Although inborn errors most commonly arise from mutations in the genes encoding metabolic enzymes themselves, they also can be caused by abnormalities in enzymes involved in the synthesis and/or recycling of essential vitamins or cofactors, in proteins involved in the recognition or transport of compounds from one cellular compartment to another, or in proteins involved in the regulation of certain pathways. This underlying complexity can be the source of considerable diagnostic confusion, because a single biochemical abnormality can result from different underlying causes, each of which may be associated with a specific treatment, prognosis, laboratory diagnostic approach, and recurrence risk. Therefore, a basic understanding of the inborn metabolic errors, including their clinical presentation and approach to laboratory diagnosis, is essential for the optimum management of patients and their families.

Finally, recent advances in human molecular genetics are providing new insights into the contribution of molecular heterogeneity [i.e., heterogeneity at the level of the deoxyribonucleic (DNA) itself] to both clinical and biochemical heterogeneity. With increasing frequency, the genes involved in metabolic disorders are being isolated and characterized, and their pathogenic mutations described. This is leading to the increasing availability of DNA-based laboratory testing for some, but not all, of the inborn errors, both to confirm a biochemical diagnosis and to test other family members (including both carrier testing and prenatal diagnosis). In some cases, results of molecular testing are being used to give patients and their families a more accurate assessment of prognosis, and to define optimum management and treatment strategies. Refer to Chapter 7 and Chapter 8 (DNA) for a thorough discussion of the applications and limitations of molecular testing for human disease.

This chapter presents clinical and classic biochemical approaches for the diagnosis and management of patients with inherited metabolic diseases. It is not meant as an exhaustive review, but rather will focus on selected disorders as a means of highlighting clinical and laboratory features that help to distinguish between the many metabolic disorders. A complete listing of human genetic disorders can be found at the Online Mendelian Inheritance in Man (OMIM) website (http://www.ncbi.nlm.nih.gov/Omim/); this useful resource contains diseases cataloged by OMIM number, referenced here in the tables for readers wanting more detailed information. Finally, although specific laboratory tests and approaches will be discussed, detailed protocols will not be given. Refer to the Suggested Readings at the end of the chapter for descriptions of these methods.

METABOLIC DISORDERS

Part of "24 - Inborn Metabolic Errors"

Disorders of Amino Acid Metabolism

Normal Amino Acid Chemistry and Physiology

Amino acids, the fundamental building blocks of proteins, are small molecules with a tetrahedral carbon covalently bound to an amino group, a carboxyl group, a variable (R) group, and a hydrogen atom. R groups are classified as hydrophobic, neutral,

acidic, or basic, and determine the unique chemical properties of the individual amino acids. There are 20 commonly occurring amino acids found in proteins and at least 150 other nonprotein amino acids.

Normal turnover of amino acids occurs through a series of enzymatically controlled steps, the first of which often is a deamination, or removal of the primary amine group, to form the corresponding α-ketoacid. These compounds are metabolized in a step-wise fashion, ultimately entering the tricarboxylic acid (TCA) cycle leading to the formation of adenosine triphosphate (ATP). Several amino acids are intermediates in the urea cycle, the main pathway of ammonia detoxification through the formation of urea. Amino acids also are found as intermediates in the transsulfuration pathway, as well as pathways involved in the generation and utilization of single carbon units.

Free amino acids normally are found in all body fluids, including plasma, urine, cerebrospinal fluid, amniotic fluid, semen, sweat, and tears. Normal levels of amino acids depend to a great extent on the age of the individual, as well as nutritional status and overall health. Circulating plasma amino acids are the products of protein catabolism, and are filtered in the kidneys by the renal glomeruli. Specific amino acids then are reabsorbed in the proximal tubules via specialized transport systems, while others are efficiently excreted in the urine. The pattern of urinary amino acid excretion therefore depends largely on the function of transport mechanisms governing tubular reabsorption.

Alterations in Amino Acid Metabolism

Abnormal concentrations of amino acids in the plasma (aminoacidemias) or urine (aminoacidurias) arise when either the normal degradation or excretion of amino acids is altered1. These disorders are classified as either primary or secondary, depending on the nature of the particular defect.

Primary Disorders of Amino Acids

A primary aminoacidemia results from a mutation in a gene governing the activity of a specific degradative enzyme. The defective enzyme cannot catabolize a particular amino acid, which in turn accumulates. These blocks can also lead to abnormal accumulations of secondary metabolites and to depletions of “downstream” intermediates. Primary aminoacidemias refer solely to metabolic blocks in the first or second step of amino acid catabolism; blocks further downstream, although in the same overall pathway, are classified as organic acidemias and often differ in both clinical presentation and laboratory approach to diagnosis (see below). The vast majority of aminoacidemias follow an autosomal recessive pattern of inheritance.

Clinical findings associated with the primary aminoacidopathies range from severe and life threatening to benign. Examples of the most commonly encountered of these disorders are highlighted below; and a summary of common aminoacidemias is presented in Table 24.1.

Hyperphenylalaninemia Syndromes

Hyperphenylalaninemia is a heterogeneous group of disorders involving either a complete or partial block in the phenylalanine hydroxylase (PH) enzyme system, which converts phenylalanine to tyrosine. Classical phenylketonuria (PKU) involves a virtual absence of PH activity, while some variant forms demonstrate significantly reduced, but not absent, activity. Other variants result from a block in either the recycling or synthesis of tetrahydrobiopterin, the cofactor necessary for PH activity. Specific diagnosis is made on the basis of phenylalanine and/or cofactor levels in plasma; assay of the PH enzyme, a liver-specific protein, usually is not performed. An elevation of phenylalanine and its metabolites and/or a deficiency of tyrosine lead to early, irreversible brain damage and impairment of mature brain function. Clinical consequences include mental retardation, microcephaly, seizures, eczema, and a characteristic musty odor. These symptoms are avoided if dietary restriction of phenylalanine is started within the first weeks to months of life and continued throughout the patient's life. Benign forms of hyperphenylalaninemia also exist in which no treatment is necessary. Finally, maternal PKU is seen in children born to women with untreated hyperphenylalaninemia; prenatal exposure to high concentrations of phenylalanine leads to congenital malformations, microcephaly and mental retardation in patients who themselves have normal levels of phenylalanine.

Hyperphenylalaninemia occurs in approximately 1/12,000 live births, and all states in the United States offer newborn screening for this disorder to identify asymptomatic infants and to institute therapy, if needed, prior to the onset of symptoms. Accurate prenatal diagnosis of PKU and its variants now is available using DNA analysis and family studies in informative families.

Homocystinuria

Excretion of homocysteine in the urine can have a variety of inherited causes. The most common of these is deficient activity of cystathionine-β-synthase, the vitamin B_6- (pyridoxine-) dependent enzyme catalyzing the conversion of homocysteine to cystathionine. This block also leads to secondary hypermethioninemia. Some, but not all, of these patients respond to therapy with pyridoxine. Clinical features include dislocation of the optic lens, skeletal abnormalities including osteoporosis, increased tendency for thromboembolitic events, and variable mental retardation. Homocystinuria with methylmalonic aciduria and hypomethioninemia is caused by defects in the remethylation of homocysteine to methionine, a reaction dependent on the vitamin B_{12} derivative, methylcobalamin (MeCbl). Remethylation defects can arise from: (i) a failure in the metabolic conversion of hydroxycobalamin to MeCbl, (ii) nutritional or intestinal absorption deficiencies of vitamin B_{12}, or (iii) failure to synthesize 5-methyltetrahydrofolate, the cosubstrate in the remethylation reaction (note that methylmalonic aciduria is not seen in this last group). Clinical findings are highly variable, and range from severe, life-threatening illness presenting in infancy, to neuropathy and thrombocytopenia of juvenile or adult onset. Some, but not all, of these patients respond to therapy with vitamin B_{12} or MeCbl.

Nonketotic Hyperglycinemia

Nonketotic hyperglycinemia (NKHG) is a primary disorder of glycine metabolism clinically characterized by neonatal seizures,

TABLE 24.1. PRIMARY DISTURBANCES IN AMINO ACID METABOLISM

Disorder	Clinical Findings	Laboratory Findings	Deficient Enzyme	OMIM entry
Defects in sulfur-containing amino acids:				
Homocystinuria (due to CBS deficiency	Marfanoid features; optic lens dislocation; ±mental retardation, osteoporosis, thromboemboli. Variable responsiveness to pyridoxine (B_6).	↑ homocyst(e)ine in plasma and urine; ↑ methionine in plasma.	Cystationine-ß-synthase (CBS)	236200
Homocystinuria (due to MTR deficiency	Mental retardation, megaloblastic anemia. Autosomal dominant inheritance.	↑ homocyst(e)ine, ↓ methionine in plasma and urine.	N^5-methyl-tetrahydrofolate-homocysteine methyltransferase (methionine synthase)	156570
Homocystinuria (due to MTHFR deficiency)	Mental retardation, thrombosis, proximal muscle weakness. Variable responsiveness to folic acid.	↑ homocyst(e)ine in plasma and urine, normal plasma methionine.	N5,10-methylene-tetrahydrofolate reductase (MTHFR)	236250
Defects in phenylalanine metabolism:				
Phenylketonuria	If untreated, profound mental retardation, microcephaly, seizures, eczema, musty odor. Treatable by dietary restriction of phenylalanine.	↑ phenylalanine (20 mg/dl), ↓ or normal tyrosine in plasma, ↑ phenylpyruvic acid in urine.	Phenylalanine hydroxylase (in liver)	261600
Persistent hyperphenylalaninemia	Normal or near-normal neurological development, dietary restriction often not required.	↑ phenylalanine in plasma; milder than classical PKU.	Phenylalanine hydroxylase (greater residual activity than classical PKU)	261600
Dihydropteridine reductase deficiency	Mental retardation despite phenylalanine restriction, seizures, eczema, hypotonia, dystonic posturing. Treated with dopamine, 5-hydroxytryptophan and carbidopa.	↑ phenylalanine in plasma, abnormal biopterin/neopterin ratio in urine.	Dihydropteridine reductase	261630
Biopterin deficiency	Mental retardation despite phenylalanine restriction, seizures, hypotonia. Treated with dopamine, 5-hydroxytryptophan and carbidopa.	↑ phenylalanine in plasma, ↑ biopterin/neopterin ratio in urine.	Dihydrobiopterin synthetase	261640
Defects in tyrosine metabolism:				
Tyrosinemia Type I (hepatorenal form)	Hepatomegaly, jaundice, hepatic failure, vomiting, bleeding abnormalities, renal Fanconi syndrome, cabbagelike odor, liver cancer. Treated with NTBC or liver transplant.	↑ tyrosine in plasma, generalized aminoaciduria. ↑ succinylacetone in urine.	Fumarylaceto-acetase	276700
Tyrosinemia Type II	Progressive mental retardation, palmar and plantar hyperkeratosis and keratitis. Treated with diet restricted in phenylalanine and tyrosine.	↑ tyrosine in plasma, ↑ urine organic acids including 4-OH-phenylacetic, 4-OH-phenyllactic and N-acetyltyrosine.	Tyrosine transaminase	276600
Tyrosinemia Type III	Mild mental retardation, acute intermittent ataxia, no liver dysfunction.	↑ tyrosine in plasma.	4-hydroxyphenyl-pyruvate dioxygenase	276710
Other amino acid disorders:				
Nonketotic hyperglycinemia	Neonatal seizures (hypsarrhythmia), absent corpus callosum, death in infancy common.	↑ glycine in CSF, urine and plasma, ↑ ratio of CSF [gly] to plasma [gly].	Glycine cleavage system (in liver); four genetically distinct subunits	238300 238310 238330 238331
Hyperornithinemia (gyrate atrophy of choroid and retina)	Progressive chorioretinal degeneration, myopia, night-blindness, loss of peripheral vision in 1st decade, blindness by 3rd or 4th decade.	↑ ornithine in plasma, urine and CSF.	Ornithine aminotransferase	258870

neurological deterioration, and death in infancy or early childhood. Patients who survive have uncontrollable seizures and are severely mentally retarded. Diagnosis is based on abnormal elevations of glycine in all fluids, particularly cerebrospinal fluid (CSF). A ratio of CSF glycine/plasma glycine in the range of 0.2 or above is considered diagnostic (normal ratio less than or equal to 0.02). Because abnormal elevations of glycine also are seen in many of the organic acidurias, a complete workup must include urinary organic acid analysis. Ketones are characteristically absent in NKHG but often are present in the organic acidurias with secondary hyperglycinemia. There is no successful treatment for altering the outcome of NKHG, although sodium benzoate and dextromethorphan are sometimes used in attempts to lower glycine levels and improve neurological function. Confirmation of the diagnosis is available through enzyme assay of biopsied liver tissue, and prenatal diagnosis is performed following chorionic villus sampling, but not amniocentesis.

Urea Cycle Defects

The urea cycle is central to the excretion of nitrogen via conversion of ammonia to urea, and is also involved in the synthesis of arginine. The urea cycle defects (see Table 24.2) represent disruptions in the five different enzymatically controlled steps in the pathway. All are inherited as autosomal recessive disorders except ornithine transcarbamylase (OTC) deficiency, which is X-linked. Although considerable clinical and genetic heterogeneity exists, the typical and most severe presentation occurs at 1 to 3 days of life with progressive lethargy, hypothermia, and apnea accompanied by profound hyperammonemia (often in the range of 600 to 1,200 µg/dL). Later-onset forms can include episodes of vomiting, abnormal mental status, and hyperammonemia, as well as developmental delay and poor growth. Patients in this latter group often have enzyme activity that is reduced but not entirely absent. Laboratory studies leading to the diagnosis of a urea cycle defect include plasma and urine amino acids and urine organic acids, including orotic acid. The specific pattern of metabolites that are elevated indicates the point at which the pathway is impaired (see Table 24.2). Confirmation of the diagnosis can be performed by enzyme assay in either liver tissue (for arginino succinate synthetase, arginino succinic lyase, or arginase deficiencies) or in blood cells (for ornithine transcarbamylase (OTC) and carbamyl phosphate synthetase (CPS) deficiencies.) Ten percent to 15% of OTC mutations are large deletions involving all or part of the OTC gene; there also is considerable molecular heterogeneity among the other urea cycle disorders. Treatment of the urea cycle defects includes a protein-restricted diet, as well the administration of various agents (e.g., sodium benzoate, sodium phenylacetate, and arginine) aimed at enhancing the excretion of excess nitrogenous compounds.

TABLE 24.2. UREA CYCLE DEFECTS

Enzyme Defect/ Disorder	Clinical Findings	Laboratory Findings	OMIM entry
Carbamylphosphate synthetase I (CPSI) deficiency	Mental retardation, failure to thrive, seizures, vomiting, encephalopathy, coma. Neonatal lethal and late-onset forms.	Hyperammonemia. ↑ plasma glutamine, alanine, glycine and lysine, ↓ to normal citrulline and arginine. Generalized aminoaciduria. Deficient CPSI activity in liver.	237300
Ornithine transcarbamylase (OTC) deficiency	X-linked inheritance. Clinically indistinguishable from CPSI deficiency. Males typically die in infancy; variable expression in females.	Hyperammonemia. ↑ orotic acid in plasma and urine. ↑ plasma glutamine, glycine and alanine. Deficient OTC activity in liver.	311250
Argininosuccinate synthetase (AS) deficiency (citrullinemia)	Neonatal form: hypertonicity, vomiting, seizures, mental retardation, coma, death. Subacute form: mental retardation and ataxia in older children. Exacerbations with vomiting and encephalopathy.	Hyperammonemia. ↑ citrulline and glutamine in plasma and urine. ↓ plasma arginine. Deficient AS activity in liver, fibroblasts and lymphocytes.	215700
Argininosuccinic acid lyase (ASL) deficiency (arginino-succinic aciduria)	Failure to thrive, seizures, coma, mental retardation, early death. Milder forms in older children include ataxia, trichorrhexis nodosa and hepatomegaly. Exacerbations with vomiting and encephalopathy.	Hyperammonemia. ↑ argininosuccinic acid in plasma and urine. ↑ plasma glutamine, glycine, alanine and citrulline. Deficient ASL activity in erythrocytes and fibroblasts.	207900
Argininase (ARGI) deficiency (argininemia)	Spasticity, ataxia, seizures, vomiting, hepatomegaly, mental retardation. Exacerbations with vomiting and encephalopathy.	Moderate hyperammonemia. ↑ arginine in CSF, plasma and urine. Orotic aciduria during exacerbations. Deficient ARGI activity in erythrocytes, leukocytes and liver.	207800

Clinically Benign Aminoacidemias

In certain instances abnormal elevations of amino acids do not lead to any predictable pattern of clinical findings, and any symptoms that are present are coincidental to, but not caused by, the metabolic disturbance. Such conditions are thought to include histidinemia, sarcosinemia, prolinemia, and hydroxyprolinemia. Therefore, caution must be exercised in attributing specific clinical findings to abnormal elevations in amino acids.

Secondary Disorders of Amino Acids

Secondary changes in amino acid concentrations result from a physiologic disturbance not directly involving the primary

metabolic pathways themselves. These include transport and absorption defects, changes secondary to prematurity, kidney and/or liver failure, starvation, overfeeding, fever, organic acidemia, hyperammonemia, and interference from drugs and their metabolites. A partial list of secondary amino acid disturbances is given in Table 24.3. Some of these disorders have significant clinical consequences, such as failure to thrive in lysinuric protein intolerance or renal stones in cystinuria. Other secondary changes in amino acids can be indicators of other clinical problems, such as generalized aminoaciduria in renal tubular dysfunction and hypertyrosinemia in liver failure. A wide variety of medications can also lead to secondary changes in amino acid profile. Those most commonly encountered include antibiotics, antiseizure medications (e.g., valproate leading to hyperglycinuria and hyperglycinemia) and contrast dyes given for diagnostic imaging procedures. Because of the large number of situations that potentially can confound the interpretation of an altered amino acid profile, it is important that the laboratory be made aware of any pertinent clinical information, including relevant physical and routine laboratory findings, drug history, and diet and nutritional status.

TABLE 24.3. TRANSPORT DEFECTS LEADING TO SECONDARY AMINOACIDURIA

Disorder	Clinical Findings	Laboratory Findings	Primary Defect	OMIM entry
Cystinuria type I	Nephrolithiasis, renal complications, variable mental retardation in some patients.	↑ excretion of cystine, lysine, ornithine and arginine.	Mutations in SLC3A1 amino acid transporter gene; impaired transport of dibasic amino acids in renal tubules and intestinal mucosa.	220100
Lysinuric Protein Intolerance (dibasic aminoaciduria II)	Failure to thrive, periods of nausea and vomiting, aversion to protein-rich foods.	↑ excretion of lysine, arginine and ornithine; low levels in plasma. Intermittent hyperammonemia.	Mutations in SLC7A7 amino acid transporter gene; impaired transport of lysine, arginine and ornithine in intestinal epithelia and renal tubules.	222700
Hartnup disorder	Intermittent and variable photosensitivity, motor abnormalities, psychotic behavior. Some patients respond to nicotinamide; others show spontaneous remission.	Characteristic aminoaciduria of the monoamino-monocarboxylic group.	Impaired transport of neutral amino acids by intestinal mucosa and renal tubules.	234500
Lowe syndrome (oculocerebrorenal syndrome)	Renal tubular damage, congenital cataracts, hypotonia, hyporeflexia and psychomotor retardation. X-linked recessive inheritance.	Generalized aminoaciduria, proteinuria, phosphaturia, renal tubular acidosis.	Mutations in OCRL1 gene leading to deficiency of phosphatidyl-inositol 4,5-bisphosphate 5-phosphatase.	309000

Laboratory Investigations of Aminoacidemias

Amino acid analysis should be undertaken in any patient (i) who presents with life-threatening illness in the neonatal period, particularly if the illness coincides with the introduction of protein into the diet; (ii) with unexplained mental retardation or developmental delay, particularly if the patient was noted to be neurologically normal early in life and/or if there is a family history of a similarly affected sibling; or (iii) with unexplained systemic findings that could be compatible with one of the aminoacidopathies. Note that abnormalities in amino acid concentration often depend on protein intake, such that important clues may be missed if the patient was restricted from protein prior to sample collection. It is therefore imperative that specimens be collected immediately prior to altering the diet or giving medication.

The specimens evaluated most frequently are plasma, urine, and, less often, CSF. The analysis of amino acids in newborns and young infants is most efficiently and accurately done in plasma. Evaluation of newborn urine can be relatively unreliable and uninformative, because the sample often is dilute and may contain interfering drug metabolites (e.g., antibiotics given because of a clinical picture resembling sepsis). Furthermore, the immature renal reabsorption systems of premature infants can lead to secondary changes in urinary amino acid profile that can mask a primary abnormality. For older children and adults, analysis of urine is a relatively reliable screening approach, although valuable information is also gained by evaluating plasma. The most appropriate specimen(s) for analysis is frequently dictated by the clinical presentation of the patient.

Urine Spot Tests

A variety of simple urinary screening tests have been developed for patients suspected of having a metabolic defect. While these tests have the advantage of being rapid and relatively inexpensive to perform, they are limited in both sensitivity and specificity. However, especially when results of several tests are taken together, these “spot tests” can give useful clues that a more extensive metabolic workup may be warranted. The most commonly used spot tests for amino acid disorders include the ferric chloride test for PKU and tyrosinemia, the nitroprusside-cyanide test for cystinuria and homocystinuria, and the silver-nitroprusside test for homocystinuria.

Thin-Layer Chromatography

More precise information about amino acid composition is gained by chromatographic separation of amino acids by thin-layer chromatography (usually on cellulose or silica) in either one or two dimensions. This technique can be applied using either plasma or urine; urine amounts usually are standardized against creatinine concentration. Samples are subjected to acid precipitation prior to chromatography, such that only free amino acids are analyzed. Amino acids are visualized with a stain such as ninhydrin, which reacts with primary amine groups of all compounds, including those of amino acids (for this reason, many drugs also give rise to seemingly abnormal bands). The resulting pattern is interpreted by comparing the position, color and intensity of the bands with the pattern of normal, age-matched controls. This semi-quantitative approach to amino acid analysis is useful for detecting relatively large changes in amino acid concentration, such as those resulting from generalized aminoaciduria, hyperglycinuria, or phenylketonuria. Many of the more subtle changes may be missed by this method, and are much more accurately assessed by quantitative amino acid analysis using liquid column chromatography.

Quantitative Amino Acid Analysis

The most accurate method of assessing amino acid abnormalities is quantitative amino acid analysis. Modern systems are fully automated high-pressure liquid chromatography analyzers. In a typical setup, amino acids are separated on a single ion-exchange column using a gradient of elution buffers of increasing pH and ionic strength, reacted with ninhydrin, and then detected spectrophotometrically. The resulting peaks then are integrated, quantified, and compared to normal, age-matched controls. The entire analysis takes less than three hours per specimen, with the most sophisticated systems resolving at least 40 different amino acids. Because of the ease of operation and accuracy of results, some laboratories with automated systems routinely offer quantitative amino acid analysis in lieu of spot tests and thin layer chromatography.

Disorders of Organic Acid Metabolism

Normal Organic Acid Structure and Function

Organic acids are carboxylic acids that contain a variety of functional groups, but do not contain the primary amine group of amino acids. (This implies that organic acids will not be stained by ninhydrin, and will therefore not be detected by standard methods of amino acid chromatography.) Organic acids are metabolic intermediates from a variety of pathways including the metabolism of amino acids, fatty acids, carbohydrates, and steroids. Organic acids also are derived from the diet (e.g., benzoic acid, a common food preservative), a wide variety of drugs, as products of bacterial metabolism in the gut, or as contaminants from plastic urine containers, lotions, or other exogenous sources.

Disorders of Organic Acid Metabolism

Primary disorders of organic acid metabolism occur as a result of an enzymatic block in a metabolic pathway, leading to the abnormal accumulation of organic acid intermediates in the blood and urine. As with the majority of inborn errors of metabolism, the block can be from a mutation in a gene either encoding the degradative enzyme itself, or an enzyme involved in the synthesis or utilization of a cofactor required for enzyme activity. There currently are 50 to 60 known disorders of organic acid metabolism and their variants, most following an autosomal recessive mode of inheritance.

Organic Acidurias Arising from Blocks in Amino Acid Metabolism

The endpoint of amino acid metabolism is the generation of carbon skeletons and acetyl-CoA, which enter the TCA cycle for reutilization or energy production. Organic acidurias arising from blocks in amino acid metabolism most often involve the degradative pathways of valine, leucine, isoleucine (the branched-chain amino acids) or lysine; the specific pattern of accumulating intermediates gives important diagnostic information about which pathway is affected. Analysis of urinary organic acids is therefore an important step in pinpointing the specific metabolic block. For many, definitive diagnosis is confirmed by demonstrating deficient enzyme activity, most often in peripheral leukocytes or cultured skin fibroblasts. Table 24.4 outlines the more commonly encountered organic acidurias arising from amino acid metabolism.

Clinical findings tend to be overlapping and nonspecific. The typical course involves acute, life-threatening illness in the neonatal period, with altered mental status (e.g., lethargy, hypotonia, coma), tachypnea, and vomiting. Symptoms usually are not seen until after the first protein feeding. A less severe form includes failure to thrive and vomiting beginning during the first year of life, progressive psychomotor delay, and episodes of ketoacidosis and neurological deterioration exacerbated by infection, immunization, or other metabolic stress. Finally, some patients may experience onset of symptoms only after the first year of life, with episodes of ketoacidosis and neurological deterioration following a minor infection, often accompanied by lethargy, seizures, or coma. Organic acidurias also may be accompanied by unusual odors caused by the accumulation of organic compounds. Because of the nonspecific clinical findings, a definitive diagnosis only can be made through evaluation of urinary organic acid profile, most often with gas chromatography and mass spectrometry, but recently also with tandem mass spectrometry.

The accumulation of abnormal organic acid metabolites and their coenzyme A derivatives also leads to secondary laboratory findings giving strong indications that an underlying organic aciduria is present. These include metabolic acidosis (including ketosis and lactic acidosis), hypoglycemia, hyperglycinuria, hyperammonemia and carnitine deficiency. An abnormal accumulation of acylcarnitine derivatives is also often seen. Abnormal metabolites may only be elevated during acute illness, and levels can return to normal between episodes.

Maple Syrup Urine Disease

Maple Syrup Urine Disease (MSUD), or branched-chain ketoaciduria, is a heterogeneous group of disorders resulting from defects in branched-chain α-ketoacid dehydrogenase. This multienzyme complex is common to the catabolic pathways of the

branched-chain α-keto acids (BCKA) derived from valine, leucine, and isoleucine. The classic MSUD phenotype occurs in neonates or young infants, and includes apnea, poor feeding, lethargy, profound ketoacidosis, and coma possibly leading to death. The urine often has a sweet smell resembling maple syrup. Intermediate, intermittent, and thiamine-dependent forms of branched-chain ketoaciduria also exist with variable onset and outcomes; clinical heterogeneity may be a reflection of molecular heterogeneity. The abnormal accumulation of BCKA in MSUD leads to secondary elevations of branched-chain amino acids, which are readily detectable in plasma by amino acid analysis; a characteristic, abnormal pattern of urinary organic acid excretion also is seen. Alloisoleucine, a transamination product of the ketoacid of isoleucine, can be identified by amino acid analysis and is essentially pathognomonic for MSUD. Treatment of classical MSUD is accomplished with dietary restriction of the branched-chain amino acids. Confirmation of diagnosis by assay of enzyme activity can be performed.

TABLE 24.4. ORGANIC ACIDURIAS ARISING FROM DEFECTS IN AMINO ACID METABOLISM

Disorder	Clinical Findings	Laboratory Findings	Primary Enzyme Deficiency	OMIM entry
Disorders of leucine metabolism:				
Isovaleric acidemia	Neonatal presentation: poor feeding, vomiting, ketoacidosis, coma, "sweaty feet" odor. Chronic course: exacerbations with encephalopathy and mental retardation.	Severe metabolic acidosis, ↑ lactate and pyruvate. ↑ isovalerylglycine and 3-hydroxyisovaleric acid in urine; pancytopenia.	Isovaleryl-CoA dehydrogenase	243500
3-Hydroxy-3-methylglutaric acidemia	Hepatomegaly, lethargy, coma. Exacerbations with a Reye syndrome-like presentation.	Metabolic acidosis, nonketotic hypoglycemia, hyperammonemia, carnitine deficiency. ↑ 3-hydroxy-3-methylglutaric, 3-methyl-glutaconic, 3-hydroxy-isovaleric, 3-methylglutaric acids in urine.	3-hydroxy-3-methylglutaryl-CoA lyase (HMG CoA lyase)	246450
3-methylcrotonyl-glycinuria	Hypotonia, muscular atrophy, variable seizures and mental retardation. Some patients respond to treatment with biotin.	↑ 3-methylcrotonylglycine and 3-hydroxyisovaleric acid in urine.	3-methyl-crotonyl-CoA carboxylase	210200
Disorders of isoleucine and valine metabolism:				
ß-methyl-acetoaceticaciduria (3-oxothiolase deficiency)	Failure to thrive, encephalopathy, mental retardation.	Ketoacidosis, variable hyperammonemia. ↑ 2-methyl-3-hydroxyburic and methylaceto-acetic acids and tiglyglycine in urine.	Mitochondrial ß-ketothiolase	203750
Propionic acidemia	Acute presentation with lethargy, vomiting and seizures. Chronic presentation with episodic vomiting, failure to thrive and mental retardation. Some patients respond to treatment with biotin.	Ketoacidosis, hypoglycemia, hyperammonemia, neutropenia, ↑ glycine in plasma and urine. ↑ 3-hydroxypropionic and methylcitric acids in urine.	Propionyl-CoA carboxylase (two genetically distinct subunits)	232000 232050
Methylmalonic acidemia (*mut* group)	Acute presentation with lethargy, vomiting and seizures. Chronic presentation with episodic vomiting, failure to thrive and mental retardation. Some patients respond to treatment with B_{12}.	Ketoacidosis, hypoglycemia, hyperammonemia, neutropenia, thrombo-cytopenia, ↑ glycine in plasma and urine. ↑ methyl-malonic, methylcitric, and lactic acids in urine.	Methylmalonyl-CoA mutase	251000
Methylmalonic acidemia (*cblA* and *cblB* groups)	Variable presentation ranging from benign to severe. Vomiting, hypotonia, failure to thrive, seizures. Responsive to treatment with B_{12}	Intermittent ketoacidosis, hypoglycemia, hyper-ammonemia, neutropenia, thrombocytopenia, ↑ glycine in plasma and urine. ↑ methylmalonic, methylcitric, and lactic acids in urine.	Defective synthesis of adenosyl-cobalamin	251100 251110
Methylmalonic acidemia (*cblC* and *cblD* groups)	Severe CNS dysfunction including psychosis and mental retardation, pulmonary thromboemboli; skeletal, eye and cardiac abnormalities.	Megaloblastic anemia. ↑ methylmalonic acid in urine. ↑ homocysteine and cystathionine in urine and plasma, ↓ methionine in plasma.	Defective synthesis of methyl- and adenosyl-cobalamin	277400 277410
Multiple carboxylase deficiency	Seizures, skin rash, alopecia, developmental delay, hypotonia; abnormal urinary odor.	Lactic acidosis; ↑ urine ß-methylcrotonylglycine and 3-hydroxyisovaleric acids.	Holocarboxylase synthetase (early onset form); Biotinidase (later onset form)	253253 260270
Maple syrup urine disease (branched-chain ketoaciduria)	Four clinical variants, ranging from severe to mild. Mental retardation, cerebellar ataxia, CNS deterioration, seizures; maple syrup odor. Some patients respond to treatment with thiamine.	Acidosis, hypoglycemia, ketonuria. ↑ urine 2-oxo-isocaproic, 2-oxo-3-methyl-valeric and 2-oxoisovaleric acids. ↑ branched chain amino acids (valine, leucine, isoleucine) in blood and urine.	Branched-chain α-ketoacid dehydrogenase complex (four genetically distinct subunits)	248600 248610 248611 246900
Disorders of lysine and tryptophan metabolism:				
Ketoadipic aciduria	Neonatal collodion skin, retarded development, hypotonia.	Acidosis. ↑ urine 2-ketoadipic, 2-ketoglutaric and 2-aminoadipic acids.	Unknown; possibly deficient activity of 2-oxoadipic acid dehydrogenase	245130
Glutaric acidemia type I	Dystonia, choreoathetosis, macrocephaly, encephalopathy. Progressive dystonic palsy.	Acidosis, hypoglycemia. ↑ urine glutaric, 3-hydroxyglutaric and glutaconic acids. Carnitine deficiency.	Glutaryl-CoA dehydrogenase	231670

Isovaleric Acidemia

Isovaleric acidemia results from a deficiency in isovaleryl-CoA dehydrogenase, the enzyme that converts isovaleryl-CoA to 3-methylcrotonyl-CoA in the degradative pathway of leucine. Roughly half of patients have the neonatal form of this disorder, similar in presentation to neonatal MSUD, with acute episodes of severe metabolic acidosis and ketosis, vomiting and coma often leading to death. A chronic intermittent form also exists, with acidotic episodes often involving neutropenia, thrombocytopenia, or pancytopenia. An abnormal "sweaty feet" odor often is present during acute episodes. The major abnormal urinary organic acid detected by gas chromatography is the glycine conjugate of isovaleric acid, isovalerylglycine; diagnosis is confirmed by demonstration of absent or severely reduced activity of isovaleryl-CoA dehydrogenase in cultured fibroblasts. Normal growth and development can be seen in patients treated with moderate protein restriction, and carnitine and glycine supplementation.

Other Organic Acidurias from Altered Amino Acid Metabolism

A variety of other organic acidurias are clinically indistinguishable from MSUD or isovaleric acidemia and are differentiated solely by organic-acid profile. These include propionic aciduria, methylmalonic aciduria, and several others (see Table 24.4 and Table 24.5). Therefore, urinary organic acid analysis by gas chromatography/mass spectrometry and/or tandem mass spectrometry is an essential step in diagnosing the individual metabolic defects in patients presenting with symptoms compatible with one of these disorders.

TABLE 24.5. ORGANIC ACIDURIAS ARISING FROM DEFECTS IN FATTY ACID METABOLISM

Disorder	Clinical Findings	Laboratory Findings	Primary Enzyme Deficiency	OMIM entry
Short-chain acyl-CoA dehydrogenase (SCAD) deficiency	Neonatal presentation: acidosis, skeletal muscle weakness, failure to thrive, seizures; Chronic course: lipid-storage myopathy in adulthood	↑ urine ethylmalonic acid; low-normal carnitine with increased acylcarnitine fraction. No hypoglycemia as in MCAD or LCAD deficiency	Short-chain acyl-CoA dehydrogenase	201470
Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency	Lethargy, coma of variable age of onset, sometimes leading to sudden death; fatty changes in liver; exacerbations with metabolic stress (e.g., infection).	Hypoketotic hypoglycemia (intermittent), secondary carnitine deficiency. ↑ urine adipic, suberic, sebacic acids, hexanoylglycine and phenylpropionylglycine.	Medium-chain acyl-CoA dehydrogenase	201450
Long-chain acyl-CoA dehydrogenase (LCAD) deficiency	Hypotonia, hepatomegaly, cardiomegaly, cardiorespiratory arrest	Hypoketotic hypoglycemia, secondary carnitine deficiency. ↑ dicarboxylic and 3-hydroxydicarboxylic acids in urine.	Long-chain acyl-CoA dehydrogenase	201460
Long-chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) deficiency	Cardiomyopathy, skeletal myopathy, Reye-like picture, unexplained sudden death. Associated with acute fatty liver of pregnancy and HELLP syndrome in mothers of affected fetuses	Hypoketotic hypoglycemia, elevated liver enzymes, lactic aciduria. ↑ 3-hydroxy-dicarboxylic acids in urine. Single base pair substitution accounts for nearly 90% of subunit A mutations	Long-chain 3-hydroxyacyl-CoA de-hydrogenase (Mitochondrial trifunctional protein, subunit A or B)	600890 143450
Glutaric acidemia type I (multiple acyl-CoA dehydrogenase [MAD] deficiency)	Neonatal hypotonia, jaundice, coma, frequently death; "sweaty feet" odor, polycystic kidneys, hepatomegaly, facial dysmorphism	↑ glutaric, lactic, ethylmalonic, butyric, isobutyric, 2-methyl-butyric, and isovaleric acids. Generalized aminoaciduria	Electron transfer flavoprotein (two subunits) or electron transfer flavoprotein dehydrogenase	231680 130410 231675
Carnitine palmitoyl-transferase II (CPT2) deficiency	Neonatal form: cardiac arrhythmia and sudden death; adult form: exercise intolerance, myoglobinuria	Hypoketotic hypoglycemia, high plasma creatine kinase, ↓ plasma total and free carnitine, ↑ lipids and long-chain acylcarnitines	Carnitine palmitoyl-transferase II	600649

Organic Acidurias Arising from Blocks in Fatty Acid Metabolism

Fatty acid oxidation is an important source of energy in heart and skeletal muscle, and occurs in all tissues except brain. The endpoint of normal fatty acid ß-oxidation is the formation of acetyl-CoA and ketone bodies. Following the entry of fatty acids into the cell, these molecules are activated to their acyl-CoA esters, and subsequently linked to carnitine via the action of carnitine palmitoyl transferase (CPT) 1. Fatty acylcarnitines then are translocated across the inner mitochondrial membrane, and then reconverted to their fatty acyl-CoA derivatives by CPT 2 prior to entry into the ß-oxidation cycle. Sequential rounds of fatty-acyl CoA dehydrogenation are catalyzed by long-, medium- and short-chain acyl-CoA dehydrogenases (LCAD, MCAD and SCAD, respectively). Genetic defects in each of these enzymes have been described, as well as defects in their common cofactor, electron transfer flavoprotein (ETF), leading to a deficiency of all three enzymes (a condition called multiple acyl-CoA dehydrogenase deficiency, or glutaric aciduria type II). Defects in fatty acid oxidation also arise because of systemic carnitine deficiency caused by a variety of genetic and nongenetic factors. Patterns of abnormal fatty acid metabolites seen on urinary organic acid analysis give important clues about which enzyme(s) may be defective; definitive diagnosis through enzyme assay in leukocytes or cultured fibroblasts is possible in most cases. Genetic defects in fatty acid metabolism are outlined in Table 24.5.

Medium-Chain Acyl CoA Dehydrogenase (MCAD) Deficiency

Medium-chain acyl CoA dehydrogenase (MCAD) is an enzyme involved in the oxidation of medium-chain (C_6 to C_8) fatty

acids. Its deficiency leads to a disorder of variable clinical severity and age of onset, characterized by episodes of vomiting, lethargy, and profound hypoglycemia accompanied by low ketones, brought on by fasting or other forms of stress. The initial presentation typically occurs within the first 2 years of life and can include hepatomegaly and fatty infiltration of hepatocytes. Patients with MCAD deficiency have been misdiagnosed as having Reye syndrome or, because the initial presentation may be as extreme as sudden death, sudden infant death syndrome (SIDS). Recent studies have suggested that MCAD deficiency may be a significant contributor to cases of SIDS (Bennett et al. 1990). Accumulation of medium-chain fatty acids leads to elevated excretion of medium chain dicarboxylic acids (adipic, suberic, and sebacic acids) in urine, which can be detected by gas chromatography. More detailed studies, including urinary acylglycine and acylcarnitine profiles, also are available for the efficient workup of patients. A specific DNA mutation, designated G985, is seen in roughly 90% of white MCAD patients. Diagnosis is confirmed by demonstrating decreased MCAD activity in cultured fibroblasts and/or the presence of the G985 mutation by DNA analysis. Treatment involves maintaining adequate caloric intake, and chances for favorable outcome appear to be excellent. Although the exact incidence of this disease is unknown, the frequency of heterozygotes may be as high as 1/100.

Systemic L-Carnitine Deficiency

Primary systemic carnitine deficiency results from inherited defects in the active transport of carnitine in the kidney and small intestine, leading to increased urinary loss of carnitine. Secondary deficiencies result from the abnormal accumulation of acyl-CoA compounds, such as those accumulating in many of the organic acidurias involving amino acid or fatty acid metabolism. Acyl-CoA derivatives readily bind to free carnitine to form acylcarnitines, which are efficiently excreted in the urine. This leads to a secondary decrease in circulating free and total carnitine, and an increase in urinary acylcarnitines. Clinical abnormalities associated with carnitine deficiency itself resemble those of Reye syndrome, including hepatic dysfunction, encephalopathy, hypoglycemia, and hyperammonemia. There also may be signs of muscle dysfunction, including cardiomyopathy. There are no known defects in the synthesis or degradation of carnitine itself. Treatment of carnitine deficiency is by supplementation with L-carnitine, with good outcome in many cases.

Laboratory Investigations of Organic Acidurias

Organic acids are excreted effectively by the kidneys and concentrated in the urine. For this reason, organic acids are most accurately measured in urine rather than blood. Patients presenting with clinical indications of an organic aciduria, especially neonates or young infants, also should be evaluated for plasma amino acids, because several of the aminoacidemias can present with similar findings. In addition, plasma and urine carnitine levels should be determined. A complete metabolic workup also should include the following routine studies: blood lactate and pyruvate, complete blood count and differential, ammonia, glucose, electrolytes, ketones, and pH, and urine-reducing substances and ketones. Because metabolic abnormalities often are not detectable once the patient has been stabilized, it is imperative that all studies be ordered at the time of acute illness, and not after the more likely possibilities (e.g., sepsis, poisoning, starvation) have been ruled out.

Urine Spot Tests

As for the aminoacidemias, a variety of simple spot tests have been developed that can give indications that an organic aciduria may be present. Although the limitations in sensitivity and specificity of these tests must be recognized, results can provide important clues that a more extensive metabolic workup is warranted. The most commonly used spot tests for the organic acidurias are the dinitrophenylhydrazine (DNPH) test for ketoaciduria, Clinitest (Ames) for urine-reducing substances, ferric-chloride test for aromatic organic acids, and the methylmalonic acid spot test for methylmalonic aciduria.

Gas Chromatography and Mass Spectrometry

The general procedure for organic acid analysis involves the acidic extraction of urine organic acids, followed by derivitization (alteration of chemical structure such that compounds are suitable for chromatography) and separation by gas chromatography (GC). Creatinine is used as a reference for urine concentration. Extraction of organic acids typically is done using a combination of organic solvents such as ethyl acetate and/or ether. Oximization, or replacement of the keto groups of α-ketoacids with oxime groups, can be done prior to extraction to decrease artifacts introduced by the solvents. Following extraction, the organic acids are converted to their trimethylsilyl derivatives and separated by GC. The separated organic acids are identified by comparing their retention times with those of known, individual, organic acids. This semi-quantitative method is the one most often employed for routine screening for organic acid disorders.

Any suspected abnormalities detected by GC should be confirmed by mass spectrometry (GC/MS) to provide positive identification of organic acids. This analysis gives rise to a characteristic mass/charge ratio for each ionizable species. Computer programs compare the resulting mass spectra to libraries containing mass spectra of known compounds.

Further Studies of Organic Acidurias

Specialized methods recently have been developed for the detailed evaluation of patients suspected of having certain organic acid defects. Tandem mass spectrometry (MS-MS) is a highly sensitive technique used for the identification and quantitation of urinary acylcarnitines and other metabolites, such as those accumulating in patients with MCAD deficiency and other inborn errors of amino acid and organic acid metabolism. Alternatively, the stable-isotope dilution measurement of urinary glycine conjugates also identifies compounds accumulating in patients with a fatty-acid oxidation disorder, particularly n-hexanoylglycine, 3-phenylpropionylglycine, and suberylglycine in MCAD deficiency. Although once performed in only a small number of specialized laboratories, these rapid and sensitive tests are being used with increasing frequency in the workup of patients suspected of having a metabolic disorder.

Disorders of Carbohydrate Metabolism

Disorders of Galactose Metabolism

Classical Galactosemia

Classical galactosemia is an autosomal recessive disorder caused by deficient activity of galactose-1-phosphate uridyltransferase (transferase, or GALT) such that galactose-1-P (Gal-1-P) and UDP-glucose cannot be converted to UDP-galactose and glucose-1-P. The resulting clinical signs and symptoms usually appear within a few days following milk ingestion, and include failure to thrive, vomiting, and diarrhea and liver dysfunction presenting as jaundice and/or hepatomegaly. Cataracts resulting from punctate lesions on the lens nucleus can be present within a few days of life. Galactosuria, generalized aminoaciduria, proteinuria, and hyperchloremic acidosis are commonly present. There is an increased susceptibility to infection, particularly by *Escherichia coli*, and an increased frequency of death from *E. coli* sepsis is seen in affected infants. If left untreated, irreversible mental retardation will develop. At the DNA level, a single base pair substitution in the transferase gene, designated Q188R, has been shown to account for approximately 67% of classical galactosemia mutations. A simple polymerase chain reaction- (PCR) based assay can be performed to confirm this genotype; other mutations have been described in patients using direct DNA sequencing or other mutation scanning approaches.

Many of the clinical symptoms appear to be reversible and/or avoidable by early diagnosis and dietary restriction of galactose by avoidance of lactose ingestion. Liver abnormalities resolve, weight gain ensues, and the aminoaciduria, proteinuria, and galactosuria decrease. In general, the cataracts will regress, and if not extensive at the outset, do not result in impaired vision. Mental retardation is not severe, but is irreversible. Long-term studies have shown that intelligence quotient (IQ) scores often are somewhat lower in treated patients than in their unaffected siblings. In addition, there may be some specific learning disorders involving spatial relationships and mathematics, as well as an increased incidence of attention deficit and behavioral problems, even in patients under strict dietary control. A high incidence of ovarian failure and hypergonadotropic hypogonadism is seen in female patients, including those that have been well managed by dietary therapy; however, pregnancies have occurred in several patients. No abnormalities have been reported in male gonadal tissue.

Because galactosemia is a treatable disease that can be diagnosed prior to the onset of symptoms, most states screen for this disorder in newborn infants using dried blood specimens.

Measurements of galactose and/or gal-1-P levels by microbiological assays or transferase activity (by enzyme activity or bacterial inhibition assays) are employed routinely. The finding of nonglucose-reducing substances in urine by Clinitest strengthens the suspicion of galactosemia. [Caution must be stressed because many hospital laboratories test only for glucose, such as with glucose oxidase test (Clinistix), and therefore do not detect elevated galactose.] When abnormal screening results are found, diagnostic tests are performed for quantitative determinations of red-cell transferase activity and gal-1-P; absence of enzyme activity accompanied by elevated levels of gal-1-P, indicates the need for immediate implementation of dietary therapy.

Transferase Variants

The implementation of large-scale newborn screening for galactosemia led to the identification of infants with abnormally low, but not absent levels of transferase activity. The subsequent workup of these patients has resulted in the characterization of a number of variants of the transferase enzyme. These classifications are based on characteristic alterations in the electrophoretic migration pattern of the enzyme. The most common is the Duarte variant, which is identified by two distinct electrophoretic bands that migrate faster than the single band seen in normal individuals. Patients homozygous for the Duarte allele have roughly 50% of normal red cell transferase activity, and Duarte/normal heterozygotes have about 75% normal activity. Patients who carry both the Duarte and classical galactosemia alleles have approximately 25% normal activity, and can have elevated gal-1-P levels, which may warrant dietary galactose restriction for the first year or two of life. At the DNA level, a single nucleotide substitution, designated N314D, underlies the Duarte biochemical phenotype, and can be detected by a PCR-based assay. Other, less-frequently encountered transferase alleles also have been described.

Other Genetic Causes of Elevated Galactose

Galactokinase and UDP-gal-4-epimerase (epimerase) deficiencies are two other inherited enzyme defects in the galactose utilization pathway that can result in elevations of galactose or gal-1-P. Galactokinase deficiency results in the inability to convert galactose to gal-1-P, in turn leading to elevations of galactose and its byproduct, galactitol, in blood and urine. The clinical findings of galactokinase deficiency are cataracts and pseudotumor cerebri, both of which can occur very early in infancy. Patients can be identified by the newborn screening tests for classical galactosemia as having elevated galactose levels but normal transferase activity. Definitive diagnosis is based on demonstration of deficient galactokinase activity in red cells. This disorder occurs much less frequently than classical galactosemia: estimates from newborn screening suggest a frequency of about 1/250,000.

Patients with epimerase deficiency can be identified by newborn screening as having elevated red cell gal-1-P but normal transferase activity. Epimerase activity is deficient in red cells and leukocytes but is thought to be normal in other tissues. This condition typically is clinically benign, and appears to be more common among African Americans. Several patients with systemic, rather than isolated red cell, epimerase deficiency have presented with symptoms similar to classical galactosemia.

Glycogen Storage Disorders

Glycogen, the principal storage form of carbohydrate in animals, is present in virtually all cells in varying concentrations. It is composed of glucose molecules linked to form a polymeric, highly branched structure. Glycogen is synthesized from glucose via a series of enzymatic steps, and is degraded by a phosphorylase, causing the stepwise cleavage of glucose (yielding glucose-1-phosphate), with other enzymes cleaving the branched linkages.

The glycogen storage disorders (GSDs) are caused by defects in the enzymes involved in glycogen synthesis and degradation. Some present mainly with liver or muscle involvement, while others have a more generalized picture. The hepatic forms are most common and are characterized by growth retardation, hepatomegaly, and fasting hypoglycemia. Other forms primarily affecting muscle usually have no liver involvement, but may present with fatigue and muscle weakness. It is the combination of clinical, biochemical, and pathological findings that must be considered in the differential diagnosis of these diagnoses. Details of the GSDs are delineated in Table 24.6.

GSD Type I (von Gierke Disease)

This disorder is caused by deficient activity of glucose-6-phosphatase, a liver enzyme that normally converts glucose-6-phosphate to glucose. This results in severe hypoglycemia and lactic acidosis (with no ketosis) that can be life threatening. Neonates often present with hepatomegaly at birth, which continue to increase in size. There is a protruding abdomen from hepatomegaly (but no splenomegaly), short stature, and a characteristic doll-like facial appearance. There is no neurological involvement. Laboratory findings include hyperuricemia, hypertriglyceridemia, and bleeding tendencies from deficient platelet adhesiveness. Progressive renal dysfunction is a frequent complication in older patients. In the second decade, a significant incidence of adenomatous nodules in the liver develops, but cirrhosis does not occur. There are several clinical variants of GSD I, caused by different genetic mutations. Definitive diagnosis is made by assaying enzyme activity in liver, but the diagnosis usually is first inferred from the clinical findings, laboratory values, and functional tests. These include glucagon or epinephrine challenges where there is little or no rise in glucose but a large increase in lactate. Dietary treatment focuses on preventing hypoglycemia and controlling the lactic acidosis. Management strategies involve frequent feedings along with nocturnal nasogastric infusion of high glucose-containing formula.

GSD Type II (Pompe Disease)

GSD type II is a lysosomal storage disorder caused by deficient activity of the enzyme acid α-glucosidase (acid maltase). The infantile form presents during the first few months of life with muscle hypotonia, macroglossia, cardiomegaly, and hypertrophic muscles. The liver is normal in function but progressively increases in size. Hypoglycemia and ketosis are not present. Infants usually die within the first year of life from cardiac failure. The juvenile or adult forms typically present with delay and/or difficulty in walking and progress much more slowly than the infantile form. Muscle weakness can develop even in the third or fourth decade, appearing as chronic myopathy. GSD

TABLE 24.6. GLYCOGEN STORAGE DISORDERS (GSD)

Disorder	Clinical Findings	Laboratory Findings	Deficient Enzyme	OMIM entry
GSD Type I (Von Gierke disease) GSD types Ia, Ib, and Ic	Typically presents in infancy, hepatomegaly, short stature, life-threatening hypoglycemia and lactic acidosis; renal dysfunction in older patients. Three different genetic types with similar phenotypes.	Hypoglycemia, Hyperuricemia, lactic acidemia, hyperlipidemia, metabolic acidosis.	*Type IA*: Glucose-6-phosphatse *Type Ib*: glucose-6-phosphate translocase *Type Ic*: glucose-6-phosphate translocase	232200 232220 232240
GSD Type II (Pompe disease)	Classic form presents in infancy. Hypotonia, macroglossia, cardiomegaly, hepatomegaly; death in first year. Later onset forms occur.	No hypoglycemia or acidosis. Abnormal enzyme activity.	Acid α-glucosidase (acid maltase)	232300
GSD Type III (Forbes disease; Limit dextrinosis)	Hepatomegaly, similar to GSD I but milder. Several biochemical subtypes.	Mild hyperuricemia, lactic acidosis, hypoglycemia.	Glycogen debrancher enzyme; amylo-1,6-glucosidase	232400
GSD Type IV (Andersen disease)	Rare disorder. Failure to thrive, early development of cirrhosis with liver failure, death by age 5 years.	Hypoglycemia, abnormal enzyme activity.	Brancher enzyme; amylo(1,4 to 1,6) transglucosidase deficiency	232500
GSD Type V (McArdle disease)	Presents in early adulthood with muscle cramps following exercise. Significant genetic heterogeneity.	Myoglobinuria common, increased serum creatine kinase.	Muscle glycogen phosphorylase	232600
GSD Type VI (Hers disease)	Growth retardation, significant hepatomegaly, normal heart and skeletal muscle; good prognosis	Mild hypoglycemia and ketosis.	Liver glycogen phosphorylase	232700
GSD VII (Tarui disease)	Very rare. Muscle cramps and weakness with exertion.	Myoglobinuria with extreme exertion.	Muscle phospho-fructokinase	232800
GSD VIII	Mild disease. Hepatomegaly, growth retardation. X-linked inheritance.	Hypercholesterolemia, hypertriglyceridemia, fasting hyperketosis.	Liver phosphorylase kinase	306000

type II is diagnosed by measuring a-glucosidase activity in leukocytes or fibroblasts.

GSD Type III (Limit Dextrinosis or Forbes Disease)

GSD type III is caused by deficient activity of the debrancher enzyme (amylo-1,6-glucosidase) resulting in the accumulation of limit dextrin, a polysaccharide derived from the breakdown of glycogen that has short outer branches. The clinical picture is similar to, but milder than, GSD type I, with hepatomegaly as the predominant feature. The clinical and metabolic abnormalities usually decrease at puberty, while muscle and heart involvement may become more pronounced in adulthood. As with most genetic disorders, there are several biochemical subtypes of GSD type III. The disorder can be diagnosed by measuring activity of the debrancher enzyme in liver, leukocytes, or fibroblasts. Management is similar to, but not as demanding as, that for type GSD type I.

GSD Type IV (Andersen Disease)

GSD type IV is a rare disorder in which patients appear normal at birth, but present with hepatosplenomegaly, progressive failure to thrive, hypotonia, and muscle atrophy within the first several months of life. Cirrhosis develops and death occurs usually by the age of 5 years. GSD type IV is caused by deficient activity of the brancher enzyme [amylo(1,4 to 1,6) transglucosidase], resulting in the accumulation of structurally abnormal glycogen. The enzyme deficiency can be demonstrated in liver or fibroblasts.

GSD Type V (McArdle Disease)

GSD type V is caused by complete deficiency of muscle glycogen phosphorylase enzyme. Characteristically, patients present in their third or fourth decade with muscle cramps following vigorous exercise. Myoglobinuria commonly is seen. Serum creatine kinase can be elevated at rest and increases dramatically following exercise. There is significant clinical and molecular heterogeneity between patients that complicates the identification of these patients. Diagnosis can be confirmed by demonstrating deficient phosphorylase activity in a muscle biopsy.

GSD Type VI (Hers Disease)

In contrast to GSD type V, this disorder is caused by a deficiency in liver glycogen phosphorylase activity. The disease is characterized by significant hepatomegaly in childhood that resolves before puberty, as well as mild hypoglycemia, hypotonia, motor and growth delay. There is no heart involvement. Diagnosis is made by enzyme analysis in liver or leukocytes. Patients with GSD type VI have a primary defect in the liver phosphorylase gene.

GSD Type VII (Tarui Disease)

GSD type VII is the rarest of the GSDs and the most severe of the muscle glycogenoses. It is caused by a deficiency of muscle phosphofructokinase.

Clinically, GSD type VII presents in a similar fashion to but more severe than GSD type V, with onset usually in childhood. There is an increased reticulocyte count; red cells are affected, causing hemolytic anemia. Diagnosis is made by measuring phosphofructokinase activity in leukocytes or fibroblasts.

GSD Type VIII

This disorder is one of the most common of the GSDs and one of the mildest. Unlike all others, it is inherited in an X-linked recessive manner. Clinically, patients present with hepatomegaly, growth retardation, along with hypercholesterolemia, triglyceridemia, and fasting ketosis. Most adult patients are asymptomatic, as symptoms gradually disappear with age. Over the years, this disorder has been classified as GSD types VIa or IX. Patients have low liver phosphorylase activity, as seen in type VI, but in this case liver phosphorylase kinase activity also is deficient and is the primary defect.

Disorders Of Fructose Metabolism

Fructose is the predominant sugar in honey, vegetables, and fruit, and is a component of the disaccharide, sucrose, one of the most common sweeteners in food, medications, and infant formula. Fructose metabolism occurs mainly in the liver, kidney, and small intestine. The clinical consequences of metabolic blocks in fructose metabolism range from life-threatening to benign.

Hereditary Fructose Intolerance

Hereditary fructose intolerance (HFI) is an autosomal recessive disorder that presents with symptoms including failure to thrive, poor eating, vomiting, jaundice and hepatomegaly, hypoglycemia, generalized aminoaciduria, and proteinuria. Symptoms appear only following ingestion of fructose or sucrose and can be so severe as to be life threatening. Because there is no fructose in breast milk, breast-fed infants with HFI do not experience symptoms until weaning or the introduction of fruits and vegetables into the diet.

HFI is caused by deficient activity of fructose-1,6-bisphosphate aldolase (aldolase B), which normally catalyzes the conversion of fructose-1-phosphate to D-glyceraldehyde and di-hydroxyacetone phosphate. Its absence leads to an abnormal accumulation of fructose-1-phosphate in liver, kidney, and small intestine, resulting in hyperuricemia, lactic acidosis, and hypoglycemia. The diagnosis of HFI is suspected in patients with clinical findings described above and fructosuria following fructose ingestion. Although enzymatic studies can be performed using liver tissue, the diagnosis usually is made by demonstrating hypoglycemia and hypophosphatemia following an intravenous fructose load. The complete elimination of fructose and sucrose from the diet results in normal physical growth and development. Interestingly, some affected individuals learn to selectively avoid fructose-containing foods.

Fructose-1,6-bisphosphatase deficiency

Fructose-1,6-bisphosphatase catalyzes the conversion of fructose-1,6-bisphosphate to fructose-6-phosphate and inorganic phosphate. Because this enzyme plays an important role in gluconeogenesis, its deficiency leads to lactic acidemia once liver glycogen stores are depleted. Symptoms are similar to those seen for many of the organic acidurias, with severe metabolic acidosis, hypotonia, apnea, and hypoglycemia usually appearing within the first week of life. These patients do not develop an aversion to fructose, and there is no renal tubular dysfunction or liver abnormality, as seen in HFI. Outcome is good with appropriate dietary management.

Essential Fructosuria

This is a benign condition usually identified by elevated reducing substances following routine urinalysis. Essential fructosuria is caused by a deficiency in hepatic fructokinase activity, the enzyme that catalyzes the conversion of fructose to fructose-1-phosphate. No clinical or pathological findings have been reported.

MITOCHONDRIAL MYOPATHIES

Part of "24 - Inborn Metabolic Errors"

The mitochondrial myopathies are a diverse group of disorders involving abnormalities in the normal function of mitochondria. These important organelles contain metabolic pathways critical to energy metabolism and the production of ATP, and their deficiencies lead to a wide range of clinical abnormalities, predominately affecting muscle, brain, and heart tissue. Muscle biopsy often reveals “ragged red fibers,” or abnormal mitochondrial morphology visualized by special staining. Lactic acidosis also is commonly present.

Some of the mitochondrial enzymes are encoded by mitochondrial DNA, and their deficiencies follow a maternal pattern of inheritance. These disorders encompass a wide range of clinical variability stemming from a phenomenon called heteroplasmy, or variability in the proportion of mutant to normal mitochondria, both within and between tissues. Examples in this category include MERRF (mitochondrial encephalopathy and ragged red fibers) and MELAS (mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes), for which specific mitochondrial DNA mutations have been identified. Another group of mitochondrial disorders arises from mutations encoded by nuclear genes, with most inherited as autosomal recessive traits. This includes disorders of pyruvate metabolism (e.g., pyruvate dehydrogenase deficiency), as well as the tricarboxylic acid cycle (α-ketoglutaric dehydrogenase deficiency) and electron transport chain (complex I deficiency).

Screening tests for the mitochondrial myopathies include blood lactate and pyruvate, free and total carnitine, plasma amino acids, and urine organic acids. More definitive testing involves direct assay of specific enzymes from either skin or muscle biopsy, and/or molecular studies of mitochondrial DNA.

PEROXISOMAL DISORDERS

Part of "24 - Inborn Metabolic Errors"

Peroxisomes are organelles that perform a variety of anabolic and catabolic functions, such as the metabolism of very-long-chain fatty acids and the synthesis of plasmalogens. One group of peroxisomal disorders, including Zellweger syndrome, arises from

inherited defects in peroxisomal biogenesis leading to deficiencies of multiple peroxisomal enzymes. Although there is a spectrum of clinical severity, patients typically have profound hypotonia, dysmorphic features, hepatomegaly, and renal cysts, with death occurring in infancy. Laboratory findings supporting the diagnosis of a biogenesis defect include abnormal levels of very-long-chain fatty acids, pipecholic acid and phytanic acid, reflecting the deficiency of multiple peroxisomal enzymes. There is no known treatment for this group of disorders.

In addition to the biogenesis defects, there are an increasing number of disorders involving the deficiency of a single peroxisomal enzyme. Included in this category is X-linked adrenoleukodystrophy, a childhood-onset illness in which affected boys suffer a rapidly downhill course of progressive neurological deterioration and death. Plasma very-long-chain fatty acids are abnormally elevated, but other peroxisomal metabolites are normal. Results of experimental treatment protocols with bone marrow transplantation and dietary supplementation are cautiously optimistic.

Disorders of Lysosomal Storage

Lysosomes are single-membrane bound, acidic organelles that contain degradative hydrolases involved in the normal turnover of macromolecules. Lysosomal storage disorders result from deficient activity of a particular lysosomal enzyme or enzymes, leading to the abnormal accumulation of macromolecular substrates and ultimately to clinical pathology. Deficient enzyme activity can be caused by a mutation in a gene encoding the degradative enzyme itself, an activator protein required for proper enzyme function, or a protein involved in the synthesis and processing of the enzyme.

The general clinical course for the lysosomal storage disorders is one of regression. Patients most often appear normal at birth and begin to experience symptoms in late infancy or early childhood (although neonatal-, juvenile- and adult-onset variants of many of these disorders are known). Symptoms can include progressive loss of developmental milestones, psychomotor retardation and variable visceral involvement, usually leading to early death. The variation in symptoms between the different disorders is explained in part by the different tissue distributions of the various substrates. For example, many of the sphingolipids are components of brain and other nervous tissue; therefore, enzymatic defects in sphingolipid catabolism often lead to mental retardation and other neurological involvement. Significant clinical variability also is seen within the various groups of disorders, and the specific clinical phenotype is, to a large extent, determined by the molecular mutation(s) present in a given patient.

Mucopolysaccharidoses

The mucopolysaccharides are a heterogeneous group of macromolecules predominantly involved in the structural integrity of the extracellular matrix. They consist of unbranched polysaccharide chains containing both acidic and amino sugars and also are referred to as acidic mucopolysaccharides, glycosaminoglycans, or, when linked to protein, proteoglycans. There is varied tissue distribution for the different mucopolysaccharides: keratan sulfate is found predominantly in cartilage, cornea, and intervertebral discs; dermatan sulfate is found in heart, blood vessels, and skin; and heparan sulfate is a component of lung, arteries, and cell surfaces in general.

Mucopolysaccharidoses (MPS) are disorders resulting from defects in the stepwise degradation of mucopolysaccharides from enzymatic blocks at various points in the catabolic pathways of keratan, heparan, or dermatan sulfate. Although clinically heterogeneous, the MPS are generally characterized by a chronic progressive course, organomegaly (including hepatomegaly and splenomegaly), dysostosis multiplex, and coarsening of facial features. Profound mental retardation is seen in types IH, II, and III (Hurler, Hunter, and Sanfilippo syndromes), but normal intelligence is retained in types IS, IV, and VI (Scheie, Morquio, and Maroteaux-Lamy syndromes). Morquio syndrome also is characterized by distinctive skeletal abnormalities arising from abnormal storage of keratan sulfate. Treatment of the MPS disorders has been attempted by both enzyme replacement and bone marrow transplant, with mixed results. Outcome may be better for those disorders lacking neurological findings, and significant research using animal models and modern molecular approaches are currently underway. There are at least 10 different known enzyme deficiencies leading to abnormal mucopolysaccharide storage, including four genetically distinct deficiencies leading to Sanfilippo syndrome and two leading to Morquio syndrome. All of the MPS are inherited in an autosomal-recessive manner except Hunter syndrome, which is X-linked. A complete list of the MPS and their subtypes is given in Table 24.7.

TABLE 24.7. MUCOPOLYSACCHARIDOSES (MPS)

Disorder	Clinical Findings	Laboratory Findings	Deficient Enzyme	OMIM entry
Hurler syndrome (MPS IH)	Onset 6–8 months. Severe mental and motor regression, coarse facial features, early corneal clouding, cardiac insufficiency, stiff joints, protuberant abdomen, deafness, dysostosis	Urinary excretion of dermatan and heparan sulfate.	α-L-iduronidase	252800
Scheie syndrome (MPS IS)	Milder variant of type IH. Late corneal clouding, stiff joints, and cardiac insufficiency. Normal intelligence. Shortened life expectancy.	Same as type IH.	α-L-iduronidase	252800
Hunter syndrome (MPS II)	Onset in infancy and childhood. Mental and motor regression, dwarfing, coarse facial features, progressive deafness, stiff joints, hepatomegaly, no corneal clouding. Mlld to severe clinical forms. X-linked.	Urinary excretion of dermatan and heparan sulfate.	Iduronate sulfatase	309900
Sanfilippo syndrome (MPS III) Types A-D	Onset in first few years of life. Progressive mental retardation, severe behavioral disturbances, hirsutism, moderate dwarfing, joint stiffness, moderate organomegaly. Four genetically distinct forms, distinguishable by enzyme assay.	Variable urinary excretion of heparan sulfate.	*Type A*: Heparan sulfate sulfatase *Type B*: N-acetyl-alpha-D-glucosaminidase *Type C*: Acetyl CoA:α-glucosaminide-*N*-acetyltransferase *Type D*: N-acetylglucosamine 6-sulfatase	252900 252920 252930 252940
Morquio syndrome A (MPS IV A)	Onset 12–18 months. Severe progressive skeletal changes, corneal clouding, progressive deafness, odontoid hypoplasia, cervical myelopathy. Usually normal intelligence.	Variable urinary excretion of keratan sulfate.	Galactosamine-6-sulfatase	253000
Morquio syndrome B (MPS IV B)	Onset in childhood. Milder form of Morquio A. Bone changes, corneal clouding, hypoplastic odontoid.	Variable urinary excretion of keratan sulfate.	ß-galactosidase	253010
Maroteaux-Lamy syndrome (MPS VI)	Onset 2–3 years. Growth retardation, coarse facial features, corneal clouding, dysostosis multiplex, cardiac failure, normal intelligence. Mild to severe clinical forms.	Urinary excretion of dermatan sulfate.	arylsulfatase B (N-acetylgalactosamine-4-sulfatase)	253200
Sly syndrome (MPS VII)	Onset in infancy and childhood. coarse facies ± mental retardation, hepatosplenomegaly and dysostosis multiplex	Urinary excretion of dermatan and heparan sulfate.	ß-glucuronidase	253220

MPS patients characteristically excrete mucopolysaccharides in their urine; the specific pattern gives an indication of the site of the enzymatic defect. The preliminary workup of a patient suspected of having an MPS therefore includes a urinary screening test for abnormal mucopolysaccharides, including thin-layer chromatography or electrophoresis to identify the specific macromolecular species. Such screening tests are subject to false negatives, particularly for Sanfilippo syndrome; if clinical findings are suggestive of an MPS, further workup is warranted even with negative screening results. Histological changes include large, empty-appearing vacuoles, inclusions resembling zebra bodies, and metachromatic granules (seen upon staining with a cationic dye such as Alcian blue) that are lysosomes distended by stored material. Definitive diagnosis of a specific disorder is based on the demonstration of deficient enzyme activity in leukocytes or cultured fibroblasts.

Prenatal diagnosis is possible for all of the MPS by enzyme assay on chorionic villus tissue or cultured amniocytes. Heterozygotes have roughly half of normal enzyme activity, and so carrier testing for these disorders is feasible. Care must be taken when interpreting results of carrier testing for Hunter syndrome because of the influence of X chromosome inactivation on enzyme activity. Testing for mutations at the DNA levels can be performed provided that the specific mutation is known (i.e., has been identified in an affected family member).

Oligosaccharidoses

Oligosaccharides, or the carbohydrate portion of glycoproteins, consist of 10 to 20 sugar residues with a characteristic branched structure. These molecules play an integral role in determining

the specificity of protein structure, function and antigenicity, and are found throughout the body. The group of disorders known as the oligosaccharidoses arises from enzymatic defects in the stepwise degradation of N-linked oligosaccharides. This degradation normally occurs by the action of (i) a series of exoglycosidases acting in sequential fashion at the nonreducing termini, and (ii) aspartylglycosaminidase, which cleaves the bond between asparagine and the adjacent sugar, N-acetylglucosamine. Defects in each of the degradative steps are known, and are summarized in Table 24.8. All oligosaccharide storage disorders follow an autosomal-recessive pattern of inheritance.

TABLE 24.8. OLIGOSACCHARIDOSES

Disorder	Clinical Findings	Laboratory Findings	Deficient Enzyme	OMIM entry
Fucosidosis	Onset at 1 year. Coarse facies, growth retardation, dysostosis multiplex and severe psychomotor regression. Variants with milder presentations with angiokeratomas and psychomotor retardation have been described.	↑ sweat electrolytes in most patients. Urinary excretion of fucosyl-rich oligosaccharides.	α-fucosidase	230000
α-Mannosidosis	Onset during infancy. Rapid progression of psychomotor and mental retardation, facial coarsening, dysostosis multiplex, marked hepatosplenomegaly and death between 3–10 years. Milder clinical variants have been described.	Vacuolated lymphocytes. Urinary excretion of mannose-rich oligosaccharides.	α-mannosidase	248500
ß-Mannosidosis	Milder than α-mannosidosis. Variable mental retardation, facial coarsening, hepatosplenomegaly and bone disease.	Urinary excretion of a characteristic mannose-containing disaccharide.	ß-mannosidase	248510
Aspartylglycosaminuria	Onset in first few years of life. Angiokeratoma, severe mental retardation, cardiomyopathy, coarse facial feature, recurrent infections	Aspartylglycosaminuria.	Aspartylglucosaminidase	208400
Sialidosis (Mucolipidosis I) Types I and II	Type I: Onset in childhood. Myoclonus, ocular cherry-red spots, decreased visual acuity, gait abnormalities, normal intelligence. Type II: Dysmorphic findings, hepatosplenomegaly, dysostosis multiplex, coarse facies, and dwarfism. Clinical variants with congenital or juvenile onset have been described.	Urinary excretion of sialyloligosaccharides.	Neuraminidase I (sialidase)	256550

Clinically, the oligosaccharidoses are similar to the MPS, with a chronic progressive course, often following a period of normal growth and development. Specific findings can include mental retardation, facial coarsening, hepatosplenomegaly, cataracts or corneal opacities, and hearing loss. There are clinical subtypes for each of the oligosaccharidoses, representing a broad range of variability in both age of onset and severity of symptoms, and probably resulting from specific molecular mutations.

Abnormal urinary oligosaccharide excretion is seen in patients with an oligosaccharidosis, as well as in some patients with mucolipidosis, sphingolipidosis, or glycogen storage disease. The

diagnostic approach for patients suspected of having an oligosaccharidosis therefore begins with screening for abnormal urinary oligosaccharides by thin-layer chromatography; these results are used as a guideline for proceeding with specific diagnostic enzyme assays in leukocytes or cultured fibroblasts. Because of the clinical overlap between the MPS and oligosaccharidoses, the MPS should be considered in the differential diagnoses and workup of these patients. Prenatal diagnosis and carrier detection are possible for all of the oligosaccharidoses.

Sphingolipidoses

The sphingolipids comprise a heterogeneous group of macromolecules containing the fatty alcohol sphingosine, which is linked covalently to a fatty-acid side chain to form ceramide. Sphingolipids can be divided into different subclasses based on chemical modifications of ceramide: glycolipids (or glycosphingolipids) result from the linkage of ceramide to one or more sugar residues; cerebrosides from the linkage of ceramide to a single sugar residue; sulfatides (or sulfolipids) from a cerebroside containing a sulfate group; and sphingomyelin from ceramide linked to phosphocholine. In addition, the gangliosides are a complex group of glycolipids containing one or more negatively charged sialic acid (N-acetylneuraminic acid) residues. The sphingolipids are distributed throughout the body as structural components of plasma membranes, and many are particularly abundant in brain and other nervous tissue. Table 24.9 summarizes the known genetic defects of sphingolipid catabolism. All follow an autosomal recessive pattern of inheritance except Fabry disease, which is X-linked. Diagnosis is made directly from results of enzyme assays ordered on the basis of clinical, laboratory, and histological findings.

Although there is clinical overlap between the different disorders, it is useful to distinguish between those with a progressive neurodegenerative course and those characterized by more severe visceral involvement. The sphingolipidoses characterized by progressive psychomotor deterioration include metachromatic leukodystrophy, Krabbé disease (globoid-cell leukodystrophy), G_{M1} gangliosidosis, and G_{M2} gangliosidosis (including Tay-Sachs disease and Sandhoff disease). These disorders usually are

TABLE 24.9. SPHINGOLIPIDOSES

Disorder	Clinical Findings	Laboratory Findings	Deficient Enzyme	OMIM entry
G_{M1} Gangliosidosis Type I	Infantile onset. Rapid mental and motor regression. Seizures to decerebrate state, blindness, coarse facies, macrocephaly, cherry-red spots, macroglossia, hepatosplenomegaly, dysostosis multiplex; death 1–2 years.	Foamy histiocytes in bone marrow.	Acidic ß-galactosidase (ß-galactosidase-1)	230500
Type II	Juvenile onset. Slower progression than type I. May present with ataxia, seizures, strabismus, loss of speech; death 3–10 years.	Same as in type I.	Acidic ß-galactosidase (ß-galactosidase-1)	230600
Type III	Adult onset. Progressive cerebellar dysarthria with spasticity and ataxia, angiokeratoma.	Same as in type I.	Acidic ß-galactosidase (ß-galactosidase-1)	230650
G_{M2} Gangliosidosis				
Tay-Sachs disease	Onset 3–6 months. Exaggerated startle reaction. Seizures, mental and motor deterioration, blindness, doll-like facies, cherry-red spots, macrocephaly; death by 3–4 years.	Deficient activity of hexosaminidase A in serum, leukocytes and fibroblasts.	Hexosaminidase A	272800
Sandhoff disease	Clinically indistinguishable from Tay-Sachs.	Deficient activity of hexosaminidase A and B in leukocytes and fibroblasts.	Hexosaminidase A+B	268800
Juvenile G_{M2} gangliosidosis	Onset 2–6 years. Ataxia, progressive spasticity, seizures, late blindness; death by 5–15 years	Same as in Tay-Sachs disease.	Hexosaminidase A	272800
Adult (chronic) G_{M2} gangliosidosis	Progressive deterioration of gait and postures, muscle weakness and wasting, and ataxia. Gradual loss of motor and intellectual milestones.	Same as in Tay-Sachs disease.	Hexosaminidase A	272800
Metachromatic leukodystrophy				
Late infantile form	Onset 6–24 months. Rapid neurological regression, gait disturbance, blindness, peripheral neuropathy and quadriparesis.	Elevated CSF protein, decreased nerve conduction velocity.	Arylsulfatase A (cerebroside sulfatase)	250100
Juvenile form	Onset 5–10 years. Mental regression, gait disturbance, blindness, peripheral neuropathy loss of speech, quadriparesis.	Same as above.	Arylsulfatase A (cerebroside sulfatase)	250100
Adult form	Onset mid-teens-60 years. Dementia, symptoms similar to juvenile form.	Same as above.	Arylsulfatase A (cerebroside sulfatase)	250100
Krabbe disease (globoid cell leukodystrophy)	Onset 3–6 months. Irritability progressing to severe mental and motor deterioration. Infantile encephalopathy, hypertonicity, blindness, cherry-red spots, deafness, peripheral neuropathy; death by 12–18 months.	Elevated CSF protein, decreased nerve conduction velocity.	Galactocerebroside ß-galactosidase	245200
Fabry disease	X-linked inheritance. Pain and paresthesia in extremities, whorl-like corneal dystrophy, angiokeratoma, cardiovascular angiectasis and renal failure.	Proteinuria, increased globotriaosylceramide in plasma or urinary sediment.	α-Galactosidase A	301500
Gaucher disease				
Type I (chronic nonneuropathic; adult form; visceral form)	Onset in childhood or later. Most patients are Ashkenazi Jews (incidence 1:2,500). Hepatosplenomegaly, episodic bone pain, hypersplenism and pulmonary infiltrates.	Typical “Gaucher” foam cells in bone marrow, thrombocytopenia, anemia.	Acidic ß-glucosidase (glucocerebrosidase)	230800
Type II (acute neuronopathic)	Onset birth-18 months. Rapid progressive neurological regression with seizures, hypertonicity, spasticity and hepatosplenomegaly; death by 12 months.	Same as above.	Acidic ß-glucosidase (glucocerebrosidase)	230900
Type III (subacute neuronopathic)	Similar to type II. Symptoms may appear early in life, but their progression is slower.	Same as above.	Acidic ß-glucosidase (glucocerebrosidase)	231000
Niemann-Pick disease				
Type A (acute neuronopathic form; classic infantile)	More common in Ashkenazi Jews. Onset in infancy. Rapidly progressive neurological regression, cherry-red spots, hepatosplenomegaly; death by 3 years.	Foam cells in bone marrows.	Sphingomyelinase	257200
Type B (chronic, visceral form)	More common in Ashkenazi Jews. Onset in infancy or childhood. Only the visceral findings of type A.	Same as above.	Sphingomyelinase	257200
Type C (chronic neuronopathic form)	Onset usually after 2 years. Spasticity, seizures hepatosplenomegaly, loss of mental and motor skills; death before age 20 years. Clinically heterogeneous.	Foamy macrophages in marrow; low HDL, LDL and ApoA1.	Mutant NPC protein leading to block in cholesterol esterification. Sphingomyelinase not deficient.	257220
Type D (Nova Scotia variant)	Onset 2–4 years. Resembles type C with a protracted course and neurological abnormalities.	Foam cells in spleen and lymph nodes.	Mutation allelic to Type C	257250
Type E (Adult form, nonneuropathic form)	Adult onset. Hepatosplenomegaly. No neurologic difficulties.	Foam cells in marrow.	Sphingomyelinase.	257200

CSF, cerebrospinal fluid; HDL, high-density lipoprotein; LDL low-density lipoprotein.

considered in the evaluation of patients with clinical findings suggestive of neurodegenerative disease.

The sphingolipidoses characterized by more severe visceral involvement include Gaucher disease and Niemann-Pick disease. Both of these diseases are characterized by hepatosplenomegaly and variable neurological involvement. Histopathological changes represent lysosomes swollen with stored material: Gaucher disease is typified by the presence of Gaucher cells, and Niemann-Pick by foamy histiocytes. More rarely encountered disorders such as Farber disease and acid lipase deficiency are discussed elsewhere (see Suggested Readings); Pompe disease (glycogen storage disease type II) is discussed elsewhere in this chapter.

Tay-Sachs Disease (TSD) Carrier Testing

Treatment strategies for the sphingolipidoses are aimed at prevention rather than cure. This involves the identification of at-risk couples (i.e., both partners are carriers for the same disorder), who can then make informed decisions regarding their reproductive future. This approach is well suited for TSD, which is found at a significantly increased frequency among Ashkenazi Jewish individuals, as well as a subset of French Canadians. There is roughly a ten-fold increase in heterozygote frequency among these groups compared to the general population, making it feasible to conduct screening programs aimed at identifying TSD carriers.

The assay for TSD carrier detection is routinely performed using serum, although certain medications (particularly oral contraceptives) and medical conditions (pregnancy, liver disease) can interfere with the serum assay. In these cases, evaluation of enzyme activity in leukocytes usually yields reliable results. Carriers of Sandhoff disease also will be detected by this assay in either tissue. Some laboratories offer confirmation of Tay-Sachs carrier status by DNA testing; three specific mutations account for roughly 93% of mutant alleles among Ashkenazi Jews.

Recently, testing for Tay-Sachs carriers has been extended to include molecular testing for other disorders occurring at a higher frequency among Ashkenazi Jews. These panels also may include carrier testing for Canavan disease, cystic fibrosis, Niemann-Pick disease Type A, and Gaucher disease, all of which are performed by DNA analysis for population-specific mutations.

Mucolipidoses (ML)

These disorders are so named because they share clinical features of both the MPS and sphingolipidoses. Mucolipidoses I and IV

represent deficiencies in sialidase (neuraminidase) activity specific for glycoproteins and gangliosides, respectively. ML-II (I-cell disease) and ML-III (pseudo-Hurler polydystrophy) are single gene disorders characterized biochemically by deficient intracellular activity of multiple lysosomal enzymes involved in the degradation of mucopolysaccharides, oligosaccharides and sphingolipidoses, leading to abnormal storage of these compounds. The primary defect for both disorders is deficient activity of N-acetylglucosaminyl-1-phosphotransferase, the enzyme necessary for normal intracellular targeting of lysosomal enzymes. As a result, the enzymes are secreted extracellularly. The distinction between ML-II and ML-III is made on clinical grounds, with ML-II patients being much more severely affected. Biochemical diagnosis is made by demonstrating low intracellular activity of several lysosomal enzymes together with abnormally high activity outside cells (serum or culture medium). Diagnosis can be confirmed by demonstration of deficient phosphotransferase activity.

NEWBORN SCREENING

Part of "24 - Inborn Metabolic Errors"

Screening for genetic disorders falls into three categories: (i) neonatal screening, or the identification of infants with treatable metabolic disorders prior to the onset of clinical symptoms; (ii) prenatal screening, or the identification prior to birth of fetuses with certain genetic disorders; (iii) and carrier screening, or the identification of phenotypically normal individuals who are heterozygous for a particular mutation and therefore at risk of having affected offspring. There are numerous issues that must be addressed in determining which disorders are amenable to population screening. This section focuses on neonatal screening and the disorders typically included in these programs.

Newborn screening first became feasible in 1961, with the development of the Guthrie bacterial inhibition assay for measuring phenylalanine in dried blood specimens. The purpose of this test was to identify infants at risk for PKU prior to the onset of symptoms, so that dietary management could be instituted and serious clinical sequelae avoided. PKU screening has subsequently been instituted throughout the Unites States and other countries, and has also been expanded to include other metabolic and nonmetabolic disorders.

At least three aspects of PKU render this disorder amenable to newborn screening. First, untreated PKU leads to severe clinical symptoms, including mental retardation. Secondly, dietary management prevents the mental retardation, provided the treatment is begun early. Finally, PKU can be detected in the newborn period as an elevation of blood phenylalanine in affected individuals. Newborns now are screened prior to discharge from the hospital at 1 to 3 days of age; some states retest at 2 weeks of age to assure that all infants have ingested sufficient protein for the screen to identify an affected infant. The most effective screening programs are those in which all newborns are tested.

Many general issues also must be considered in the establishment of an optimal screening test. The test must be simple and inexpensive to perform, as well as reliable, reproducible, and accurate. Ethical, legal, and economic considerations also determine the appropriateness of a given test, including the estimation of both the adverse and advantageous consequences of the proposed screening. These are determined in part by the frequency of the disorder in the population, as well as by the number of affected individuals who will be successfully identified. In addition, the practicality and cost of preventive treatment must be considered and constantly reevaluated in the context of advances being made in biomedical technology. Finally, the test sensitivity (the proportion of people with the disease whose results are positive) and specificity (the proportion without the disease whose results are negative) must be balanced for each screening test, thereby minimizing the number of false negative results while holding false positive cases to an acceptably low number, which for newborn screening is to accomplish no false negatives.

Many of the newborn screening tests rely on the detection of abnormally elevated blood metabolites that accumulate as a consequence of a particular metabolic block. These include the tests for PKU, homocystinuria, and tyrosinemia. In contrast, other metabolic disorders are instead identified by the deficient activity of a specific enzyme, rather than the accumulation of substrate. These include the screening tests for biotinidase deficiency and the Beutler test for classical galactosemia (i.e., assay of gal-1-P uridyltransferase activity by the Beutler test). For these disorders in which enzyme activity itself is measured from the dried newborn blood spot, the accuracy of the test is independent of both dietary intake and in most cases clinical status of the infant.

The recent application of tandem mass spectrometry techniques to newborn screening has significantly increased the number of metabolic disorders potentially detectable by newborn screening, as well as decreased the number of false positives and negatives. At this time, only a few states utilize this technology, although this number will certainly grow in the future. Several resources are available with excellent summaries of newborn screening principles and procedures (see Suggested Readings).

SUGGESTED READINGS

Comprehensive metabolic testing and other genetics services are available at many major medical centers throughout the world; a number of excellent on-line resources are accessible from:

http://www.ncbi.nlm.nih.gov/Omim/ The Online Mendelian Inheritance in Man (OMIM) website contains a complete listing of human genetic disorders catalogued by OMIM number.

http://www.geneclinics.org/ GeneClinics is a clinical information resource relating genetic testing to the diagnosis, management, and genetic counseling of individuals and families with specific inherited disorders.

http://www.genetests.org/servlet/access GeneTests, providing current genetic testing information for families and healthcare providers.

http://mcrcr2.med.nyu.edu/murphp01/homenew.htm An Internet jump-station to sources of information on rare genetic diseases affecting children (New York University Medical Center).

http://biochemgen.ucsd.edu/ University of California San Diego Medical Genetics, with links to laboratories providing specific metabolic tests and to the Society of Inherited Metabolic Disease.

ADDITIONAL SUGGESTED READINGS

Additional suggested readings on inherited metabolic disorders include:

Bennett MJ, Rinaldo P, Strauss AW. Inborn errors of mitochondrial fatty acid oxidation. *Crit Rev Clin Lab Sci* 2000;37:1–44.

Burnila AB, Bonafe L, Zacchello F. Clinical and biochemical approach to the neonate with a suspected inborn error of amino acid and organic acid metabolism. *Semin Perinatol* 1999;23:162–173.

Burton BK. Inborn errors of metabolism in infancy: a guide to diagnosis. *Pediatrics* 1998;102:E 69.

Hommes FA, ed., *Techniques in Diagnostic Human Biochemical Genetics: A Laboratory Manual*. New York: Wiley-Liss, 1991.

Moser HW. Molecular genetics of peroxisomal disorders. *Front Biosci* 2000;5:D298–D306.

Saudubray JM, Martin D, deLonlay P, et al. Recognition and management of fatty acid oxidation defects: a series of 107 patients. *J Inher Metab Dis* 1999;22:488–502.

Seashore MR. Tandem spectrometry in newborn screening. *Curr Opin Pediatr* 1998;10:609–614.

Seashore MR, Wappner RS. *Genetics in Primary Care and Clinical Medicine*. Stamford, CT: Appleton & Lange, 1996.

Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Basis of Inherited Disease*. 8th ed. New York: McGraw-Hill, 2001.

Shapira E, Blitzer MG, Miller JB, and Africk DK. *Biochemical Genetics: A Laboratory Manual*. New York: Oxford University Press, 1989.

Wallace DC. Mitochondrial diseases in man and mouse. *Science* 1999;283:1482–1488.

The distinction between aminoacidemia and aminoaciduria is somewhat arbitrary, since for many disorders abnormal concentrations of amino acids occur in both plasma and urine. The terms are therefore often used interchangeably.

[1] The distinction between aminoacidemia and aminoaciduria is somewhat arbitrary, since for many disorders abnormal concentrations of amino acids occur in both plasma and urine. The terms are therefore often used interchangeably.

25
Point-of-Care Testing

Elizabeth Lee-Lewandrowski

Kent B. Lewandrowski

INTRODUCTION

Part of "25 - Point-of-Care Testing"

Point-of-care testing (POCT) is an emerging concept in laboratory medicine and has attracted considerable interest in the medical literature. Recent experience has revealed a variety of applications for these new technologies in hospitals, outpatient clinics, and home-care settings (1). Diagnostics manufacturers continue to invest significant resources in developing point-of-care (POC) technologies creating new opportunities to redesign the delivery of laboratory services. Some of the current uses of POCT have been discussed in other chapters in this book. In this chapter, the discussion will focus on general concepts relating to key issues relevant to medical, operational, and regulatory concerns. No attempt will be made to review the current literature on individual POCT technologies. For a more in-depth discussion of POCT applications, see the excellent presentation in reference 1.

POINT-OF-CARE TESTING: A HISTORICAL PERSPECTIVE

Part of "25 - Point-of-Care Testing"

Point-of-care testing (also called bedside or near-patient testing) has been defined by Price and Hicks as “all testing not undertaken in a central laboratory… that is undertaken close to the patient, with the result leading to a possible change in the care of that patient” (1). According to the Joint Commission on Accreditation of Healthcare Organizations, POCT is differentiated from central-laboratory testing in that it does not require a physical laboratory; that is, the laboratory exists at each point-of-care. The specimen is obtained, the test performed, and the result reviewed where care is provided. Despite various attempts, a rigorous definition of POCT is somewhat elusive, if not artificial. However, differentiating POCT from central laboratory testing is a useful concept from the management perspective because it involves testing by nonlaboratory personnel using specialized technologies adapted for near-patient bedside applications.

POCT had its beginnings many decades ago when physicians performed testing in hospital laboratories with or without the aid of a technologist. An archaic example of POCT was the practice of tasting a patient's urine to detect the presence of glycosuria indicative of diabetes mellitus. In the traditional academic medical center, the task of performing laboratory testing fell to the junior resident or intern house officer. The menu of laboratory tests was small, the available technology limited, and regulatory and medical-legal issues were of little concern. Subsequently, hospitals established clinical laboratories capable of performing an increasing diversity of tests. With the use of automated technology, the volume of testing also increased markedly. Over time, medical technology evolved into a formal profession complete with undergraduate and graduate training programs and a system of certification examinations designed to ensure competency. For many years, POCT remained unrecognized and limited to a few selected types of tests such as fecal occult blood testing, physician performed microscopy, and urinalysis. Under the traditional fee-for-service system of reimbursement, there was little incentive for doctors to evaluate and treat patients rapidly or to expedite noncritical laboratory testing. This mentality began to change as managed care became an increasingly important force influencing the medical reimbursement system. Efficient cost-effective management of patients now is essential to quality medical care and in many cases to the financial success of the provider. This shifting landscape has resulted in a reassessment of the role of the clinical laboratory and of the opportunities afforded by point-of-care testing.

POINT-OF-CARE TESTING: GENERAL CONCEPTS

Part of "25 - Point-of-Care Testing"

For the clinical pathologist responsible for providing laboratory services, there are three options for performing a test as shown in Table 25.1. These include reference laboratories, hospital laboratories, or testing at the POC. Reference laboratories include national commercial operations such as Quest Diagnostics and Laboratory Corporation of America. These laboratories boast a large test menu and a significant economy of scale owing to the volume of testing performed. Economy of scale translates into a low unit cost (although charges may be much greater). The major disadvantage of reference laboratories lies in the fact that their remote location results in long turnaround times that may be inappropriate for some types of testing. At the opposite end of the spectrum is POCT. In the case of POCT, the available menu is limited and the economy of scale is minimal because tests are performed one at a time. Consequently, the unit cost of POCT tends to be higher than other alternatives. However, the turn around time for POCT is usually only a few minutes. The hospital laboratory presents another option. On the one hand, some hospital laboratories have grown to impressive size with large

menus, automated instrumentation, and significant economy of scale. In this respect, hospital laboratories may resemble regional reference operations. On the other hand, many laboratories have set up STAT laboratories and urgent-care facilities with pneumatic tube transport systems to facilitate rapid transport and short turnaround time. These STAT laboratories provide a reasonable alternative to POCT in many applications, and may offer some cost and quality advantages. The needs of each institution for laboratory testing are highly variable depending on geography, available infrastructure, logistics, the type of medical services, patient acuity, reimbursement, regulatory requirements, and a variety of other factors. The three options for providing laboratory testing (reference laboratories, hospital laboratories, and POCT) compete to serve the needs of the patient and physician. Many solutions therefore are possible depending on a number of complex and at times contradictory factors. In the final analysis, most hospitals utilize all three options to a greater or lesser degree.

TABLE 25.1. OPTIONS FOR PROVIDING LABORATORY SERVICES

	Reference Laboratory	Hospital Laboratory	Point-of-Care
Menu	+++	++	+
Economy of Scale	+++	++	+
Turnaround Time	+	++	+++

Enthusiasm concerning the opportunities provided by POCT to achieve rapid turnaround time has resulted in substantial and continued investment in POCT technologies by the diagnostics industry. At the same time, hospitals are learning how to manage POCT performed by nonlaboratory personnel using interdisciplinary teams composed of the laboratory, nursing, and hospital administration. Consequently, the opportunities to implement successful POCT programs have been expanding as technology has improved progressively (Table 25.2). However, POCT continues to present major challenges including issues of cost, medical necessity, analytical quality, the performance of tests by nonlaboratory personnel, the impact of POCT on hospital operations, and regulatory concerns.

TABLE 25.2. MAJOR TRENDS IN THE EVOLUTION OF POINT-OF-CARE TESTING

- Expanding menu
- Consolidated platforms
- Improved analytical technologies
- Improved data management systems
- Product-line orientation of major suppliers
- Improved management of POCT using multidisciplinary teams

TABLE 25.3. EXAMPLES OF TESTS AVAILABLE AT THE POINT-OF-CARE

Chemistry
- Glucose
- Blood gases and electrolytes
- Cholesterol
- Cardiac markers
- Urinalysis
- Urea nitrogen
- Creatinine
- Ionized calcium
- Drugs of abuse
- Pregnancy testing

Hematology:
- Prothrombin time
- Partial thromboplastin time
- Activated clotting time
- Blood count

Microbiology:
- Physician performed microscopy
- Rapid group A Strep testing
- H. pylori testing
- Influenza testing
- Sexually transmitted diseases

Other:
- Fecal occult blood
- Gastric occult blood

Menu For Point-of-Care Testing

The menu of tests available at the POC has grown rapidly as shown in Table 25.3 to include many of the commonly performed tests in the chemistry, hematology, and microbiology laboratories. Some of these tests are performed on instrument-read quantitative systems with electronic data management capability (e.g., bedside glucose testing, blood gases, and electrolytes) while others utilize visually read devices (noninstrumented tests) producing a semiquantitative result (dipstick urinalysis) or a simple positive-negative result (*Helicobacter pylori* serology, fecal occult blood, pregnancy testing). As the menu of tests available at the POC has expanded, some manufacturers are consolidating multiple tests on a single platform. For example, devices are available that can perform blood gases, electrolytes, blood-urea nitrogen, creatinine, ionized calcium, glucose, activated clotting time, and other tests on a single hand-held device. As a consequence, most of the test menu required for emergency-department and acute-care applications now can be performed at the POC. The future will inevitably produce further expansion of the test menu and consolidation of platforms. It is reasonable to predict that most of the high volume and critical turnaround-time-dependent tests eventually will be available at the POC and that the majority of the conventional laboratory test volume will not necessarily be performed in the central laboratory. Where testing will ultimately be performed may depend on many variables, but lack of available technology will not be a limiting factor.

Utility Of POCT

POCT may be used in a variety of applications as shown in Table 25.4. These include patient screening, diagnosis, medical management, and management of patients. Medical management refers to testing performed for the purpose of managing a clinical condition such as bedside capillary glucose testing to maintain glycemic control in diabetic patients. In contrast, management of patients refers to testing performed to assist in managing the patient within the hospital. For example, cardiac markers may be used in conjunction with clinical data to triage patients to one of several different types of hospital beds such as an observation

unit, a general medical unit, or intensive-care setting. The decision is more related to hospital operations than medical issues with the aim being to provide the most cost-effective and clinically appropriate setting in which to care for the patient. In all of the above applications, the key to the selection of POCT versus the central laboratory is the priority of rapid turnaround time over other considerations. Although the unit cost of POCT tends to be higher than central-laboratory testing, when the total cost of care is considered, POCT may in some cases prove to be more cost effective.

TABLE 25.4. CLINICAL UTILITY OF POINT-OF-CARE TESTING: SOME EXAMPLES

Screening: Fecal occult blood, cholesterol, capillary blood glucose testing for diabetic screen at a health fair
Diagnosis: Pregnancy testing, cardiac marker, rapid Strep A
Medical Management: Bedside glucose testing, gastric occult blood, blood gases, electrolytes, activated clotting time
Management Of Patients: Cardiac markers used to triage patients to an observation unit versus an inpatient medical unit, urine pregnancy screen as a pre-radiology test in a trauma unit
Home Use: Patient Self Monitoring: Capillary glucose testing, ovulation testing

A growing application for POCT is home testing used for patient self-monitoring. Self-monitoring by diabetic patients using hand-held over-the-counter glucose meters has become an accepted standard of care. Capillary blood-glucose testing has radically improved the ability of diabetic patients to maintain glycemic control compared to obsolete methods such as urine glucose testing using dipstick methods. In the future, noninvasive and continuous glucose monitoring devices will become available further enhancing the ability of the patient to manage their condition and reducing the discomfort of obtaining specimens using the fingerstick method. A number of other tests now are available for home use including urine pregnancy testing, ovulation monitoring kits, fecal occult blood, and drug-screening kits using hair specimens to name just a few. Patient self-monitoring of anticoagulation testing using simple hand-held devices has become available, but remains controversial. Home testing using over-the-counter kits presents a radical departure from the traditional method of obtaining laboratory test results in that the patient, not the physician, chooses the test and interprets the result. The future of home testing remains uncertain and presents many challenges to the traditional approach endorsed by the medical establishment.

Analytical Performance

Many studies have reported on the analytical performance of testing devices at the POC. Evaluations on the performance of these technologies must take into account the capabilities of the devices under controlled conditions as compared to the performance of testing in actual practice in the hands of nonlaboratory personnel (2). Many POC devices are not as accurate or precise as testing performed in the clinical laboratory by skilled medical technologists. There are several reasons for this including limitations imposed by the capabilities of the technologies themselves and operator-related errors in the preanalytical and analytical phases of testing. Some early POC devices exhibited significant issues with analytical performance. In many cases improved engineering has reduced operator errors and improved the precision and accuracy of POCT devices. When evaluating a laboratory or POCT device, it is important to be aware of the limits of performance necessary for the clinical care of the patient. POCT systems do not necessarily need to be as accurate or precise as testing performed in the clinical laboratory provided the performance characteristics of the system are sufficient for the clinical application. For example, bedside glucose testing is widely used for the management of patients with diabetes as an aid to direct therapy (3). These technologies have revolutionized the care of the diabetic patient and have facilitated the ability of patients to maintain tight glycemic control. Most of these devices are only accurate to within +/- 10% to 15% of a reference value, but this level of performance is adequate for the clinical application. In contrast, laboratory-based glucose measurements using reference methods are considerably more accurate, but the turnaround time typically is too long for hospitalized patients and is hopelessly inadequate for outpatient or home-use settings. A different situation occurs in the diagnosis of diabetes mellitus. In this application, accurate reproducible results are essential to reliably classify patients and avoid over- or under-diagnosis. POC devices are not sufficiently reliable for use in this application and laboratory-based measurements are strongly preferred. In contrast, POCT glucose meters have been used in public-health applications to screen populations for diabetes mellitus. Individuals that test positive on the screen then are referred for follow up visits and formal laboratory testing. Thus the analytical performance of a given technology must be understood in terms of how the device will be used in clinical practice rather than a dogmatic adherence to the doctrine that greater accuracy or precision is itself the desired endpoint. This does not mean that all POC technologies are “good enough” for clinical application rather that these technologies must be evaluated in the context in which they will be used. Some devices currently available and approved by the Food and Drug Administration are not necessarily reliable. Therefore, the laboratory should assist other medical professionals in evaluating and validating POCT devices. Various parameters should be addressed when validating a new technology for laboratory testing (Table 25.5). The extent of the validation should at a minimum fulfill basic regulatory requirements. Beyond this, the validation should be extensive enough to ensure the reliability of the technology in clinical practice. This may require judgment and considerable expertise and represents one area where the laboratory can be particularly helpful in supporting the POCT program. Two caveats to consider when evaluating POCT technologies is how the devices

compare to tests offered in the clinical laboratory and what differences in sample types may be encountered. For example, the laboratory may offer troponin T for the evaluation of patients with chest pain whereas most POCT systems use troponin I. Reference ranges for the same tests (e.g., two troponin I methods) may be different, resulting in confusion among physicians. Finally, most POCT systems use urine or whole blood whereas central laboratories use plasma or serum. Implementing POCT therefore may necessitate careful consideration of many factors unique to the individual site and requires close communication with the clinical staff.

TABLE 25.5. FACTORS TO CONSIDER IN THE VALIDATION OF POINT-OF-CARE TECHNOLOGIES

- Ease of use, ease of training
- Data management
- Accuracy
- Precision
- Reference ranges
- Interferences
- Crossover studies to existing methods
- Quality control frequency and ranges
- Sample selection

MANAGEMENT OF POINT-OF-CARE IN THE HOSPITAL

Part of "25 - Point-of-Care Testing"

Traditionally, POCT was unregulated and consequently little effort was made by hospitals to establish an approach to managing these technologies. Predictably, issues concerning the quality of test results became apparent (2) resulting in the establishment of regulatory requirements (see below) that have been enforced by hospital accreditation organizations. Failure to comply with regulatory mandates may have major consequences. The need to address regulatory requirements combined with rapid expansion of POCT technologies has resulted in a concerted effort by institutions to manage POCT as a formal hospital program.

The first step in developing a strategy to manage POCT involves setting up an interdisciplinary POC management team, including the laboratory, physicians, and nurses (4). The POC team has the responsibility of determining the test menu, selecting technologies, establishing policies and procedures, ensuring training and regulatory compliance, and providing advisory assistance to the end users of POC technologies. Establishing a POC team requires financial support from the institution. At Massachusetts General Hospital nearly $100,000 is spent on salary support for the POCT management team. This figure does not include the cost of office and laboratory space, consumables and office supplies, administrative overhead, and other nonsalary expenses. Once a POC team has been established, a management structure must be put in place to ensure that the team has the authority to implement new initiatives and to perform corrective action where necessary. Figure 25.1 shows the management structure for POCT at the Massachusetts General Hospital. Oversight of POCT is the responsibility of the Point-of-Care Steering Committee, which includes the chief of pathology, the director and program manager of POCT, the chief medical officer for the hospital, the vice president of ancillary services, and the associate director of nursing. This committee provides high-level oversight and has interdisciplinary authority to mandate compliance with the POC program. Beneath the steering committee is the Point-of-Care Management Team, composed of a medical director, a laboratory program manager, a nurse manager, and a medical technologist. The POC management team is responsible for direct oversight of POCT, determining medical need, selecting technologies, establishing policies and procedures, overseeing regulatory compliance and quality assurance, and ensuring that individuals performing POCT have been adequately trained. The POC management team also ensures that each site performing POCT has assigned a medical director and a manager who are responsible for supervising testing activity, ensuring regulatory compliance, and serving as a liaison to the POC management team. In this system, responsibility is placed with the end user. The POC management team serves in a consulting role and coordinates the POC program throughout the institution. However, responsibility for compliance rests with the individual testing site, not the laboratory or administration.

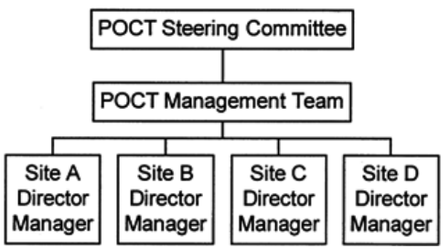


FIGURE 25.1. Management structure for point-of-care testing at the Massachusetts General Hospital: circa 1999.

The first task of the POC management team is to assess what testing is being performed in the institution. This may include a combination of e-mail surveys and site visits to inspect testing areas and interview staff. Many types of tests may be discovered that were unknown to the POCT team. In a number of cases, we discovered POCT during site inspections that were not reported during e-mail surveys. This is particularly true for manual POC tests such as fecal and gastric occult blood, various urinalysis products, and urine pregnancy testing.

The next step is to assess the medical need for each POC test and determine which technologies are most appropriate for the intended application. This requires interviews with physicians or nurses and is best performed by the POCT medical director. Questions to be considered when assessing medical need include assessing why the site is performing POCT, what turnaround time is required both from a medical perspective and from the perspective of unit operations, and finally what analytical performance is acceptable for the intended application.

The next task is to develop policies and procedures. This includes establishing a quality control and quality assurance program to ensure regulatory compliance and to maintain quality testing. A corrective action plan to remedy ongoing deficiencies is essential to the success of the program. In some cases, this may require mandates to discontinue testing in noncompliant sites. The corrective action plan is particularly important for manual POCT where auditing of quality control and quality assurance activities is more difficult than for instrumented tests because of the necessity of maintaining manual records. In contrast, oversight of instrumented tests can be facilitated by electronic data management systems and therefore requires less onsite assessment of compliance. The key to successful corrective action is to ensure that the POC team has the authority to make decisions concerning the suitability of testing on any given site. When a site is found to be noncompliant, an effort should be made to educate the personnel who are performing tests as to the necessary requirements for compliance with the program. If remedial action is not successful, consideration should be given to removing

testing when compliance cannot be achieved with the quality assurance program. One recommendation to ensure the authority of the POCT management team is to have an institutional policy whereby the POCT team controls the Clinical Laboratory Improvement Amendment (CLIA) certificates (see below) for all testing sites.

Once a POCT program has been established, the team should begin a continuous process-improvement effort designed to enhance the efficiency and ease of compliance with the program. This requires ongoing site inspections, continuing education, and may involve implementing new technologies, consolidation of technologies, and in some cases, elimination of POC tests that do not meet quality standards. For example, our site inspections revealed that a number of urinalysis dipstick products were being used in our hospital for a variety of applications. These include screening for hematuria and urine pH on oncology units, monitoring hydration status in pediatric patients and those with diabetes insipidus, and general urinalysis applications. Each product required its own policy and procedure, quality control, training, and competency assessment programs. First, the medical need for each type of test was studied to ensure that the application was appropriate and that the technologies performed at an acceptable level for the intended application. For example, pH paper to measure urine pH had not been validated and was eliminated. Owing to the availability of multiple products, maintaining compliance with the POC program was extremely difficult for users of urinalysis test strips because there were too many different products, standards for performance were not established, and some of the applications were inappropriate. The management decision was to consolidate all of these technologies to a single urinalysis strip and to establish one set of policies and procedures, training, and quality control. This radically simplified the management of POC urinalysis testing in our institution.

REGULATORY ISSUES

Part of "25 - Point-of-Care Testing"

Regulatory issues that pertain to POCT include requirements mandated by the CLIA 88, and subsequently by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the College of American Pathologists (CAP), and in some localities, state or other organizational requirements. Prior to these regulations, there were few guidelines for POCT. Predictably, major concerns regarding test quality, operator competency, and management of test results became apparent. Regulations imposed by CLIA 88 divide POC tests into two categories, CLIA-waived and nonwaived tests, the latter including moderate and highly complex subtypes. Waived tests are defined as tests that use simple accurate methods that make the possibility of errors unlikely or pose no significant risk of harm if performed incorrectly (5). Examples of CLIA waived tests are shown in Table 25.6. According to CLIA 88, all laboratory testing with the exception of waived tests are subject to certain basic requirements for quality control, patient records, and quality assurance (Table 25.7). Waived tests are exempt from these regulations, but are still subject to guidelines imposed by the JCAHO or the CAP. These requirements have evolved over time reflecting the need to make guidelines concerning POCT realistic and practical. In the past, both the JCAHO and the CAP imposed requirements that were arbitrary at best and were unachievable in many situations. Regulatory guidelines continue to evolve as accreditation agencies learn what guidelines promote practical, high quality POCT versus those that are superfluous.

TABLE 25.6. EXAMPLES OF CLINICAL LABORATORY IMPROVEMENT AMENDMENT WAIVED POINT-OF-CARE TESTS

- Dipstick urinalysis (but not microscopic analysis)
- Dipstick microalbumin
- Bedside glucose testing (if the device is FDA approved for home use)
- H pylori serology (selected kits)
- Fecal occult blood
- Hemoglobin (copper sulfate or Hemocue)
- Spun hematocrit
- Urine pregnancy testing
- Ovulation test kits
- Sedimentation rate (nonautomated)
- Group A strep antigen (selected kits)

TABLE 25.7. SUMMARY OF REQUIREMENTS FOR NON-WAIVED TESTING MANDATED BY THE CLINICAL LABORATORY IMPROVEMENT AMENDMENTS OF 1988 (CLIA 88) (3)

- Testing covered under appropriate CLIA 88 certificate
- Personnel requirements: Specific training requirements for director, technical consultant, clinical consultant and testing personnel
- Proficiency testing standards
- Patient test management standards: Specimen collection, handling, reporting, records
- Quality control standards: Including procedure manuals, per-formance verification, equipment maintenance, calibration, specific quality control, and corrective action
- Quality assurance: Systems established to monitor quality (pranalytic, analytic, postanalytic) and assessment of staff competency.

In order to perform testing, all sites must be covered under a CLIA 88 certificate. Different types of certificates are available including a waived testing certificate, a provider performed microscopy (PPMP) certificate and a certificate for moderate and high complexity testing. Waived tests are not covered under federal law concerning personnel or testing requirements. However, both the JCAHO and CAP have instituted standards for waived tests as part of the hospital/laboratory accreditation process. The standards of the JCAHO and the CAP differ somewhat and in some facilities there also may be state requirements as well. An excellent review of the perspective of the JCAHO concerning requirements for POCT is provided in reference 6 and of requirements mandated by CLIA 88 in reference 7. Further details may be obtained directly from the JCAHO or CAP.

PERSPECTIVES ON COST

Part of "25 - Point-of-Care Testing"

Conventional wisdom concerning the cost of POCT is that testing performed at the bedside is more expensive on a unit-cost basis than testing performed in the clinical laboratory where the

economy of scale is typically higher. For example, in one study at a Veterans Affairs medical center, the cost of bedside capillary glucose testing was reported to be 3.6 times more expensive ($11.50 per tests versus $3.10) than testing performed in the clinical laboratory. Other studies have shown that POCT is not invariably more expensive than testing performed in the clinical laboratory depending on test volume, technology, program management, and other factors (8). In another example, Bailey et al. reported an institutional savings of $392,336 with a reduction in unit cost from $15.33 per panel to $8.03 following institution-wide implementation of POCT for blood gases and electrolytes (9). Despite a growing literature on this subject, a determination of the actual cost of POC testing versus the central laboratory remains an enigma. Other factors besides unit cost also must be considered such as the impact of reduced turnaround time on the time to diagnosis and overall cost of care, issues of workflow on clinical units, impact on test utilization, and benefits to patient outcomes to name only a few. Collectively, these factors are too complex to be understood solely in terms of unit cost. Kost has pointed out that “there is no adequate generalization. Cost effectiveness analysis, consideration of marginal costs of competing alternatives, and other means of economic assessment must be individualized to the institution and setting.” Lacking clear consensus in the literature, the decision of whether to implement POCT versus using a central laboratory will depend on a number of features unique to the individual hospital or care unit. Lacking reliable outcomes data in the literature, the approach will be in part intuitive and may require trial and error to arrive at a satisfactory solution.

THE FUTURE OF POCT: POSSIBLE SCENARIOS

Part of "25 - Point-of-Care Testing"

Technological improvements in POCT devices have resulted in greater ease of use, improved analytical performance, and lower unit cost because of the development of less expensive, more efficient platforms. Many hospitals have learned to manage POCT successfully and to improve compliance with regulatory guidelines. As the menu available at the POC expands and platforms are consolidated into multianalyte systems, new opportunities to exploit these technologies will be possible. Currently, the rate-limiting factor for most POCT programs is the ability to manage quality control and patient data in an efficient reliable manner. Also, problematic is the fact that many technologies cannot be integrated electronically (including both instrumented and noninstrumented tests) into a turnkey POC solution. Until these challenges are overcome, POCT will be limited to specialized applications in selected settings.

Assuming further improvements in POCT technology, various scenarios are possible. At one end of the extreme, virtually all turnaround-time-dependent tests could be moved out of the central laboratory to the bedside where immediate test results would be available to providers of clinical care. At the opposite end of the spectrum, POCT may remain only a minor component of total testing volume because of limitations imposed by cost, technology, and system management. The most likely outcome probably will lie between these two extremes with continued growth in POCT and expansion of the menu. Each institution will seek its own solutions depending on its unique circumstances resulting in a limitless number of permutations. Clearly, POCT will challenge laboratory professionals for many years to come.

REFERENCES

1. Price C, Hicks J. Point-of-care testing: an overview. In: Price C, Hicks J, eds. *Point-of-Care Testing*. Washington D.C.: AACC Press, 1999:3–16.

2. Lewandrowski K, Cheek R, Nathan D, et al. Implementation of capillary blood glucose monitoring in a teaching hospital and determination of program requirements to maintain quality testing. *Am J of Med* 1992;93:419–426.

3. Lewandrowski E, Lewandrowski K. Bedside capillary glucose testing in the management of patients with diabetes mellitus: clinical utility and relationship to patient outcomes. In: Hicks J ed. *Point-of-Care Testing*. Washington D.C.: American Association for Clinical Chemistry Press, 1999:319–336.

4. Lewandrowski E, Lewandrowski K. At the bedside: managing manual point-of-care testing. *Advance For Admin of the Lab* 2000;8:8–10.

5. Seiler J. At the bedside: regulatory issues for POCT. *Advance For Admin of the Lab* 2000;7:8–10.

6. *Quality Point-of-Care Testing. A Joint Commission Handbook*. Oakbrook Terrace, Ill.: Joint Commission on Accreditation of Healthcare Organizations, 1999.

7. Belanger A. Alternate site testing. The regulatory perspective. *Arch Path Lab Med* 1995;119:902–906.

8. Lewandrowski E, Laposata M, Eschenback K, et al. Utilization and cost analysis of bedside capillary glucose testing in a large teaching hospital: implications for the management of point-of-care testing. *Amer J Med* 1994;97:222–230.

9. Bailey TM, Topham TM, Wantz S, et al. Laboratory process improvement through point-of-care testing. *Journal of Quality Improvement* 1997;23;7.

26
Tumor Markers

Morton K. Schwartz

The use of cancer markers goes back more than 150 years. In 1847, Sir Henry Bence-Jones described a protein in urine that could be identified by its unique thermal properties. It becomes slightly opaque when cooled. When citric acid is added, it effervesces, assumes a reddish hue, and becomes quite clear, but as it cools it becomes opaque. This protein, known as “Bence-Jones protein” or, as Sir Henry Bence-Jones described it, “hydrated deutoxide of albumen,” served as a definitive diagnostic test for myeloma for 100 years. With the introduction, first of electrophoresis and then of immunological procedures, the Bence-Jones protein was identified as the light chain of immunoglobulin G (IgG). These techniques allowed the identification of IgG light chains in the urine of some individuals who did not exhibit Bence-Jones protein or have myeloma (1).

Another early application of testing in cancer was by Sir Michael Foster who in 1867 reported the presence of amylase in blood from patients, and proposed the use of urinary amylase as a marker for cancer of the pancreas. As more information was obtained about the pathology and physiology of the pancreas, it was realized that pancreatic cancer originates in the ductal cells and not in the acinar tissue, which produces amylase. Amylase elevations therefore were only significant in cancers with a large tumor burden, which impinged on acinar tissue (2). Other early markers include acid phosphatase (prostate cancer), alkaline phosphatase (bone cancer), catechols (pheochromocytoma and neuroblastoma), 5-hydroxyindole acetic acid (carcinoid), and pituitary gonadotrophins (choriocarcinoma). These markers are still in use, but the techniques for their measurement all have been improved and both their sensitivity and accuracy increased (3).

DEFINITIONS

Part of "26 - Tumor Markers"

It was not until the invention of the radioimmunoassay and its application in immunoassays permitting quantitative assay of nanogram concentrations of protein that the term “tumor marker” was introduced and became a part of the everyday laboratory vocabulary. Tumor markers originally were defined as substances that can be measured quantitatively by immunochemical means in tissue or body fluids to identify the presence of a cancer and possibly the organ where it resides, to establish the extent of tumor burden before treatment, to predict prognosis, and to monitor the response to therapy. This definition now must be modified to include markers in solid tissue that can be identified semiquantitatively or quantitatively by immunochemical staining, cell sorting, or molecular diagnostic techniques as well as substances released from the cell (usually from the surface membrane) and found circulating in body fluids, where they can be measured by a variety of techniques. The range of molecules that now are classified as tumor markers is considerable, including tumor-associated antigens, nuclear and specific proteins, enzymes, isoenzymes, genes, oncogenes and their products. The performance of tumor markers in the United States requires use of a Food and Drug Administration (FDA) approved assay (4, 5). At Memorial Sloan Kettering Cancer Center, this volume has increased dramatically every year. The three tumor markers with the most remarkable growth are CEA, which has increased from 10,925 tests in 1987 to 22,125 in 1999; PSA from 569 tests in 1987 to 23,693 in 1999; and CA15-3 from 0 in 1987 to 14,661 in 1999. CA125 has increased in that time from 1,246 to 7,185, but AFP and hCG have remained constant at about 3,500 tests. In 1999, free PSA and CA19-9 were introduced.

Clinical application of tumor markers requires an understanding of the epidemiological and analytical definitions of sensitivity and specificity and the importance of prevalence. In addition, we must appreciate the analytic limitations of the assay and be aware of preanalytical and postanalytical effects on the marker. Epidemiological sensitivity is the percentage of true positives. Specificity is the percentage of true normals. The positive predictive value is a percentage of the true positives compared to positive values. The negative predictive value is the true negatives compared to all negatives. The efficiency of the test is the true negatives and true positives divided by the number of tested individuals.

Prevalence of the disease plays an important role in the use of a marker for cancer detection. In a test with 95% sensitivity and specificity, the predictive value would be only 16.2% when the prevalence was 1%, 50% when the prevalence was 5%, and 95% when the prevalence was 50%. Analytical sensitivity can be defined as the lowest detectable amount of marker measured by the method in use, while analytical specificity reflects the extent of exogenous material interferences in the assay (4, 5).

It also is of extreme importance to note that observed values in the same specimen measured by different methods may be different and these values cannot be used interchangeably. These differences may be as a result of poor calibration, differing antibody

specificities, variation in quoted reference ranges, and assay robustness. Robustness is related to hook effects, linearity, and the possible presence in the patient's serum of human antimouse antibody (HAMA). This is becoming more important as cancer patients in larger and larger numbers are treated with monoclonal antibodies of mouse origin (6).

Carcinoembryonic Antigen (CEA)

When first described, CEA fulfilled all the requirements of an ideal tumor marker. It apparently was elevated in almost all patients with colorectal cancer, but was normal after successful removal of the tumor. 30 years later, it now is well-understood that CEA is elevated in all solid-tissue tumors, not only those of the colon or rectum and is elevated in only 4% of patients with early stage diseases compared to 65% of patients with distant metastases (4). Following successful therapy, elevations of CEA may fall and subsequent elevations suggest recurrence. These elevations may be seen months, if not years, before there is clinical evidence of disease. It is important to understand that elevations will not be seen in about 30% of individuals with recurrent metastatic disease. In two patients with equally progressive disease, a crescendolike rise in serum CEA concentration may be seen in one patient while little or no increase is observed for the other. This probably reflects the pathology of the tumor. Well-differentiated tumors produce CEA while those that are undifferentiated or poorly differentiated produce little, if any. In patients with metastatic gastric cancer, CEA elevations never were observed over a period of 1 year in individuals with undifferentiated cancers, whereas there was a mean rise of more than 20 times the pretreatment level in patients with highly differentiated tumors. It certainly would make interpretation of CEA data more understandable and useful if the pathology of the tumor was reported with CEA values (7).

During therapeutic monitoring, it is important to understand the significance of changes in CEA values. There is no definition of a meaningful decrease or increase of CEA. In general, a 25% to 35% change has been termed significant. Four patterns of change of serum CEA values may be observed. Two patterns are expected: an uninterrupted rise in patients who do not respond to treatment, and a fall in those who do. There are two other patterns. In patients who are responding to treatment, there may be a surge in CEA for weeks posttreatment, followed by the conventional fall. In others, there may be an immediate but sustained fall in serum CEA and then the rise that indicates lack of response. It is apparent that if CEA is to be used in evaluating response to therapy, an understanding of these kinetics is essential (7, 8).

An expert panel of the American Society of Clinical Oncology (ASCO) has concluded that CEA cannot be used for screening for colon-rectal cancer, but preoperative CEA may assist staging and surgical treatment planning. The panel also concluded that CEA elevations detect recurrence earlier than other techniques and CEA monitoring should be done at the start of treatment and then every 2 to 3 months thereafter (9). CEA also has been found to be useful in monitoring cancer of the breast, ovary, and pancreas. Elevated CEA values are useful prognostic indicators (4).

CA19-9

CA19-9 is the marker most useful in adenocarcinoma of the pancreas. It is more sensitive (70% to 95%) and specific (72%) than CEA (40% to 60% and 70% respectively). Combined measurement of both markers does not significantly augment the sensitivity of the test. CA19-9 is not elevated in islet cell carcinoma of the pancreas. The value of CA19-9 is obscured by the fact that benign conditions such as acute or chronic pancreatitis or cholelethiasis may cause elevations. CA19-9 may be useful in monitoring patients who are receiving treatment. Many other markers have been studied in patients with pancreatic cancer, but none has the sensitivity or specificity of CA19-9. Of considerable interest is that CA19-9 is only found in the serum of individuals who secrete Lewis antigen. It is not found in serum of nonsecretors (10).

CA72-4

CA72-4 is a glycoprotein with a molecular weight of 220,000 to 400,000. It is clearly different from CEA, CA19-9, or other glycoprotein tumor markers. Although elevated in serum of patients with many different gastrointestinal cancers and only rarely in benign gastrointestinal diseases, it is more sensitive and specific than any other marker in gastric cancer. In one study, it was clearly shown that it is inferior to CA19-9 in pancreatic cancer (22% versus 82%) and to CEA in colorectal cancer (32% versus 58%). In gastric cancer, it identified 59% of the patients compared to 52% with CA19-9 and 25% with CEA. When combined with CA19-9, 70% of the gastric cancer patients were identified (11).

Alpha Feto Protein (AFP)

AFP has been used for more than 30 years in screening for hepatocellular cancer (HCC) and in diagnosis and monitoring of patients with germ-cell tumors. When screening for HCC, a disease of great importance in sub-Saharan Africa and Southeast Asia where hepatitis B and hepatitis C are endemic, the sensitivity of the assay is important. In an early study with an immunodiffusion method, 39 of 56 patients with HCC were detected and there were no false positives. When a radioimmunoassay was used in these same specimens, 49 of the 56 patients were detected but there were so many false positives that screening was not practical. The solution is either use of a less sensitive method or an increase in the cut-off to eliminate the false positive. When monitoring germ-cell tumors, particularly embryonal testicular cancer, the most sensitive method available is needed because successful therapy is defined as a return of the marker to no detectable level (12, 13).

In China, there have been widespread AFP screening programs in high-risk areas to detect HCC. Between 1971 and 1976, 1,967,511 persons were screened in Shanghai and 3,090 cases were detected (15.3 cases/100,000). In Qidong Province between 1976 and 1992, there were 499 cases identified in the 1,310,871 persons tested (38.1 cases/100,000). If only those individuals with liver diseases were considered, AFP detected 253 cases of HCC in 76,600 persons in Shanghai (330 cases/

100,000) and 48 cases in 4,875 persons (98.6 cases/100,000) in Qidong. In these studies, the 3- to 4-year survival of detected and treated patients was 57.1% and the 2-year survival, 69%. In a limited study of 3,454 hepatitis B antigen-positive patients, there were 31 HCC cases (897 cases/100,000).

When a highly sensitive AFP method was used, normal levels were seen in all of 450 normal individuals, in 98.2% of 564 patients who were hepatitis-B antigen positive and in 96.4% of 536 patients with nonmalignant disease or other cancers. With this method, elevations were seen in 23/28 (82%) HCC patients from Africa who were hepatitis-B antigen positive and in 45/57 (79%) hepatitis-B antigen positive HCC patients from the Far East. Abnormal values were observed in 38% of 16 HCC patients who were hepatitis-B antigen negative (14).

Although individuals with cirrhosis are at high risk for HCC, studies have suggested that AFP screening is not useful in detecting cirrhotic patients who will develop HCC (15). In addition to screening, AFP is of use in monitoring patients with HCC. The concentrations rise and fall, reflecting the course of the disease, and may be reflective of recurrence before any other clinical or diagnostic indication.

AFP is an important marker in germ-cell tumors, but only when used in conjunction with the ß-subunit of choriogonadotropin (ß-hCG) (16). The AFP observed in HCC is of liver origin and the AFP in germ-cell tumors originates in the yolk sac. It may be possible to differentiate the liver form of AFP from yolk-sac AFP by their different affinities for lentil lectins. Lectin binding of the AFP in HCC has been documented whereas the yolk-sac form does not react. This difference has been exploited for more precise identification of the source of AFP elevations.

Chorionic Gonadotropin (hCG) and its Beta Subunit (ßhCG)

hCG, which is produced by the syncytiotrophoblast cell after trophoblastic differentiation has been used for more than 50 years in the diagnosis and monitoring of trophoblastic cancers such as choriocarcinoma and hydatidiform mole. An elevation is essential for the confirmation of diagnosis, and changes in the concentrations clearly reflect the success or failure of therapy. AFP is not elevated in these cancers (16).

In nonseminomatous germ-cell tumors of the testes, a combination of AFP and hCG — more particular the ß subunit of hCG (ß-hCG)— are essential for the management of the patient. Presumably because of the heterogeneity of the cancer cell, it is essential to measure both ß-hCG and AFP when monitoring embryonal testicular cancer. Successful therapy is indicated by a fall of the elevated markers to normal. A rise in one or the other or both suggests recurrence, usually before there is clinical evidence of metastases. Very often during an exacerbation, only one of the markers will become elevated (16).

AFP and ß-hCG have been used to predict therapeutic response and to evaluate prognosis of testicular cancer. If the serum concentrations fall to 50% of the pretreatment value during a biostatistically defined period [7 days for AFP and 3 days for (hCG)], there is much poorer prognosis. In individuals with longer fall-off periods, the 9-year survival was 29% compared to 84% in men with a shorter fall off. A multivariate equation has been established that uses the concentration of ß-hCG, the activity of serum lactate dehydrogenase, and the number of metastatic sites to predict whether conventional therapy will be effective. Neither AFP nor CEA offered additional help. When this equation was used, the prediction of those who will respond is >90% accurate, but unfortunately, the equation properly predicts only approximately 50% of those men who do not respond (17, 18).

CA-125

CA-125 has been observed in the blood of patients with ovarian cancer but also is elevated in a variety of cancers (uterus, pancreas, liver, lung, and a variety of benign diseases). The assay is only approved by the FDA for monitoring second-look operations in women who have been treated for ovarian cancer. Elevations of CA-125 may be seen in many nonmalignant conditions. These include pregnancy, menstruation, ovarian cysts, endometriosis, and peritoneal or pleural inflammation. Because of the elevations in benign conditions and because CA-125 elevations are related to tumor burden, CA-125 cannot be used for early diagnosis (19).

However, CA-125 is an important marker because elevations after treatment suggest presence of residual tumor, and second-look exploratory surgery probably is unwarranted. In a literature survey, elevations were seen in all but one of 500 women with residual tumor prior to further treatment. About 50% of women with normal serum concentrations of CA-125 after treatment still had residual tumor. An elevated CA-125 is highly suggestive of residual tumor, but a normal value should not give a false sense of security that the tumor has been removed completely. It now is accepted to describe responses or recurrences when the only indication of a change in the patient's condition is the CA-125 (20).

CA-125 may be a prognostic predictor. In women in whom the fall of serum CA-125 after therapy followed a half-life of five days, prognosis was much better than in women in whom it did not. Eighty percent of women who experienced a return to normal CA-125 values within 30 days after therapy survived for more than 5 years, compared to only 10% of women who did not achieve normal values in this time.

It has been proposed that use of an algorithm and sequential CA-125 results may make it possible to use CA-125 for early detection (21). If the logarithm of CA-125 concentrations is plotted against time (taking two annual specimens initially and then after the first elevation, a further specimen every 3 months), the slope of the line becomes predictive. In a study of 1,855 women in the “training set” (the data used to develop the algorithm) and an additional 1,840 women in the validation set, the sensitivity for the early diagnosis was 83%, the specificity 99.8%, and the positive predictive value 16% (a value that exceeds the 10% expected in a screening program). Prospective clinical studies now are required to establish the true utility of this algorithm.

There is controversy over the role of CA-125 in screening and there is general agreement that single CA-125 values cannot be used for this purpose. Increases of CA-125 are observed in approximately 50% of patients with mucinous early (stage I or II) cancer and to a much lesser extent in women with nonmucinous

disease. A retrospective study was carried out with retrospective specimens obtained from the Norwegian Janus Serum Bank (specimens from 39,300 women collected in 1974 to 1986). Of these, 105 women developed ovarian cancer 1 to 143 months after the specimen was collected. The cancer patients were compared with 323 matched normal controls randomly selected from the bank. Fifty percent of the 105 women with ovarian cancer had abnormal CA-125 values as many as 18 months before diagnosis; 7% (23) of the control women had elevated CA-125 values. Twenty-five percent of the cancer patients had increased CA-125 levels 60 months before, compared with only 0.9% in the controls. The conclusion of this study was that CA-125 is useful in screening (22). In another study of 915 women without clinical symptoms, 36 had elevated CA-125 values; none had ovarian cancer but many had benign gynecological problems. Four who did not have an increased CA-125 died of nonovarian neoplasms. In screening of 1,010 asymptomatic postmenopausal women with CA-125, 43 abnormalities were observed. Abnormal pelvic examinations were seen in 28, over half related to fibroids or cysts. Thirteen women had abnormal ultrasound patterns; in the 12 who consented to laparotomy, one cancer was established.

In a study of 1,082 asymptomatic women over age 40, abnormal CA-125 values were found in 36. Additional specimens were collected and repeat assays were performed on those with the increases. A doubling of the value was considered a true positive. One of the 1,082 women had a clinically confirmed stage II ovarian cancer, diagnosed 21 months after the last specimen was collected. The incidence of detection in this study (1 in 1,082) would be 93 cases/100,000 screened, compared with the reported clinical incidence of ovarian cancer of 42.2 cases/100,000 (23). Evaluation of these screening studies suggests that if CA-125 is the only criterion; 30 laparotromies would be performed to confirm each ovarian cancer. Screening asymptomatic women in the United States by pelvic ultrasonography and CA-125 would cost 45 billion dollars. In women with operable pelvic masses, CA-125 can be very valuable. In a large study of women in Scandinavia, the sensitivity of CA-125 was 87% and the specificity 88%. The prevalence was 50% and the positive predictive value 88% (24).

Prostate-Specific Antigen (PSA)

The tumor marker that has received the most public attention over the past few years is prostate-specific antigen (PSA). PSA is prostate-tissue specific but not prostate-cancer specific. Elevations (PSA >4.0 µg/L) are seen in about 40% of men with early prostate cancer, 70% of men with more advanced cancer, and in 20% of men with benign prostatic hypertrophy (BPH). After surgical removal of the prostate, the PSA in serum falls with a half-life of 3.2 ± 0.6 days. Following radiation therapy the fall-off is much greater, and the half-life approaches 2 to 6 months (25).

PSA originally was approved by the FDA for monitoring therapy. In men whose cancer has successfully responded to treatment, serum PSA will reach very low (nondetectable) levels. A subsequent elevation to very low levels suggests recurrence. At 0.4 µg/L each of 16 men studied had recurrent disease. In another study, using a “super-sensitive” method with a reported minimum detection level of 0.025 µg/L, a rise to 0.04 µg/L was related to recurrence (26). The problem with the “super-sensitive” method is that many other tissues elaborate small amounts of PSA and if this assay is used in screening, it could conceivably create a large number of false positives (27).

In addition to monitoring, preoperative PSA values may be important prognostic markers and may be useful in influencing clinical decisions about orders for bone scans. Only three of 561 men with PSA <10 µg/L and only one of 467 patients with PSA <8.0 µg/L had an abnormal bone scan. These investigators concluded that when serum PSA is <10 µg/L bone scans are not necessary. It was estimated that if this criterion were used, 50,000 patients each year in the United States would not require radionuclide bone scans, which would represent a savings of $30 million (28).

The FDA recently has approved the use of PSA for screening. This use in screening for prostate cancer is supported by a large body of data suggesting that 60% to 70% of men with PSA values >10 µg/L will have biopsy-proven cancer. The problem in initiating large-scale screening programs is their cost and the fact that at autopsy, the large majority of men will have cancer cells in their prostates even though only 2% to 3% of these men will ever have a clinically significant prostate cancer. An absolute marker is not yet available to establish which prostate cancer will be clinically meaningful (29).

In an early study, 1,600 men who did not have urologic symptoms were compared to a control group of 200 men with urologic symptoms. In men with PSA values between 4.1 and 9.9 µg/L, 22% of the study group had biopsy-proven prostate cancer compared with 26% in the control group. In men with PSA values >10 µg/L, 64% of the screened population had biopsy-proven cancer compared with 64% in the control group (30). Other studies have found that the combination of digital rectal examination (DRE) and PSA detects 70% of cancers, compared with only 24% in individuals with abnormal DRE and normal PSA. DRE and PSA complement each other, and their combined use has been recommended (31).

The most compelling argument for the use of PSA in screening is a retrospective study of serum specimens collected over a 20-year period at a geriatric institution. In men who did have demonstrable cancer, there was an age-related small linear incremental increase in PSA over time. In men with BPH, the increase was still linear, but the slope was somewhat higher. In men who developed cancer, the rate of increase was exponential. In men with malignant disease, the exponential rise was much more rapid than those with localized cancer. Elevations preceded clinical evidence of prostate cancer by many years. It is conceivable that mathematical consideration of several PSA values collected over time may permit prediction of the eventual development of localized or malignant prostate cancer (32).

We now know a great many things about the physiology and pathology of PSA (33, 34, 35, 36 and 37). There is a physiologic variation of PSA concentrations in sequential specimens drawn over a short period of time. It was suggested that a meaningful change between successive specimens is ±30%. Hormone therapy (finasteride) may reduce PSA concentrations by as much as 50% without any effect on the cancer (36). Men with low endogenous

testosterone concentrations also may have low PSA values, and normal values may occur in such men even if they have prostate cancer. PSA values are age-related, young men having lower reference ranges and older men higher values. At age 70, the upper limit of the reference range is 6.5 µg/L. Black males have higher values than white males. Exercise may affect the values. In men who spent 30 minutes on an ergometer bicycle, there were elevations to twice the preexercise level.

Free and Complexed PSA

The question has been raised as to how PSA measurements can be used more effectively in the management of men with prostate cancer and made more accurate in distinguishing between benign disease and cancer. PSA exists in serum primarily as three forms. One form is complexed to the protease inhibitor α l-antichymotryspin, a second to α-2 macroglobulin and the remainder is noncomplexed (free PSA). The total PSA is presumably the free and the antichymotrypsin-bound form because the α-2 macroglobulin form is not immunoreactive. One proposed approach is to use the ratio between free and total PSA. In an original report, it was found that if the ratio was greater than 0.154, benign disease was present, and if it was lower, the patient suffered from cancer. With this cutoff, 93% of men with BPH were identified. Various cut-offs have been proposed and there is continuing controversy over the role of the ratio in the management of patients. When a cutoff of 0.197 was used, the sensitivity was 84.1% and the specificity was 46.8%. In another study, free and total PSA were studied in 773 men of whom 379 had histologically confirmed prostate cancer and 394 had benign prostatic hypertrophy. All had PSA levels between 4.0 µg/L and 10 µg/L. The normal/abnormal cutoff for free PSA was established emperically as that required to maintain 95% sensitivity for detection of prostate cancer in the prostate cancer patients. A 0.25 free PSA/total PSA ratio was determined and with it 95% of the cancers in the BPH patients were found. The procedure probably avoided 20% of unnecessary biopsies. The cancers in men with greater ratios than 0.25 were usually older men with less-aggressive tumors (38, 39, 40). Recently it has become possible to measure the complexed PSA by itself. With this technique, it has been found that in combination with total PSA, it may be possible to stratify men with total PSA between 2 to 4 µg/L (41).

Efforts have been made to use algorithms to improve the sensitivity of PSA testing, and to improve discrimination between prostate cancer and BPH, to define the aggressiveness of the malignancy and to establish whether a prostate cancer has extended beyond the capsule. Many of the equations include PSA, staging, histology and other nonlaboratory parameters. One proposed equation employs age and PSA, prostatic-acid phosphatase (PAP), and creatine phosphokinase (CPK) isoenzymes. With this equation, the sensitivity was 85.3% in Stage T2 cancers and 66.12% in BPH. It is claimed that the algorithm can double or triple the sensitivity of serum PSA. Seventeen of 28 patients (61%) with early stage cancer and 93 of 94 (99%) of men with more advanced cancer was identified. It should be emphasized that the computer program for these algorithms is proprietary. The important point is that neither CPK isoenzymes, total CPK, or PAP alone are useful markers in detecting prostate cancer (42).

Acid Phosphatase

Acid phosphatase is one of the oldest cancer markers and has been used for more than 50 years in monitoring patients with prostate cancer. Elevations are seen in about 80% of men with bone metastases but in 20% or less of men with localized cancer. In a study of 80 patients with stage D prostate cancer, abnormal levels of PSA were seen in 76% of the men compared to elevations of prostatic acid phosphatase in 49%. The conclusion of the authors is that PSA is more sensitive and more useful than acid phosphatase and will be elevated in patients with smaller tumors (43). In our experience, there is a small subset of patients who do not show clinically related PSA levels who have acid phosphatase levels that reflect the clinical course. We continue to offer PAP assay.

Human Kallikrein-2 (hK2)

PSA is a kallikrein (hK-1). Another form hK-2 also is elevated in prostate cancer. Both are serine proteases, but have different biochemical and catalytic properties. In a study of 324 men, 159 of whom had prostate cancer, mean hK-2 levels and the ratio of hK-2 to free PSA were significantly higher than in serum from control subjects. The conclusion was that in men with elevated PSA, those with high hK-2 have a 5- to 8-fold increase in risk for prostate cancer. hk-2 may be a useful adjunct to PSA in selective patients with elevated PSA who require a biopsy (44).

Breast Antigen (CA15-3, BR27.29)

Circulating breast cancer antigens are used in monitoring therapy and recurrence. These antigens are all high-molecular-weight glycoproteins (mucins) coded by the MUC-II gene and expressed on the ductal cell surface of most glandular epithelia. Although many of these markers have been studied in the United States, only two (CA15-3 and BR27.29) are approved by the FDA and only for monitoring. The antibody configurations in the assays are slightly different, and different numerical results may be obtained. Elevations are directly related to stage. In one study of CA15-3, 9% of women with stage I disease and 19% of those with stage II disease had elevations in serum marker concentrations, as compared to 38% in women with stage III cancer and 75% with stage IV. Similar distributions were observed with BR27.29 assay. The detection rate in early-stage disease may be improved with automated methods, which are more precise at low concentrations. The major application of the markers is in monitoring disease progression and response to therapy. The efficiency for identifying progression during therapy was >90% with a lead time of 35 days, and about 85% during follow-up, with a lead time of 75 days. In a study of BR27.29, all patients who experienced a regression of disease had a falling breast cancer antigen value, those with stable disease had values that did not change, and 94% of women with progressive disease showed an increase (45, 46). In a summary of the literature, it was found that of 1,672 patients with breast cancer, there were 352 who

had recurrences. Of these, 235 (67%) had elevations in the breast antigen before or at the time of recurrence. In 1,320 patients without recurrence, 92% had normal values. The mean lead-time from marker elevation to clinical diagnosis of relapse ranged from 2 to 90 months (47).

The ASCO panel, which made CEA recommendations, also reviewed breast cancer markers. The panel recommended hormone receptor assays be used in management, but concluded that present data are insufficient to recommend CA15-3 or BR27.29 for screening, diagnosis, staging, or surveillance following primary treatment. An increasing CA15-3 or BR27.29 level can detect recurrence following primary treatment, but the clinical benefit is not established. Present data are insufficient to recommend routine use of CA15-3 or BR27.29 alone for monitoring response to treatment, but in the absence of readily measurable disease, a rising CA15-3 or BR27.29 concentration may be used to suggest treatment failure (9).

The ASCO recommendations are based on the panel's view that insufficient clinical data are available to support use of serum markers as treatment indicators in breast cancer. One study contradicts the opinion and demonstrates the effectiveness of serum CEA and CA15-3 in management of breast cancer. In a series of lymph-node-positive, high-risk breast cancer patients who had a rise in either marker to the 97th percentile of normal, a rigorous conventional search for metastases was made. Women in whom metastases was not found were randomized to receive chemotherapy or no therapy. In the no-therapy group, the median time to metastases was 4 months, and the median survival 12 months. In the treatment group, the median time to metastases was 24 months and the median survival was 60 months (48).

These markers are not useful in screening because their presence in the serum is related to the extent of tumor burden and the stage of the cancer nor are they useful prognostic markers. However, tissue analysis of markers in primary surgical specimens can provide useful information and practical information concerning prognosis-total survival, disease-free survival, and risk of recurrence. In addition to the well-established estrogen and progesterone receptor assays, these tests include Cathepsin D, HER-2/neu oncogenes, epidermal growth factor and urokinase-type plasminogen activator antigens. With these markers, it may be possible on the day of biopsy or surgery to identify the subset of women with negative lymph nodes and small tumors who will have an early recurrence of cancer (49, 50 and 51).

HER-2/neu

More than a decade ago, the neu oncogene found in rats was reported to encode an epidermal growth factor receptor-related protein with a molecular weight of 185Kd (p185) (52). The human homologue of the rat neu oncogene has been referred to as C-erb-B-2 or HER-2/neu. It is closely related to the epidermal growth factor receptor (C-erb-1). The HER-2/neu gene is amplified in a variety of epithelial-cell tumors. Most attention has been paid to this amplification in breast cancer and the increases in the protein product. Immunohistochemistry studies have shown that normal epithelial cells have approximately 100,000 molecules of p185 per cell whereas in human epithelial tumors, such as breast cancer, the number of molecules may be elevated as much as 30- to 45-fold. The HER-2/neu protein is composed of the cytoplasmic domain, the transmembrane domain, and the extracellular domain. The C-terminal or cytoplasmic end of the molecule is a region containing 580 amino acids and demonstrates tyrosine kinase activity. The extracellular portion of the molecule is heavily glycosylated and constitutes the ligand-binding domain. It is shed from the cell surface by proteolytic cleavage and released into culture supernatants as well as into the circulation of both normal individuals and patients with cancers of epithelial origin. The intact native molecule found in epithelial cells has been widely reported to be an 185Kd protein. Reports now have confirmed the original report in 1989 that the extracellular domain of HER-2/neu is approximately 105Kd with a range from 97-115Kd. Breast cancer patients with HER-2/neu positive tumors (identified by immunohistochemistry) in their primary tumors have been shown to have a poor prognosis with shorter disease-free and overall survival than patients who do not express HER-2/neu. Many reports suggest that patients with HER-2/neu-positive primary tumors should be monitored for elevated serum HER-2/neu levels as a means of detecting early recurrence. In addition elevated levels of serum HER-2/neu correlate with the presence of metastatic disease and poor prognosis and may be valuable in predicting response to various forms of therapy. Combining HER-2/neu serum measurements with other tumor markers (CEA and/or CA15-3) may improve the sensitivity of detection of recurrence (53, 54 and 55).

In our studies of serum HER-2/neu, the specificity in normal women was 100% and those with benign disease 95%. In women with breast cancer, the sensitivity was 1.7% in stage I disease, 3.0% in stage II disease and 35.5% in stage V. There were 56 elevations in 285 women with breast cancer (19.7%). HER-2/neu was monitored over time in patients with metastatic breast cancer who were receiving therapy. All patients had a pretherapy serum specimen and four post-therapy specimens. The tissue HER-2/neu and hormone receptor status was known for all patients. The tissue HER-2/neu was positive in seven patients, of whom five received Herceptin (Genentech, Inc., San Francisco, California) and Taxol (Bristol-Myers Squibb). One patient received only Taxol and one only Adriamycin (Pharmacia Inc., Kalamazoo, Michigan). In six of these seven patients, the pretherapy HER-2/neu was elevated. In five of these patients, there was a rapid fall in HER-2/neu following therapy and lower levels of HER-2/neu were reflective of a response to therapy. There did not seem to be a relationship between serum HER-2/neu levels and hormone-receptor status. A combination of serum HER-2/neu, CA15-3 and CEA allowed a more precise evaluation of response to therapy. The one patient in whom the pretherapy HER-2/neu was not elevated had bone metastases and initially progressive disease. After receiving Herceptin and Taxol, the HER-2/neu remained within normal limits. The patient remained stable for several months and then experienced a recurrence and the therapy was changed to Adriamycin. The HER-2/neu remained in the normal range of five separate occasions between November 1998 and June 1999, but the CA15-3, which remained at about 45 U/mL until the clinical progression, then rose to more than 100 U/mL. The serum CEA also was within normal limits for the entire study. None of these patients

were studied sufficiently long to see if serum HER-2/neu would reflect escape from therapy and a recurrence.

In four patients, tissue HER-2/neu was negative. Only one of these patients received Herceptin and Taxol. One other patient received Taxol alone, one adriamycin, one Tamoxifin, and one a combination of cytoxan and adriamycin. In none of these patients was the preoperative HER-2/neu level elevated, nor did the marker reflect the clinical course of the patient. Of considerable interest is that in two of these patients, CEA was much more indicative of the course of the patient than either HER-2/neu or CA15-3.

There are many other studies related to the prognostic significance of tissue HER-2/neu and possible role of serum HER-2/neu. Numerous workers have shown that the expression of tissue HER-2/neu is related to poor prognosis and predicts the subset of women who will respond to Herceptin therapy. CA15-3, CEA, and HER-2/neu were compared in 200 patients with primary breast cancer who were followed up to 2.5 years. At recurrence, 32.6% of the patients had an elevation of HER-2/neu. This was the first marker elevated in 18.6 of the women whereas CA15-3 was the first indicator in 18% and CEA in 20%. Elevations were seen in 8 of 10 women who expressed HER-2/neu, but in only 1 of 30 women who did not express the oncogene in their tissue. In another study, serum HER-2/neu was measured in 50 controls, 112 patients with benign breast disease and 45 patients with no evidence of breast cancer after therapy. None of these women had elevations. Elevations were seen in 38.5% of patients with liver cirrhosis and in 26.5 of patients with primary liver cancer as well as in 7% of women with local-regional breast cancer and 41.5% of women with metastatic breast cancer. Elevations also were seen in 6 of 28 women with ovarian cancer, 3 of 14 of those with colorectal cancer. Elevations also were seen in 11.5% of patients with localized lung cancer and 16% of those with metastatic disease. Elevations were seen in 22% of men with prostate cancer and 11% of gastric cancer patients. These studies suggest the use of HER-2/neu in other cancers as well as breast cancer and the possible interference in both noncancer- and cancer-related liver disease (56).

In 35 patients receiving fractionated paclitaxel therapy, 10 had HER-2/neu levels above 15 ng/mL (the positive/negative cutoff) (54). The overall therapy-induced remission rate was 36%, but 62% in the women with normal HER-2/neu levels. Sequential changes of the serum HER-2/neu were similar to those of the breast antigen CA27.29, but the HER-2/neu seemed to be more sensitive and reflective of clinical changes. Serum HER-2/neu was measured in 300 patients with metastatic breast cancer who received aromatase inhibitors during therapy. Fifty-eight of these women had elevated serum HER-2/neu values. The clinical response rate was 40.9% in the 242 patients who had “normal” concentrations of HER-2/neu and only 20.7% in the 58 women with elevations. In a retrospective study of 200 women who had received tamoxifen, there was no statistical relationship between HER-2/neu levels and clinical response. Pretherapy HER-2/neu levels did not predict response rates to chemotherapy. Elevations were observed in 10 of 22 women with metastatic breast cancer, but the levels neither predicted response nor reflected therapeutic response (57, 58).

In a study of 42 women with primary breast cancer and 62 patients with stage IV breast cancer, elevated values of serum HER-2/neu were seen in six of the 42 primary cancer patients (14.2%) but in 92.8% of the women who expressed HER-2/neu in their tissues and in 8.2% of women who did not express the oncogene. Serum concentrations were correlated to tumor size and node involvement. In the women with stage IV disease, 27 of 62 (45.5%) had elevations. There was a correlation with CA15-3 as well as to hormone receptor status. Elevations of serum HER-2/neu were found in 32 of 94 patients (34%) with metastatic breast cancer. Low values were a powerful indicator of response to antiestrogen therapy. Three of 32 women with elevated levels responded to therapy compared to 35 of 62 with nonelevated values. In women with elevated values, there was a shorter time to progression. Regardless of other predictive factors, elevated HER-2/neu levels indicated a lack of response to hormonal therapy. Serum HER-2/neu levels before chemotherapy were found to correlate with the number of positive nodes, but there was no correlation to age, receptor, or disease status. After nodal status, the most significant variable in predicting disease-free survival was a low level of HER-2/neu (57, 58).

Cathepsin D

Cathepsin D is measured in tissue but although reported present in serum, assays in serum have not been found to be clinically useful. Breast tissue cathepsin D may be an important prognostic indicator of recurrence, disease-free, and overall survival. However, there is controversy over its use for this purpose and the conflicting reports may be a result of methodology differences. Assays have been developed for a variety of different portions of the protein molecule (50).

Total cathepsin D was studied in 738 primary tumors. High levels were associated with peritumoral vascular invasion, high-grade infiltrating duct cancers, tumors >2 cm in size, and metastases in auxiliary lymph nodes. There was low activity in *in-situ* carcinomas and there was no correlation to menopausal or hormone receptor status. With the same method, but with a cut-off of 60 pmol/mg cytosol protein, an additive prognostic value for cathepsin D with c-erb-B-2 and C-myc oncogenes were observed. There was no correlation to node invasiveness, hormone receptor, tumor grades, or the menopausal status of the patients. In a study of 267 patients, a cut-off of 31 ng/mL cytosol protein was used and it was concluded that cathepsin D is the best prognostic marker for predicting disease-free survival in women with breast cancer and that pS2 is the best marker for predicting overall survival.

Multivariate analysis was used in a study of 710 primary breast cancers. The cathepsin D values ranged from 2.8 to 531 pmol/mg cytosol protein with a mean value of 59 ± 40 pmol/mg. Cathepsin D correlated with hormone receptors and weakly to pS2. There was no relationship to age, menopausal status, tumor size, or differentiation. Higher concentrations of cathepsin D were observed in tissue from lymph-node-positive patients. There was a positive correlation with relapse-free survival ($p = 0.001$) but not with overall survival. In 155 women with cathepsin D values <30 pmol/mg cytosol protein, the survival was 63 ± 5 months; whereas in 189 women with values >70 pmol/mg cytosol protein,

it was 42 ± 4 months. Many workers agree that cathepsin D is a potentially important prognostic marker in breast cancer but its clinical application awaits clearer definition. High cytosolic cathepsin D concentrations are an indicator of poor prognosis in breast cancer. Cathepsin D concentrations are independent of other prognostic markers (59, 60 and 61).

In studies with a commercially available immunoradiometric kit for measurement of the cathepsin D, a cut-off was established at 40 nmol/mg cytosol protein in primary breast cancer tissue. The median follow-up was 48 months in women with low concentrations of cathepsin D and 45 months in women with high concentrations. Patients with high concentrations had significantly shorter disease-free interval ($P<0.05$) and overall survival ($P<0.01$). Women with high concentrations of cathepsin D and low concentrations of estrogen receptor protein had highly significant shorter disease-free intervals ($P<0.01$). Women with high concentrations of cathepsin D and low concentrations of estrogen receptor protein had highly significant shorter disease-free intervals ($P<0.001$) and overall survival ($P<0.001$) when compared to women with high concentrations of estrogen receptor protein but low enzyme concentrations.

Cathepsin B

Serum and urine cathepsin B were reported to be significantly higher in patients with distant metastases from a variety of cancers than in control noncancer patients or cancer patients without distant metastases. Successful treatment resulted in a fall in the cathepsin B to control values. In lung cancer, the expression of tissue cathepsin B was an independent factor associated with death from the disease. Survival rate was significantly poorer in these individuals. In another study in 65 normal/malignant matched lung tissue samples, the cancerous tissue had 4.6 times the cathepsin B concentration. There was no correlation to cell type, tumor stage, or cell differentiation. Increased activity was related to shorter survival. In colon-rectal cancer tissue, cathepsin B was 1.4 times higher in activity than in normal mucosa, but elevations in cancer tissue only were observed in patients with early stage disease. There was no correlation with clinical or pathology prognostic indicators. With a monospecific polyclonal antibody raised against human liver cathepsin B, it was observed that there is a significant correlation to survival in patients with advanced colon cancer, who demonstrated a low immunohistochemical expression of the enzyme. In 41 patients with advanced cancer, 21 exhibited a high expression, 17 exhibited a low expression, and three were negative. In 28 patients with early stage disease, six were negative and 17 demonstrated a low immunohistochemical expression of the enzyme. In 41 patients with advanced cancer, 21 exhibited a high expression, 17 exhibited a low expression, and three were negative. In 28 patients with early stage disease, six were negative, 17 had normal tissues, and 17 from those benign adenomas were negative (50).

Urokinase Plasminogen Activator (uPA) and Plasminogen Activator Inhibitor

uPA activator is a serine protease involved in cancer invasion and metastases. It is an independent prognostic factor in breast cancer and when observed in high concentrations in breast tissue, is an indicator of shortened disease-free survival and overall survival (49).

Enzymes as Tumor Markers

Many of the markers already discussed are enzymes (acid phophatase, prostate specific antigen, human kallikrein, urokinase plasminogen activator, cathepsin D and B). However, many other enzymes are valid markers, some of which have been used for decades and are still of clinical relevance.

Alkaline Phosphatase (ALP)

Serum alkaline phosphatase (43) is elevated in patients with primary bone cancer as well as in individuals with cancer metastatic to bone. Elevations are greater in persons with osteoblastic bone lesions than in patients with osteolytic disease. In osteoblastic disease, serum levels can be as much as 40-fold the upper reference level. Because the majority of metastatic bone lesions in breast cancer are osteolytic and in prostate cancer osteoblastic, elevations of ALP in prostate cancer usually are much higher than in breast cancer. Serum levels reflect regression and progression, but there may be a paradoxical rise in ALP during the early phase of disease regression, presumably reflecting an attempt to repair the damaged bone.

Alkaline phosphatase exists in forms that are organ related; primarily intentine, bone, liver, and placenta. These can be differentiated by specific inhibitors, differences in electrophoretic mobility, heat denaturation, and immunochemical techniques. In establishing whether an elevated ALP is related to bone, liver, or both, traditional isoenzyme measurements have not been too useful when there is minimal involvement of either bone or liver. This is primarily because of similarity and overlap between the bone- and liver-derived enzymes, which are in fact differently glycosylated forms of the same protein. However, it has been demonstrated that wheat-germ lectin preferentially reacts with the bone enzyme and permits excellent differentiation and quantitation of bone and liver alkaline phosphatases. More recently, monoclonal antibodies that can discriminate between the isoforms have been introduced.

The bone alkaline phosphatase measured immunochemically has been successfully used to monitor metastases to bone and differentiate bone verses liver-elevated serum alkaline phosphatase levels.

Lactate Dehydrogenase (LDH)

The role of LDH (62) in algorithms already has been discussed. Serum levels of total LDH have been found useful in hematologic cancers. Patients with lymphoma have been stratified based on LDH levels and remission rates are related to this level.

Neuron-Specific Enolase

Neuron-specific enolase (63) is the γγ isoenzyme of the glycolytic enzyme enolase. Enolase exists as three dimeric subunits γ, β, α, which give rise to five isoenzymes: αα, ββ, γγ, αβ, and αγ. The

γγ isomer (NSE) is the predominant form in brain. In serum, it is a specific marker for the family or neuroendocrine tumors referred to as the amine precursor uptake decarboxylase (APUD) tumors. These include neuroblastoma, medullary carcinoma of the thyroid, and small-cell carcinoma of the lung (SCCL). Elevations initially were observed in 90% of patients with neuroblastoma, primarily those with extensive disease, and in 70% of patients with SCCL. Ninety percent of individuals with extensive SCCL have NSE elevations compared with only 40% of patients with limited disease. In another study, elevations were not found in normal persons or in patients with non-small-cell lung cancer, but in 72% of patients with SCCL before treatment — 90% of patients with extensive disease, and 65% of those with limited disease. In a survey of the literature, elevations were reported in 103 of 169 (61%) patients with limited-stage disease and in 204 of 231 (88%) of those with extensive-stage disease. The authors of these various studies used cutoff values ranging from 7.5 µg/L to 20 µg/L. Neuron-specific enolase is useful in monitoring patients. Falls reflect response to successful therapy and subsequent elevations reflect an exacerbation. Longitudinal studies in 23 patients receiving chemotherapy indicated an excellent correlation between disease state and enzyme concentration. In 25 of 29 patients, regression was accompanied by elevations of NSE activity.

CONCLUSION

Part of "26 - Tumor Markers"

Circulating and tissue-tumor markers have been proposed as clinically useful in screening, diagnosis, prediction of prognosis, and patient management. In screening, the marker should merely answer the question as to whether a cancer is present or not. The use of markers in diagnosis should aid in confirmation of the cancer and also provide information on how severe or extensive the malignancy is (staging). In prognosis, the marker should assist in predicting the aggressiveness of the tumor. The most significant and accepted use of markers is in assistance in the therapeutic management of the cancer patient. In this case, there are two questions the marker can help to answer: (i), what is the prognosis before therapeutic intervention; and (ii), what is the probability of success of specific therapy.

Marker levels may suggest a need for a change in therapy or for additional therapy and provide lead-time for initiation of new or more aggressive therapy. They also may provide assistance in early assessment of clinical trials. It must be emphasized that markers may not be elevated in some patients with extensive and progressive tumors. The important point is that a positive marker may be very meaningful from a clinical point of view, but a negative marker should never create a false sense of security that tumor is not present or is not progressing. Another problem in the use of tumor markers in the United States is FDA approval. Non-FDA–approved diagnostic tests usually are not reimbursable by third-party payers.

Although tumor markers can provide information on regression or progression of disease, the question arises whether the costs of the assay, and more importantly, the cost of other diagnostic procedures triggered by positive marker results, are justifiable in the treatment of ultimately incurable disease and what level of intervention is acceptable among false positives. The acceptance or rejection of these costs is related to the individual's definition of cost effectiveness. Is cost effective the expenditure of whatever is necessary to achieve the best possible result for the patient, or is it taking the most economical course consistent with good medical practice? The nihilist will say that marker assays are not useful or necessary until absolute therapy is available. I contend that markers provide clinically useful information and whenever there is even the remotest possibility of altering the clinical course in a positive way, they should be included in the treatment plan (64).

REFERENCES

1. Kahn SN. Dear Dr. Bence Jones. *Clin Chem* 1991;37:1557–1586.

2. Schwartz MK. Enzymes in Cancer. *Clin Chem* 1973;19:10–22.

3. Schwartz MK. *Tumor Markers. Diagnostic Endocrinology and Metabolism (in-service training and continuing education program AACC)*. 1997;15:367–387.

4. Schwartz MK. Current status of tumor markers. *Scand J Clin Lab Invest* 1995;55:5–14.

5. Schwartz MK. Specialized techniques of cancer management and diagnosis: section 3, cancer markers. In: De Vita V, Hellman S, Rosenberg S, eds. *Principles and Practice of Oncology*, 4th ed. Philadelphia: JB Lippencott Company, 1993:531–541.

6. Sturgeon CM, Seth J. Why do immunoassays for tumor markers give differing results. A view from the UK national quality assessment schemes. *Eur J Clin Chem Biochem* 1996;34:755–759.

7. Ladenson JH, McDonald JM, Schwartz MK. Colorectal carcinoma and carcinoembryonic antigen (CEA). *Clin Chem* 1982;26:1213–1220.

8. Kiang DT, Greenberg LJ, Kennedy BJ. Tumor marker kinetics in the monitoring of breast cancer. *Cancer* 1990;65:193–199.

9. American Society of Clinical Oncology (ASCO) Tumor Marker Expert Panel. Clinical practice guidelines for the use of tumor markers in breast and colorectal cancer.

10. Steinberg W, Gelfand R, Anderson KK, et al. Comparison of the sensitivity and specificity of the CA19-9 and carcinoembryonic antigen assay in detecting cancer of the pancreas. *Gastroenterology* 1986;90:343–350.

11. Byrne DJ, Browning MCK, Cuschiere A. CA72-4: a new tumor marker for gastric cancer. *Br. J. Surg* 1992;77:1010–1013.

12. Schwartz MK. Screening for cancer: Is it cost effective? *Clin Chem* 1993;39:2397–2403.

13. Schwartz MK. Tumor markers in diagnosis and screening. In: Ting SW, Chen JS, Schwartz MK eds. *Human Tumor Markers*. Amsterdam: Elsevier Science Publishers, 1989:3–16.

14. Bellet DH, Wands JR, Isselbacher KJ, et al. Serum α-fetoprotein levels in human diseases: prespective from a highly specific monoclonal radioimmunoassay. *Proc Natl Acad Sci USA* 1989;81:3869–3873.

15. Oka H, Kuroka N, Kim K, et al. Prospective study of early detection of hepatocellular carcinoma in patients with cirrhosis. *Hepatology* 1991;12:680–687.

16. Nisselbaum JS, Bosl GJ, Golbey RB, et al. Changes in serum α-fetoprotein and chorionic gonadotropin in response to cancer therapy. *Annals Clin Lab Sci* 1984;14:178–188.

17. Bosl GJ, Geller NL, Cirrincione L, et al. Multivariate analysis of prognostic variables in patients with metastatic testicular cancer. *Cancer Res* 1983;43:3403–3407.

18. Toner C, Geller NL, Tan C, et al. Serum tumor marker half-life during chemotherapy allows early prediction of complete response and survival in nonseminomatous germ cell tumors. *Cancer Res* 1990;50:5904–5910.

19. Daoud E, Bodar G, Weaver C, et al. CA125 concentrations in malignant and nonmalignant disease. *Clin Chem* 1991;37:1968–1974.

20. Rubin SC, Hoskins WJ, Hakes TB, et al. CA125 levels and surgical findings in patients undergoing secondary operations for epithelial ovarian cancer. *Am J Obstet Gynecol* 1989;160:667–671.

21. Skates SJ, Xu FJ, Yu JH, et al. Toward an optimal algorithm for ovarian cancer screening with longitudinal tumor markers. *Cancer* 1997;76:2004–2010.

22. Zurawski VR Jr., Orjoseter H, Anderson A, et al. Elevated serum CA125 levels prior to diagnosis of ovarian neoplasia relevance for early detection of ovarian cancer. *Int Cancer* 1988;42:677–680.

23. Zurowski VR, Sjorall K, Schoenfeld DA, et al. Prospective evaluation of CA125 levels in a normal population. Phase 1: the specificities of single and serial determinations in testing for ovarian cancer. *Gynecol Oncol* 1990;36:299–305.

24. Hunter VJ, Daly L, Helms M, et al. The prognostic significance of CA125 half life in patients with ovarian cancer who have received primary chemotherapy after surgical cytoreduction. *Am J Obstet Gynecol* 1992;163:1164–1167.

25. Oesterling JE. Prostate specific antigen. A valuable clinical tool. *Oncology* 1991;5:107–126.

26. Furguson RA, Yo H, Kalyras M, et al. Ultrasensitive detection of prostate specific antigen by a time resolved immunofluorometric assay and the immulite immunochemiluminescent third generation assay: potential use in prostate and breast cancer. *Clin Chem* 1996;42:675–684.

27. Lange PH, Ercole CJ, Lightner DJ, et al. The value of serum prostate specific antigen determinations before and after radical prostectomy. *J Urol* 1989;141:873–879.

28. Partin AW, Oesterling JE. The clinical usefulness of prostate specific antigen: Update 1994 (Review). *J Urol* 1994;152:1358–1368.

29. Optenberg SA, Thompson TM. Economics of screening for carcinoma of the prostate. *Urol Clin North Am* 1990;17:719–737.

30. Cooner WH, Mosley BR, Rutherford CL Jr., et al. Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate-specific antigen. *J Urol* 1991;143:1146–1152.

31. Catalona WJ, Richie UP, Ahmann PR, et al. Comparison of digital rectal examination and serum prostate specific antigen: results of a multicenter clinical trial of 6630 men. *J Urol* 1994;151:1283–1290.

32. Carter HB, Morrell CH, Pearson JD, et al. Estimation of prostatic specific antigen measurement in men with and without prostate disease. *Cancer Res* 1992;52:3323–3328.

33. Stenman VH, Leinonen J, Zhang WM. Problems in the determination of prostate specific antigen. *Eur J Clin Chem Clin Biochem* 1996;34:735–740.

34. Semjonow A, Weining C, Operpenning F, et al. Pit falls in the interpretation of prostate specific antigen levels. In: Sturgeon CM, ed. *Proc UK NEQUAS Endocrinology Meeting*. London: Association of Clinical Biochemistry, 1988;146–154.

35. Moul JW. Incrased risk of prostate cancer in African-American men. *Molecular Urology* 1997;2/3:119–125.

36. Guess HA, Gormley GJ, Stover E, et al. The effect of finasteride on prostate specific antigen: review of available data. *J Urol* 1996;155:3–9.

37. Prestigiacono AF, Stamey TA. Clinical usefulness of free and complexed PSA. *Scand J Clin Lab Invest* 1995;55S221;32–34.

38. Vogl M, Muller MM, Hotl W. Clinical usefulness of percentage of free serum prostate specific antigen. *Clin Chim Acta* 1997;258:79–90.

39. Elgamal AA, Cornille FJ, Von Pappel HP, et al. Free to total prostate specific antigen ratio as a single test for determination of significant TIC prostate cancer. *J Urol* 1996;156:1042–1049.

40. Catalona WJ, Partin AW, Slawin KM, et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease. *JAMA* 1998;279:1542–1547.

41. Allard WJ, Cheli CD, Morris DC, et al. Multicenter evaluation of the performance and clinical utility in longitudinal monitoring of the Bayer Immuno 1™ complex PSA assays. *Int J Biol Markers* 1999;14:73–83.

42. Schwartz MK. Biochemical markers in staging clinically organ-confined prostate cancer. *Molecular Urology* 1997;2/3:129–132.

43. Schwartz MK. Enzyme tests in malignant disease. In: Moss DW and Roslaki SB, eds. *Enzyme Tests in Diagnosis*. New York: Oxford University Press, 1995;187–214.

44. Saedi MS, Rittenhouse HG, Finlay JA, et al. The prostate-localized kallikreins: prostate specific antigen and human kallikrein 2. *J Clin Ligand Assay* 1999;22:371–375.

45. Dnistrian AM, Schwartz MK, Greenberg EF, et al. BR27.29 as a marker in breast cancer. *J Tumor Marker Oncology* 1995;10:92–97.

46. Dnistrian AM, Schwartz MK, Greenberg EJ, et al. CA15-3 and carcinoembryonic antigen in the clinical evaluation of breast cancer. *Clin Chim Acta* 1991;200:81–93.

47. Molina R, Zanon G, Fillela X. Use of serial carcinoembryonic antigen and CA15.3 assays in detecting relapses in breast cancer patients. *Breast Cancer Res and Treatment* 1995;36:41–48.

48. Jager W, Kramer S, Palapelas V, et al. Breast cancer and clinical utility of CA15-3 and CEA. *Scand J Clin Lab Invest* 1995;55S221:87–92.

49. Anderson PA, Kjoller L, Christinseri L, et al. The Urokinase-type plasminogen activator system in cancer metastasis: a review. *Int J Cancer* 1997;72:1–22.

50. Schwartz MK. Tissue cathepsins as tumor markers. *Clin Chim Acta* 1995;237:67–68.

51. Schwartz MK, Smith C, Schwartz DC, et al. Monitoring therapy by serum HER-2/neu. *Int J Biol Markers* 2000;15:124–129.

52. Slamon DJ, Clark GM. Amplification of C-ERB-2 and aggressive breast tumors? *Science* 1988;240:1795–1798.

53. Ross JS, Fletcher JA. The HER-2/neu oncogene: prognostic factor, predictive factor and arget for therapy. *Cancer Biology* 1999;9:125–138.

54. Luftner D, Schrabel S, Possinger K. C-erb-B-2 in serum of patients receiving factionated paclitaxel chemotherapy. *Int J Biol Markers* 1999;14:55–59.

55. Molina R, Jo J, Filella X Broix J, et al. Utility of C-erb-B-2 in tissue and serum in the early diagnosis of recurrence in breast cancer patients: comparison with carcinoembryonic antigen and CA15-3. *Br J Cancer* 1996;74:1126–1131.

56. Milina R, Jo J, Filella X, et al. Serum levels of C-erb-B-2 (HER-2/neu) in patients with malignant and non-malignant diseases. *Tumor Biology* 1997;18:188–196.

57. Revillon F, Hebbar M, Boneterre J, et al. Plasma C-erb-B-2 concentrations in relation to chemotherapy in breast cancer patients. *Eur J Cancer* 1996;32:231–234.

58. Leitzel K, Teramoto Y, Konrad K, et al. Elevated serum C-erb-B-2 antigen levels and decreased response to hormone therapy of breast cancer. *J Clin Oncol* 1995;13:1129–1138.

59. Koblinski JE, Ahram M, sloane BF. Unraveling the role of proteases in cancer. *Clin Chim Acta* 2000;291:113–135.

60. Duffy MJ, Reilly D, Brouillet JP, et al. Cathepsin D concentration in breast cancer cytosols: correlation with disease-free interval and overall survival. *Clin Chem* 1992;38:2114–2116.

61. Duffy MJ. Prognostic and predictive markers in cancer. *J Clin Ligand Assay* 1999;22:331–334.

62. Schwartz MK. Lactic dehydrogenase: an old enzyme reborn as a cancer marker. *Am J Clin Path* 1991;96:441–443.

63. Jorgensen LGM, Husch FR, Skor BG, et al. Occurrence of neuron specific enolase in tumor tissue and serum in small cell lung cancer. *Br J Cancer* 1991;63:151–153.

64. Schwartz MK. Tumor markers: what is their role? *Cancer Invest* 1990;8:439–440.

Section 4
Medical Microscopy and Urinalysis

Medical Microscopy and Urinalysis - Introduction

Patrick C.J. Ward

Section Chief

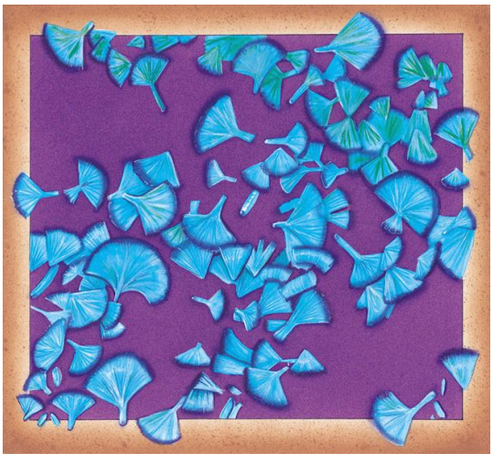


Figure.

27. Synovial, Pleural, and Peritoneal Fluids

28. Urine

Medical microscopy can be defined as the discipline of clinical laboratory medicines that deals with the examination of body fluids other than blood, using macroscopic, physiochemical, and microscopic techniques. The term “microscopy” therefore is a misnomer, because the modern clinical laboratory performs many other important nonmicroscopic examinations on specimens. However, the old terminology has remained and serves a purpose in distinguishing analyses performed on body fluids from those performed on blood.

Urine is the most frequently examined body fluid, but other fluids such as joint, pleural, peritoneal, cerebrospinal, amniotic, seminal, and ocular also receive a fair degree of attention, especially in larger laboratories. Most often, when examination of a particular fluid includes microbiological examination with culture or requires cytologic appraisal, all or a portion of the fluid sample is sent to the appropriate laboratory for that specific analysis.

The urine reagent strip, a 20th century invention, has all but replaced complicated individual chemical analyses for the determination of the presence of various bodily products in urine. Estimations of glucose albumin, hemoglobin, and bile levels, for example, as well as the determination of physical properties such as pH and osmolality, all can be accomplished using a dipstick methodology. Advances in the analysis of urine are constantly being achieved, especially in the area of immunodiagnostic testing. Abnormal products of body metabolism may be expected to be found in the various fluids of the body, and fluid, especially urine, is more easily obtained than tissue. Medical microscopy and the physiochemical examination of body fluids therefore will continue to be an important source for obtaining information to assist in diagnosis and patient care.

27
Synovial, Pleural, and Peritoneal Fluids

Jonathan L. Curry

Hesham M. Amin

Karl S. Theil

The synovial cavity is a space between diarthrodial joints, containing fluid formed by the synovial membrane. The synovial membrane covers the surface of the joint space except the contact surfaces of articular cartilage and menisci. This membrane is composed of a fibrillar interstitial matrix of collagen and glycosaminoglycans containing loosely interdigitating synovial cells, one to three cell layers thick (1). Unlike usual epithelial cells, synovial cells do not have desmosomal junctions, and there is no basement membrane between the synovial lining and subsynovial tissue. Nerves, lymphatics, and capillaries with fenestrated endothelial cells also are located in synovial tissue (Fig. 27.1).

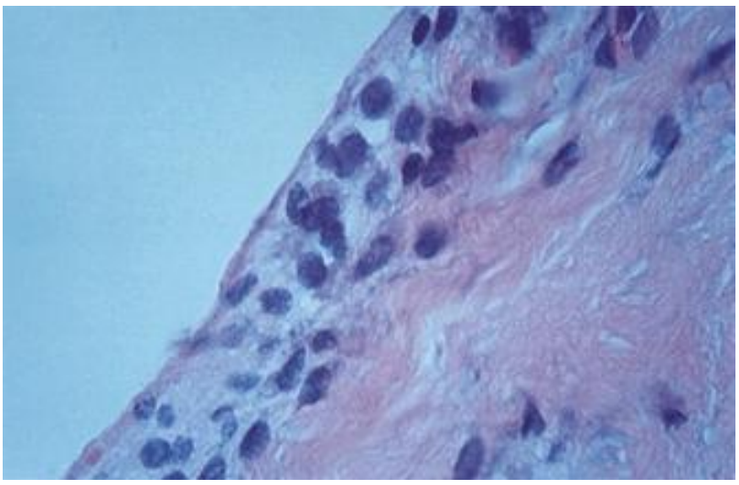


FIGURE 27.1. Synovial membrane from a normal knee joint shows joint space, synovial membrane composed of synovial cells embedded in a loose connective tissue stroma overlying dense collagen (hematoxylin and eosin).

Synovial fluid is colorless to light yellow and highly viscous. It acts as a lubricant and source of nutrients for articular cartilage in diarthrodial joints. It also is an ultrafiltrate of plasma, containing trace amounts of nonplasma proteins, lipids, hyaluronic acid, and high-molecular-weight proteins such as fibrinogen, $\alpha_2$macroglobulin, β_2-macroglobulin, and β_1-lipoprotein. Normal synovial fluid does not clot. The paucity of high-molecular-weight proteins in the ultrafiltrate may be related to endothelial permeability in synovial capillaries, charge interactions in the synovial matrix, and metabolism of proteins by synovial cells (2). Lubricin, a protein produced by synovial cells, appears to aid in lubricating cartilage-cartilage interaction (3). Hyaluronic acid, a high-molecular-weight glycosaminoglycan, is secreted by synovial cells and increases the viscosity of synovial fluid (4).

Synovial fluid analysis provides critical diagnostic information in the evaluation of patients with joint disease. There are several causes of arthritis that can be categorized broadly into infectious, crystal-induced, inflammatory, noninflammatory, and hemorrhagic. Synovial fluid analysis allows for distinction between the types of arthritis, narrowing the differential diagnosis and allowing early initiation of treatment.

Synovial fluid analysis may include macroscopic, microscopic, microbiologic, and chemical assessment. Fluid volume, color, clarity, and viscosity are observed. Total red blood cell count (RBC) and total white blood cell count (WBC) with differential are obtained. A wet preparation is viewed using compensated polarized microscopy to search for crystals or other abnormalities. Microbiological studies include Gram strain and culture. The value of additional studies, such as glucose, protein, and lactate dehydrogenase, is less certain (5).

Disease States Associated with Significant Changes in Synovial Fluid

Trauma

Acute trauma to a joint may result from a variety of causes. Hemorrhage is the most frequent consequence of trauma. Acute hemorrhage is characterized by frank blood. Evidence of previous bleeding may be inferred by finding hemosiderin-laden macrophages or erythrophagocytosis.

Septic Arthritis

Bacterial infections are the most common cause of septic arthritis. Early diagnosis of septic arthritis is important to prevent rapid destruction of the joint by proteolytic enzymes released from neutrophils. Bacteria most frequently reach the joint space by hematogenous spread across the synovial membrane. The synovial membrane is quite vascular and lacks a basement membrane and intercellular bridges; these features facilitate the spread of bacteria into the joint. Direct inoculation through trauma, contiguous wound extension, arthroscopy, or surgery is a less frequent mechanism for septic arthritis. All bacteremic patients are at risk to develop septic arthritis; however, most patients with bacteremia do not develop septic arthritis. Host interactions with the infecting organism modify the likelihood of developing a septic joint: immunologic defects, impaired defense mechanisms, trauma, and joint damage from surgery or chronic inflammation all increase the risk (6).

Neisseria gonorrhoeae is the most common cause of septic arthritis in young adults, especially in women. Gonococcal arthritis may be mimicked by disseminated meningococcal infections, so culture identification of the organism is critical. Nongonococcal septic arthritis in adults is most often a result of *Staphylococcus aureus*, followed by group A *Streptococcus* species, Gram-negative bacilli (*Escherichia coli, Salmonella* species, *Pseudomonas aeruginosa*), *Streptococcus pneumoniae*, and *Staphylococcus epidermidis*. The Gram-negative bacilli tend to occur in patients

with impaired resistance to infection, intravenous drug abuse, or underlying infections of the urinary, biliary, or intestinal tracts (7). *Haemophilus influenzae* and *Streptococcus pyogenes* are the most frequent isolates in neonates and children, respectively (7).

A variety of other organisms also can cause septic arthritis and may not be detected unless special cultures are initiated (8). Tuberculous arthritis may be the result of underlying osteomyelitis. The most common organism is Mycobacterium tuberculosis. Culture of a synovial biopsy specimen may be necessary to establish a diagnosis. Fungal arthritis may be a result of any of the deep-tissue, invasive fungi, including histoplasmosis, *Candida* species, coccidiodomycosis, blastomycosis, aspergillosis, and sporotrichosis. Actinomycosis and nocardiosis also are causes of septic arthritis. *Mycoplasma pneumoniae* has been isolated from the synovial fluid of patients with atypical pneumonia on rare occasions (9).

Viral arthritis can develop as a primary manifestation of the virus itself or as a complication related to immune-complex deposition. Viral causes of arthritis include coxsackieviruses, hepatitis B virus, lymphocytic choriomeningitis virus, mumps, parvovirus (10), rubella, and varicella (7). Diagnosis is established through serology rather than culture of synovial fluid. Although there is no evidence that human immunodeficiency virus (HIV) causes an arthritic process directly, HIV has been isolated from synovial fluid (11, 12). HIV antigen p24 has been demonstrated in synovial fluid when peripheral blood was negative for the antigen (13).

Crystal-Induced Arthritis

Monosodium urate (MSU) and calcium pyrophosphate dihydrate (CPPD) are the two most common crystals encountered in synovial fluid, and are associated with disorders termed gout and CPPD deposition disease, respectively. While joint manifestations of both disorders may be clinically indistinguishable, treatment is different. Other crystals, including calcium hydroxyapatite and calcium oxalate, usually are associated with osteoarthritis.

Gout is a multisystem disorder characterized by hyperuricemia (14). The clinical spectrum of hyperuricemia varies from asymptomatic to fullblown gout with arthritis; crystalline deposits of MSU (tophi) in and around joints, soft tissues, and renal interstitium; and nephrolithiasis. Gout is primarily a male disease, with the highest incidence in the fifth decade. Gouty arthritis usually is monoarticular at first, exquisitely painful, and characterized by granulomas containing MSU crystals within neutrophils or macrophages. Joints of the lower extremity (instep, ankle, heel, knee) are most commonly involved; wrists, fingers, and elbows also may be affected. The diagnosis of gouty arthritis is established by demonstrating MSU crystals in synovial fluid by compensated polarized microscopy. Over 95% of patients with acute gout have demonstrable crystals. Treatment is directed toward relieving the acute attack with antiinflammatory agents in the short term and lowering serum uric acid in the long term. MSU crystals can persist in synovial fluid for months in patients following treatment, even when the WBC is only slightly elevated (15).

CPPD deposition disease is most common in the elderly. The cause of crystal deposition is unknown but may be associated with age-related changes in cartilage, because many patients have preexisting joint damage. Clinical presentations vary from asymptomatic to subacute, acute, or chronic arthritis. The most common joint affected is the knee; the wrist, shoulder, elbow, ankle, and hand also may be involved. Acute attacks (pseudogout) may be precipitated by trauma, infection, joint surgery, or even walking (16). The definitive diagnosis of pseudogout is based on identification of characteristic CPPD crystals by compensated polarizing microscopy of synovial fluid. Nonsteroidal antiinflammatory drugs are the mainstay of treatment.

Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a multisystem disorder of unknown cause characterized by chronic synovial inflammation (17). Although the disease has a varied course from patient to patient, the potential for joint destruction and deformity is a major cause of morbidity. RA affects approximately 1% of the population, and women are three times more frequently affected than men. Most patients develop the disease between the ages of 35 and 50. Joints typically are affected in a symmetrical fashion, with the hands (especially proximal interphalangeal and metacarpophalangeal joints), wrists, knees, and feet most commonly involved. Signs and symptoms include joint pain aggravated by movement, and generalized stiffness greatest after periods of inactivity (“morning stiffness”).

Pathological examination of rheumatoid joints shows hypertrophy and hyperplasia of synovial fining cells, edema, vascular proliferation, and infiltration of synovium by macrophages and T lymphocytes. The release of immune mediators and chemotaxins in the synovium causes an acute inflammatory response in the synovial fluid. Proliferation of the inflammatory synovium gradually results in destruction of articular cartilage and limitation of movement as a result of pain. Clinical diagnostic criteria are well established (18).

Laboratory Evaluation

Specimen Collection

The techniques for physical examination and needle aspiration of joint fluid are outlined in textbooks of rheumatology and orthopaedics (19, 20 and 21). The sample should be collected in a sterile, plastic, disposable syringe using aseptic technique. Care should be taken to ensure that a sterile joint is not iatrogenically contaminated with bacteria from a local infection in adjacent soft tissue or from bacteremia (21). The syringe may be moistened with sodium heparin (22) (25 U/mL synovial fluid). Lithium heparin, oxalate, and powdered ethylenediaminetetraacetic acid (EDTA) should be avoided because artifacts that confound crystal analysis may result (22, 23 and 24).

The specimen is dispensed into three sterile tubes: 5 to 10 mL in a heparinized tube for microbiological studies, 5 mL in a tube containing heparin or liquid EDTA for microscopic examination, and the remainder in a nonanticoagulated tube for other studies. If lesser amounts are obtained, fluid still may be examined. Only several drops are necessary for a wet mount, leukocyte count and differential, and bacteriologic studies (25). When selecting studies on limited amounts of fluid, examinations for crystals and for microorganisms (Gram stain and culture) should receive top priority (Table 27.1).

TABLE 27.1. LABORATORY EXAMINATION OF SYNOVIAL FLUID

Macroscopic examination
Volume
Viscosity
Clarity
Color
Cell counting
Leukocyte count
Erythrocyte count
Microscopic examination
Wet preparation
Stained preparation
Polarized microscopy for crystals
Microbiological examination
Gram stain
Culture
Special studies
Chemical examination
Immunologic examination

Gross Examination

Gross examination of joint fluid is the first step in establishing a differential diagnosis of arthritis. While a precise diagnosis rarely is made from gross examination alone, a careful look at the fluid can provide valuable clues to underlying pathology.

Volume

Normal joints contain very small amounts of fluid; the normal knee joint — the largest synovial joint — contains from a few drops to as much as 4 mL. There is sufficient fluid to coat synovial surfaces and little more, thus making it difficult to sample fluid from a normal joint. Increased amounts of fluid can be seen in conditions causing edema such as myxedema, congestive heart failure, or anasarca. High-dose corticosteroid therapy may be associated with transient, asymptomatic, noninflammatory effusions (26). Difficulty in obtaining fluid from an obvious effusion may occur because of rice bodies, thick fibrin exudates, or loculations that prevent aspiration.

Viscosity

Normal joint fluid is viscous because of a high concentration of hyaluronate. An estimate of viscosity can be made by watching the fluid “string” as it is expressed from a syringe, or by suspending a drop between a glass slide and the tip of a glass rod. Normal fluid will stretch to a string of 1 to 2 inches in length, having a viscosity like egg white (Fig. 27.2). A watery consistency indicates low viscosity, and usually is associated with inflammation. Enzymes released by inflammatory cells, primarily neutrophils, degrade hyaluronate, and visibly reduce the joint fluid viscosity. Low viscosity often is caused by inflammation; however, edema or a sudden effusion after trauma can decrease viscosity by dilution in the absence of inflammation. Very viscous fluids can be seen in hypothyroidism and in fluid aspirated from ganglia. More precise quantification of viscosity generally is not necessary.

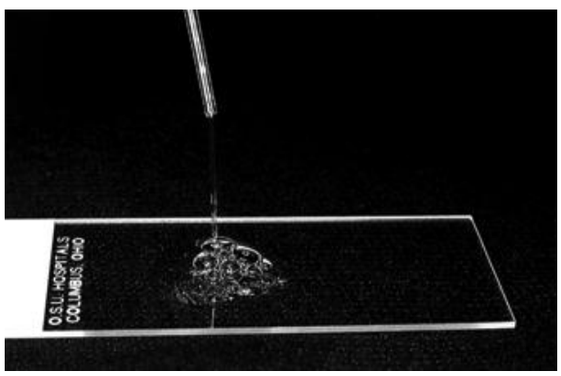


FIGURE 27.2. The viscosity of synovial fluid is demonstrated by the “string test” — the normally viscous synovial fluid forms a thin strand between the pipette tip and the surface of the slide.

The mucin clot test is a qualitative test used to estimate the degree of polymerization of hyaluronic acid. Both protein and hyaluronate are required for a mucin clot to form (27). The supernatant from a centrifuged specimen is transferred to a clean glass tube. A few drops of glacial acetic acid are layered gently on the surface of the fluid. As the acid settles to the bottom of the tube, a dense white precipitate (clot) of protein hyaluronate forms. The quality of the clot is graded as “good,” “fair,” or “poor.” A “good” clot will remain intact even when the tube is shaken; a “poor” clot resembles a few formed shreds in a turbid solution (Fig. 27.3). An alternate method is to prepare a 1:4 solution of joint fluid and 2% acetic acid. The solution is mixed and examined for the presence of a mucin clot. The quality of the clot is graded as just described. The cause of poor mucin clot formation is uncertain. The quality of the mucin clot is not solely

related to the fluid leukocyte count and the percentage of neutrophils; abnormal synovial cell metabolism resulting in production of altered forms of hyaluronate also may be a factor (28).

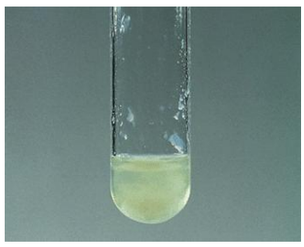


FIGURE 27.3. Poor mucin clot formation is marked by friable, poorly formed shreds.

Synovial fluid is metachromatic if stained with 0.2% aqueous toluidine blue. This feature can help identify synovial fluid if there is doubt whether the joint or soft tissue has been aspirated (dry tap). To perform the metachromatic stain test, spot a filter paper with a drop of fluid. Use a drop of normal saline as a control. After drying, add a few drops of toluidine blue. A positive reaction is marked by a color change to purple when the control spot remains blue (Fig. 27.4). This procedure may be misleading if the fluid has come in contact with heparin, because heparin also is metachromatic with toluidine blue (24).

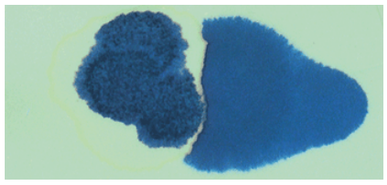


FIGURE 27.4. Metachromatic stain test illustrates metachromasia of normal synovial fluid. Toluidine blue stained synovial fluid (left) is purple. Compare with toluidine blue alone (right).

Clot Formation

Normal joint fluid does not clot because it lacks high-molecular-weight coagulation proteins, including fibrinogen, prothrombin, factors V and VII, antithrombin, and tissue thromboplastin (29). The absence of a clot can help distinguish a truly bloody effusion from a traumatic aspiration or inflammatory effusion.

Clarity

Normal joint fluid is transparent. Because erythrocytes and leukocytes may settle out in the collection tube after aspiration, it is important to ensure that the fluid is well mixed before assessing clarity. If newsprint cannot be read easily through it, the fluid is designated cloudy or opaque. Cloudy fluids may reflect an increase in cellular or protein components, and usually are a result of inflammation. Rice bodies, products of degenerating, fibrin-infiltrated synovial membrane, may appear as flecks resembling polished white rice. Flecks of dark particles resembling ground pepper may be a clue to ochronosis (30). Black or gray pigmented debris from metal or plastic fragments after prosthetic surgery has been described (31). Cholesterol crystals can give an oily quality to the fluid.

Color

Normal synovial fluid is colorless to light yellow. Grossly bloody fluid can be seen in a variety of conditions listed in Table 27.2. Streaks of blood, or the appearance of blood at the beginning of aspiration that decreases as aspiration continues, suggests a traumatic tap. Examination of the supernatant of centrifuged fluid may help distinguish true hemorrhage from a traumatic tap: if the supernatant is xanthochromic, blood probably has been present for sometime. Comparison of the hematocrit of the fluid with the venous hematocrit can help resolve the question of whether venous blood (acute bleed) or a hemorrhagic effusion is present. Noninflammatory fluids generally are yellow or straw-colored. Massive numbers of crystals can make the fluid appear opalescent. Cream-colored to off-white fluids usually contain pus with or without crystals. Various bacterial pigments in septic arthritides may discolor fluid. Milky white or chylous effusions typically are related to chronic arthritis or lymphatic obstruction secondary to filariasis, but can be seen following trauma and pancreatitis (32, 33). The presence of free-floating fat droplets suggests trauma with or without fracture. Golden cloudy fluids may contain cholesterol crystals. A gray fluid has been attributed to a retained bullet fragment (34).

TABLE 27.2. CAUSES OF HEMARTHROSIS

Trauma
Neuroarthropathy
Bleeding disorders
Malignancy (primary vs. metastatic)
Chrondrocalcinosis
Anticoagulant therapy
Joint prostheses
Thrombocytosis
Sickle cell trait
Sickle cell disease
Pigmented villonodular tenosynovitis

Cell Counting

In addition to the macroscopic examination, the leukocyte count is critically important in classifying joint fluids as noninflammatory, inflammatory, or septic. The total leukocyte count and percentage

of neutrophils are good discriminators between inflammatory and noninflammatory disease (5). Normal synovial fluids are paucicellular; most studies indicate that a WBC of less than 0.1 × 10^9/L (100/μL) is normal (35), with an upper limit of less than 0.2 × 10^9/L (200/μL) being generally accepted (Table 27.3) (24). RBCs are valuable in distinguishing frank blood from a hemorrhagic effusion.

TABLE 27.3. NORMAL JOINT FLUID WBC AND DIFFERENTIAL VALUES

Item	Mean	Range
WBC (× 10^9/liter)	0.063	0.013–0.180
Differential count (%)		
Neutrophils	7	0–25
Lymphocytes	24	0–78
Monocytes	48	0–71
Histiocytes	10	0–26
Synovial lining cells	4	0–12

From McCarty DJ. Synovial fluid. In: McCarty DJ, ed. *Arthritis and allied conditions: a textbook of rheumatology*. 11th ed. Philadelphia: Lea & Febiger, 1989: 70.

WBC, white blood cells.

The total cell count is obtained using a hemacytometer chamber viewed with a standard light or phase-contrast microscope. The count should be performed promptly to minimize the effects of cell clumping and cell death, which can decrease the total leukocyte and neutrophil counts within several hours (36). The presence of clots should be noted as a sign of spuriously low cell counts. The fluid should be well mixed to avoid cell sedimentation before counting. In contrast to other body fluids, normal saline (0.9 N) rather than acetic acid must be used as a diluent to avoid clogging of the pipette and consequent cell entrapment by a protein clot. Staining with 0.1% methylene blue dye aids in identifying leukocytes if a phase-contrast microscope is not available. Hypotonic saline (0.3 N) can be employed to lyse erythrocytes when performing a total leukocyte count. If the fluid is too viscous to permit dilution and cell counting, 0.05% hyaluronidase in phosphate buffer may be added as a diluent (24); however, leukocyte counts are said to be slightly higher in hyaluronidase-treated fluids (37).

In the absence of a hemacytometer chamber, counting leukocytes per highpower (400 x) microscopic field (HPF) can make a rough estimate of the leukocyte count in undiluted fluid. When a WBC of less than 2/HPF is noted, the chamber count frequently shows less than 1,000 cells/μL. In cases where the WBC is higher than 2/HPF, chamber counts should be done to ensure accuracy (38). Electronic counters have been used as a substitute for manual counts, but they are subject to counting errors if debris, lipid droplets, or proteinaceous material are mistaken for leukocytes (39).

Microscopic Examination

The microscopic examination includes analysis of a wet preparation, leukocyte differential count, cytologic evaluation using stained slides, and crystal examination using compensated polarized microscopy.

Wet Preparation

A wet preparation is made for immediate examination by placing a drop of fresh, well-mixed fluid onto the center of a clean glass slide. The drop of fluid is covered with a coverslip, the edges of which may be sealed with nail polish to prevent drying. Examine the center of the coverslipped area for white blood cells, red blood cells, crystals, and other materials. Be careful to avoid artifacts introduced during slide preparation, such as paper fibrils or drying artifacts near the nail polish seal. The wet preparation may reveal crystals that cannot otherwise be seen after routine fixation with alcohol and staining with Wright's stain. Examination of the sediment following centrifugation can help concentrate cellular material and crystals, especially in the clear fluids. Addition of a drop of synovial fluid to special slides coated with dry methylene blue and cresyl violet offers promise as a rapid supravital stain (40).

Light Microscopy

Air-dried Wright-stained cytocentrifuge preparations are used for cellular differential counts. In the absence of a cytocentrifuge, smears of the centrifuged sediment are acceptable, provided there is no delay in processing the sample and the smears are as thin as possible. Thick smears may have a bluish background from heparin or mucopolysaccharides in the fluid, which can prevent cytoplasmic spreading and make examination difficult (24).

Normal cellular elements found in synovial fluid (Table 27.3) include monocytes, lymphocytes, histiocytes, neutrophils, and synovial lining cells or synoviocytes (Fig. 27.5). The appearance of blood cells in Wright-stained preparations is identical to that of peripheral blood. The predominant cell type is a monocyte. Phagocytic vacuoles help identify histiocytes, but the distinction from monocytes is not always clear-cut. Occasional lymphocytes may show reactive changes, including larger nuclear size, nuclear indentations, and a small nucleolus, but the majority are small with round nuclei similar to those in peripheral blood. Synovial lining cells resemble mesothelial cells and do not have diagnostic significance. A few red blood cells are almost always present in joint effusions.

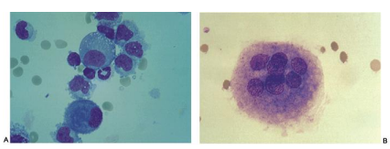


FIGURE 27.5. A: Normal cellular elements found in synovial fluid include neutrophils, lymphocytes, monocytes, histiocytes, and synovial lining cells. A few red blood cells are almost always present in joint effusions (Wright-Giemsa). **B:** Synovial lining cells may be multinucleate (Wright-Giemsa).

Abnormal cellular elements can suggest a specific diagnosis in some cases. *Erythrophagocytosis* by macrophages indicates previous hemorrhage (Fig. 27.6). *Hemosiderin* pigment and *hematoidin crystals* are breakdown products of hemoglobin, and are seen after hemorrhage. Hematoidin crystals are golden yellow, refractile, rhomboidal crystals that may be seen intracellularly or extracellularly. Cartilage cells, or their fragments, are not normally present in synovial fluid and are recognized as individual cells with deep purple cytoplasm and small nuclei surrounded by halos (Fig. 27.7). Cartilage may be seen following trauma or in osteoarthritis. Iron-laden chondrocytes can point to a diagnosis of hemochromatosis. Yellow chondrocytes are described in ochronosis (41). *Ragocytes* or *RA cells* are neutrophils containing refractile round cytoplasmic inclusions and are best seen in wet preparations using phase-contrast microscopy. These cells frequently are found in inflammatory effusions and are not diagnostic of a specific disorder. The inclusions represent phagocytosed immunoglobulin, immune complexes, DNA particles, rheumatoid factor, fibrin, and antinuclear factor (25). Lupus

erythematosus (LE) cells are neutrophils that contain a phagocytized homogenized nuclei (Fig. 27.8). While LE cells may be present in the synovial fluid of patients with systemic lupus erythematosus (SLE), even when peripheral blood LE preparations are negative (24), they are not diagnostic. LE cells also have been described in rheumatoid arthritis (42), and their absence does not exclude a diagnosis of SLE. LE cells should be distinguished from *tart cells*, which are neutrophils or macrophages containing a phagocytized nucleus that retains some chromatin detail (Fig. 27.9). Tart cells are a nonspecific finding without diagnostic significance. So-called *Reiter's cells* are macrophages containing one or more phagocytized neutrophils (Fig. 27.10). They are not diagnostic of Reiter's syndrome, and can be found in a variety of inflammatory effusions.

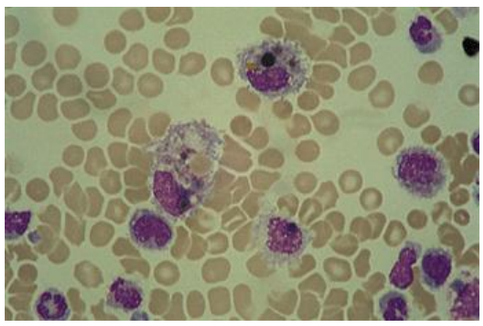


FIGURE 27.6. Indicators of previous hemorrhage in synovial fluid include erythrophagocytosis by a macrophage (center) and breakdown products of hemoglobin, including hemosiderin pigment (brown) and hematoidin crystal (yellow) in macrophage (top) (Wright-Giemsa).

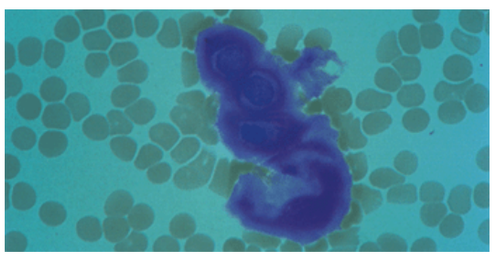


FIGURE 27.7. Cartilage is recognized by its deep purple cytoplasm and nucleus in a lacuna. Chrondrocytes also may be seen as single cells rather than the cluster shown here (Wright-Giemsa).

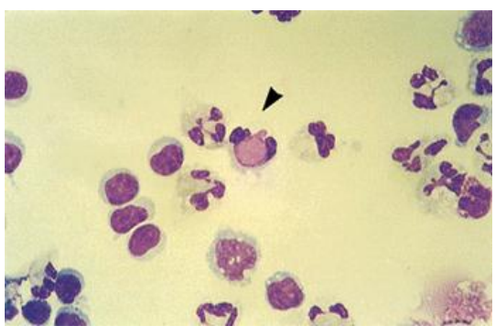


FIGURE 27.8. LE cell (arrow) is a neutrophil containing a phagocytized homogeneous nucleus (Wright-Giemsa).

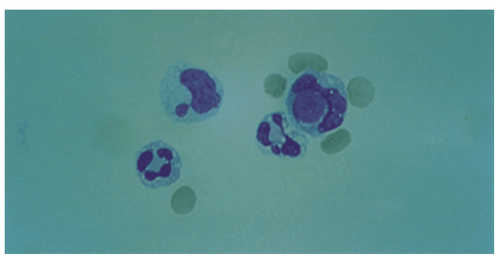


FIGURE 27.9. Tart cell is a macrophage containing a phagocytized nucleus that retains some nuclear detail (Wright-Giemsa).

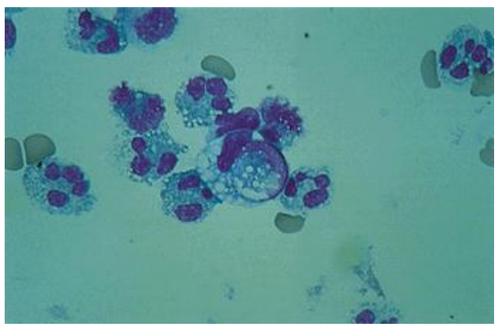


FIGURE 27.10. Rieter's cell (*center*) is a macrophage that has phagocytosed one or more neutrophils. This finding is not specific for Reiter's syndrome.

Lipid-laden macrophages (Fig. 27.11) may be seen in traumatic arthritis as a late finding (43), in chronic arthritis with chylous effusion, and in pancreatitis (33). The lipid in chronic arthritis probably originates from a breakdown of cells damaged by inflammation. Lipid may appear as extracellular or intracellular isotropic fat globules or as anisotropic lipid droplets that show “Maltese cross” birefringence under polarized light. Oil red O or Sudan black B fat stains may be used to identify lipid material.

Prominent intracellular and extracellular *lipid microspherules* or lipid liquid crystals may be observed under polarized microscopy as small Maltese crosses (44, 45). They are of uncertain origin and have been seen in pigmented villonodular synovitis (46) and chronic arthritis (47). The role of the microspherules in the pathogenesis of the arthritis remains to be determined.

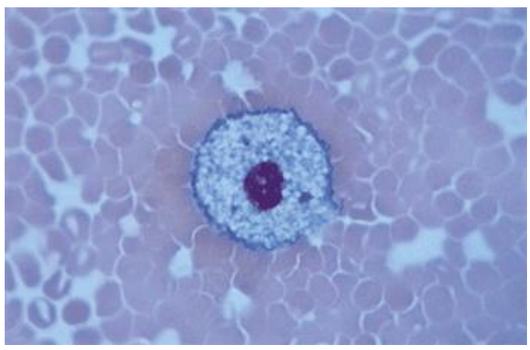


FIGURE 27.11. Lipid-laden macrophage in synovial fluid (Wright-Giemsa).

Eosinophils are not found in normal synovial fluid, and are uncommon in pathological fluids. Increased eosinophils (more than 2% of the leukocytes) are noted in a variety of disorders listed in Table 27.4 (48). *Charcot-Leyden crystals* are elongated, diamond-shaped crystals of lysolecithinase from eosinophils that stain pink or purple with Wright stain (Fig. 27.12). They have been noted in conditions associated with chronic eosinophilia, but their appearance in joint fluid is very rare. Charcot-Leyden crystals have been described in a patient with an allergic reaction following intraarticular injection of steroids (49) and also in a patient with chronic myelogenous leukemia (50). *Mast cells* (Fig. 27.13) have been identified in small numbers in synovial fluid, with variable quantities observed in patients with diverse disorders, including ankylosing spondylitis, Reiter's disease, psoriatic arthritis, and enteropathic arthritis (51). Mast cells have been observed in greatly increased numbers in patients with systemic mastocytosis (52).

TABLE 27.4. CAUSES OF SYNOVIAL FLUID EOSINOPHILIA

Rheumatic diseases
Hypereosinophilic syndromes
Bacterial infections
Allergic disease
Parasitic arthritides
Metastatic adenocarcinoma
Arthrography
Air
Dye
Therapeutic X-irradiation
Urticaria
Acute
Chronic
Idiopathy

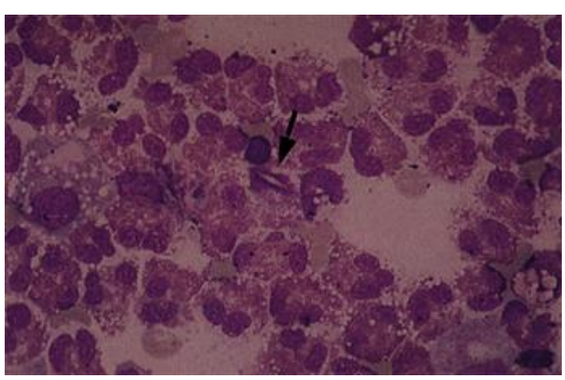


FIGURE 27.12. Diamond-shaped Charcot-Leyden crystal (arrow) within an eosinophil in eosinophilic effusion (Wright-Giemsa).

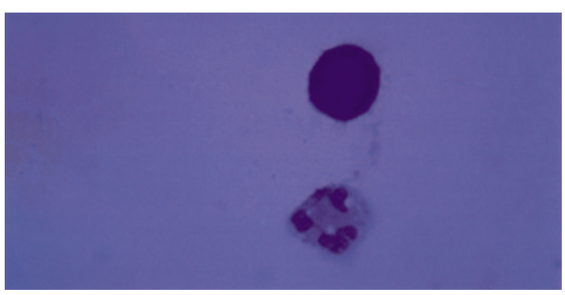


FIGURE 27.13. Mast cell in synovial fluid (Wright-Giemsa).

Sickled erythrocytes may provide a clue to an underlying sickle disease, trait, combination SC or S-thalassemia (53, 54). The

presence of sickled cells does not necessarily indicate that the hemoglobinopathy is responsible for the joint pathology. *Platelets* may be found in synovial fluid of patients with osteoarthritis and rheumatoid arthritis (55, 56). Marrow spicules are a clue to fracture with involvement of joint (57).

Although they are rare, primary and metastatic malignancies can involve the joint space, and the malignant cells can be recognized on microscopic examination. *Non-Hodgkin's lymphoma* (58), *Hodgkin's lymphoma* (59), blast transformation of chronic myeloid leukemia (60, 61), *acute lymphocytic leukemia* (62, 63 and 64), adult T-cell leukemia (62), *acute myelomonocytic leukemia*(65), *chronic lymphocytic leukemia* (66, 67), metastatic squamous carcinoma (68, 69), and *adenocarcinoma* (70, 71) have been described.

Collagen can be seen to advantage by lowering the microscope condenser. Under polarized light, collagen is weakly birefringent. *Amyloid* appears as amorphous deposits that are green and weakly birefringent when stained with Congo red and viewed under polarized light. Identification of amyloid in synovial fluid correlates well with synovial biopsy findings, so biopsy may be avoided in patients with positive fluid findings (72, 73). Rarely, fragments from metal or polymer prostheses can be seen (31, 74).

Some medications and radiographic contrast media can affect the leukocyte differential count. One study found significantly lower numbers of lymphocytes in patients with rheumatoid arthritis treated with nonsteroidal antiinflammatory drugs (NSAIDs) plus methotrexate, gold, or hydroxychloroquine than in those receiving NSAIDs alone (75). Injection of ionic contrast media with epinephrine was associated with a leukocytosis, whereas no inflammatory changes were noted after injection of nonionic contrast media (76).

Polarized Microscopy

A polarizing light microscope is equipped with two polarizing filters that have fixed axis of light transmission: the polarizer is located over the light source, and the analyzer is between the stage and the observer's eye. For crystal analysis, a first-order red compensator is placed between the polarizer and the analyzer for compensated polarized microscopy. Phase-contrast optics may be added to improve the visibility of cytoplasmic granularity.

To view crystals in a wet preparation or stained smear, the polarizing filters should have their axes of transmission oriented perpendicular (90°) to each other so that light transmitted by the polarizer will not be transmitted by the analyzer. The background of the field is pitch black when the polarizer and analyzer are oriented at 90°. If a birefringent crystal is placed between the polarizer and analyzer, polarized light entering the crystal is refracted into two rays (perpendicular to each other), nonparallel to the plane of the entering light. Refracted light shifted to the plane of the analyzer passes through and the crystal appears as a white object on a dark background. Only birefringent materials behave this way in polarized light.

Birefringence is a property of materials that have two optical axes (biaxial). The optical axis of an object is the path through which light is transmitted unrefracted. Plane polarized light entering a birefringent object out of alignment with the optical axis is refracted into two rays. The one that deviates the greatest degree from the optical axis is designated the slow ray, whereas the fast ray deviates the smallest degree (Fig. 27.14). The exiting rays, by virtue of traveling at different velocities through the object, are out of phase with each other. The difference in the degree of refraction between the fast and slow ray, along with the thickness of the object, determines the strength of birefringence. Not all crystals are birefringent, and not all birefringent objects are crystals.

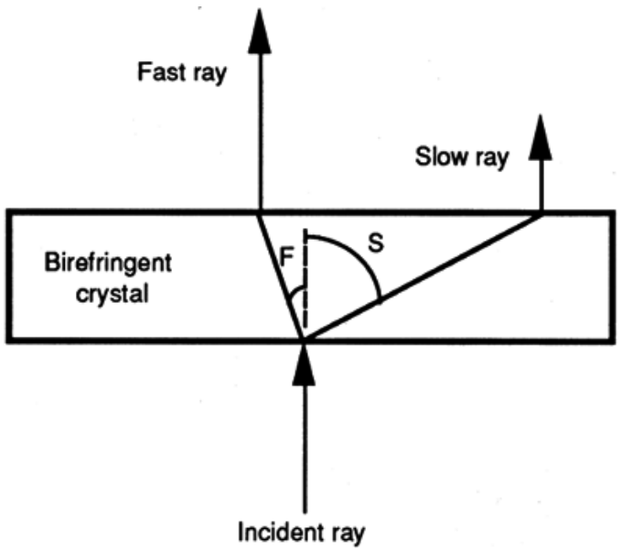


FIGURE 27.14. Light entering a birefringent crystal is refracted into fast and slow rays that are mutually perpendicular. The fast ray has the lesser degree of deviation from the optical axis of the crystal. (Modified from Gatter RA. Use of the compensated polarizing microscope. *Clin Rheum Dis* 1977;3:91–103. Used with permission.)

A compensator, depending on its birefringence and thickness, selectively retards the transmission of a particular wavelength of light. Compensators are named based on the color of light transmitted. Thus, a red compensator orientated 45° to the polarizer will selectively retard the transmission of green light (540 to 575 nm), and the transmitted light from the compensator through the analyzer appears red or rose-colored. The orientation of the slow ray (axis) of the compensator is marked for proper alignment. Most are fixed in the 45° position with respect to the polarizer and analyzer.

Crystals are classified optically as positively or negatively birefringent depending on the orientation of the fast and slow refractive rays to the optical axes of the crystal. The optical axes of biaxial crystals correspond with the physical length of the crystal,

and cross obliquely in the center of the crystal. By definition, when the fast ray of the crystal bisects the acute angle of the optical axes, the crystal is termed negatively birefringent. When the slow ray of the crystal bisects the acute angle of the optical axes, the crystal is termed positively birefringent (Fig. 27.15).

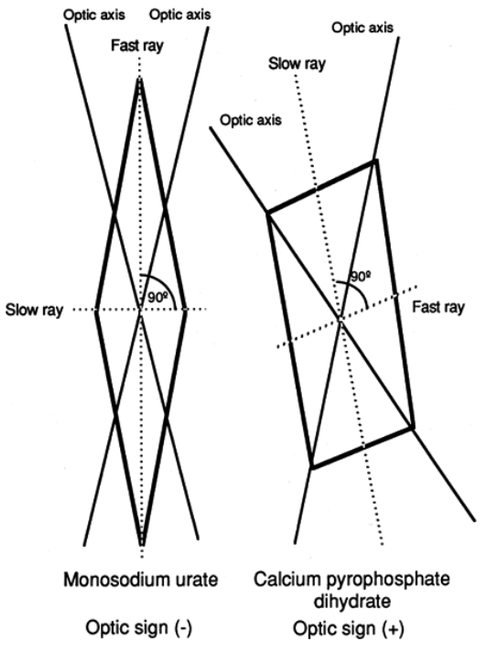


FIGURE 27.15. The optical sign of biaxial crystals is determined by the relationship of the fast and slow axes to the optical axis of the crystal. If the fast axis of the crystal bisects the acute angle of the optical axis, the crystal is optically negative. If the slow axis of the crystal bisects the acute angle of the optical axis, the crystal is optically positive. (Modified from Gatter RA. Use of the compensated polarizing microscope. *Clin Rheum Dis* 1977;3:91-103. Used with permission.)

Rays of light passing through the crystals cannot be seen; therefore, the physical orientation and color of the crystal compared with a known crystal determine birefringence. A birefringent crystal produces interference colors by adding to or subtracting from the incident ray when that ray is oriented parallel to the slow ray (axis) of the compensator. The wavelength of emitted light is shifted higher, to blue, if the slow ray of the crystal is parallel to the slow axis of the compensator, and is shifted lower, to yellow, if the fast ray of the crystal is parallel to the slow axis of the compensator. The colors are described with respect to the wavelength of the red compensator and represent the sum of the wavelengths produced by the compensator and crystal when both are placed between two polarizing filters. As seen in Figure 27.16, blue and yellow are spectrally equidistant from the red of the red compensator and reflect the additive or subtractive effect of the crystal on the red wavelength. The color of the crystal when its long dimension is parallel to the slow axis of the compensator defines the direction of the crystal's slow axis. A crystal that appears blue when its long dimension is parallel to the slow axis of the compensator (the slow ray bisects the acute angle of the optical axes) is positively birefringent. A crystal that appears yellow when its long dimension is parallel to the slow axis of the compensator, and its slow ray oriented perpendicular to the long dimension of the crystal (the fast ray bisects the acute angle of optical axes) is negatively birefringent. Several reviews provide detailed discussions of polarized microscopy (21, 77, 78 and 79).

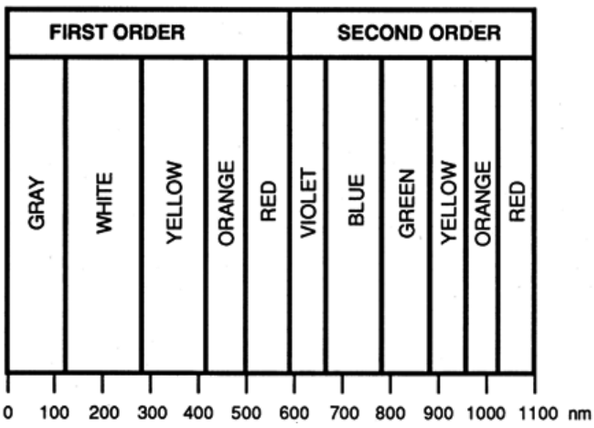


FIGURE 27.16. Orders and colors resulting from the use of compensators in a polarized light system theoretically from 0 to 1,100 nm. Yellow and blue are spectrally equidistant from the red of a red compensator. (From Gatter RA. Use of the compensated polarizing microscope. *Clin Rheum Dis* 1977;3:91-103. Used with permission.)

Weakly birefringent crystals may be difficult to see if the light is not properly adjusted; too bright a background can “overexpose” the crystals; too little light makes the already weak birefringence too weak to see. The slide should be screened under high power even if no crystals are noted under low power, with the knowledge that light microscopy alone cannot definitively exclude the presence of crystals (25, 80). Estimate the number of crystals as few, moderate, or many, and note their location as intracellular, extracellular, or both. Next, with a crystal in view, insert the red compensator. Align the long dimension of the crystal with the slow axis of the compensator, using a rotating stage or by rotating the slide, and note the color and shape of the crystal. Remember that more than one type of crystal may be present (81). Unidentifiable crystals may be subjected to x-ray diffraction, electron microscopy, or other esoteric tests when the crystals are suspected to be responsible for the joint complaint (21).

Monosodium urate (MSU) crystals are needle-shaped, 5 to 20 μm long, and strongly negatively birefringent (Fig. 27.17). MSU crystals may be located intracellularly or extracellularly (25). When the long axis of the crystal is oriented parallel to the slow axis of the red compensator, the crystal is bright yellow (Fig. 27.18 and Fig. 27.19). Rotating the axis of the compensator by 90° causes the crystal to change to blue. These crystals usually are easy to see under low power with polarized light; occasionally, microcrystals 1 to 2 μm in length may only be present requiring high-power examination for detection (82). MSU crystals can be provisionally recognized by their size and color in Wright-stained preparations in the absence of a polarizing microscope

(Fig. 27.17) (83). Rarely, spherules of MSU are present, which may be difficult to recognize as urate (84).

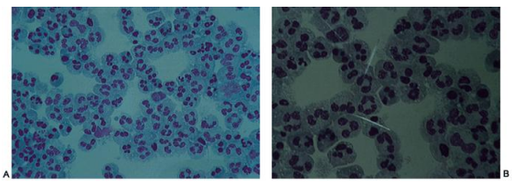


FIGURE 27.17. A: Synovial fluid with acute inflammation and monosodium urate crystals. The needle-shaped crystals are more difficult to recognize without polarized microscopy (Wright-Giemsa). **B:** Monosodium urate crystals observed with polarized light are clearly visible and brightly birefringent.

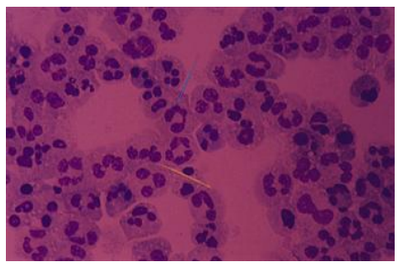


FIGURE 27.18. Monosodium urate crystals observed with compensated polarized light. The slow axis of the compensator is oriented parallel to the long axis of the yellow crystal.

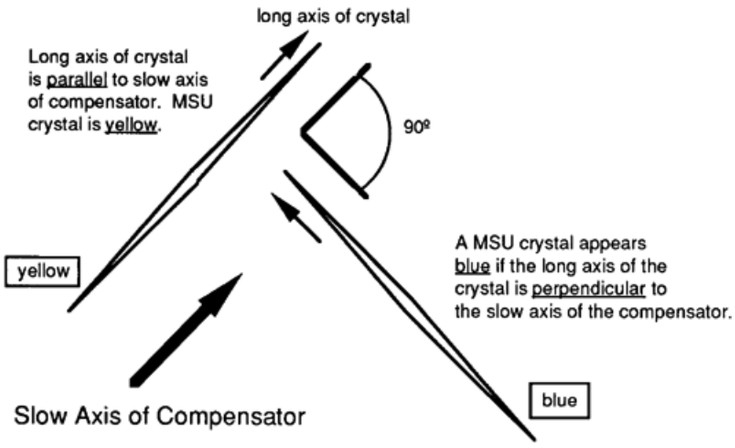


FIGURE 27.19. Polarization characteristics of monosodium urate crystals.

Calcium pyrophosphate dihydrate (CPPD) crystals are round to rhomboidal-shaped, 1 to 20 µm in length, and up to 4 µm in width. These weakly positive birefringent crystals (Fig. 27.20) often are more difficult to see than MSU crystals. CPPD crystals can be intracellular or extracellular. CPPD crystals have polarization characteristics opposite those of MSU (Fig. 27.21). When the long axis of the crystal is oriented parallel to the slow axis of the compensator, the crystal appears blue (Fig. 27.22). Rotating the axis of the compensator by 90° will cause the crystal to turn yellow. When CPPD crystals are nearly square, it can be difficult to locate the long dimension for orientation.

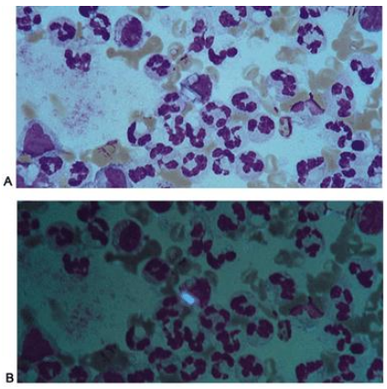


FIGURE 27.20. A: Synovial fluid with acute inflammation and calcium pyrophosphate dihydrate crystals. The rhomboidal intracellular crystals (center) are characteristic (Wright-Giemsa). **B:** Calcium pyrophosphate dehydrate crystals observed with polarized light show weak birefringence.

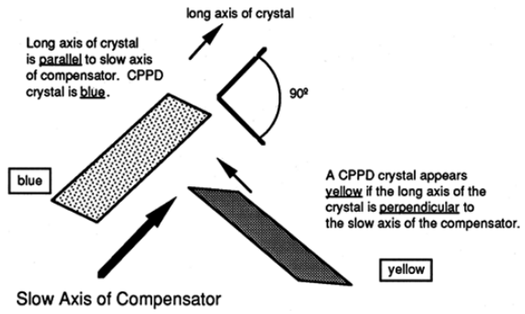


FIGURE 27.21. Polarization characteristics of calcium pyrophosphate dehydrate crystals.

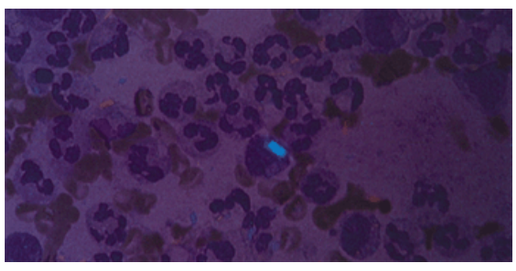


FIGURE 27.22. Calcium pyrophosphate dihydrate crystals observed with compensated polarized light. The slow axis of the compensator is oriented parallel to the long axis of the blue crystal.

Cholesterol crystals usually appear as large, flat rhomboidal plates and have notched corners. They are strongly positively birefringent, especially when stacked on top of one another. Negatively birefringent, needle-shaped forms have been described

that resemble MSU (85). Cholesterol crystals are never phagocytosed (25), and have no known etiologic role in the pathogenesis of arthritis. Both rhomboidal and needle-shaped forms can be seen in the same specimen.

Basic calcium phosphate (BCP) crystals, including hydroxyapatite and calcium phosphates, are ultramicroscopic in size and can be diagnosed as individual crystals only with transmission electron microscopy. Under the light microscope, these crystals appear as spheroidal clumps that vary from 1 to 50 µm in diameter (25). BCP crystals usually are not birefringent unless they are oriented along a common axis. Alizarin red S, a stain for calcium salts, has been used to identify clumps of crystals within cells, though its value as a screening test for these crystals has been questioned (80).

Corticosteroid crystals can mimic the morphology and polarization characteristics of MSU. Triamcinolone hexacetonide may be seen following intraarticular steroid injection therapy, and can persist in synovial fluid for months. These crystals are 10 to 20 µm in length, rectangular to needle-shaped, and negatively birefringent, and can be phagocytized by neutrophils. In contrast to MSU, they have ragged edges (25). Other steroid crystals can have different shapes and be positively birefringent (21). Steroid crystals may dissolve in alcoholic fixatives, making examination of a wet preparation mandatory for their identification.

Calcium oxalate (CO) crystals are bipyramidal, 1 to 2 µm in size, and positively birefringent. *Lithium heparin* crystals resemble CPPD, are 2 to 5 µm in length, and are positively birefringent (88). Other crystalline objects have been described in synovial fluid, including cryoprotein, immunoglobulin (86), and *aluminum phosphate* (87). Additionally, *metal fragments* from prostheses also have been reported in synovial fluid (31).

Artifacts frequently are seen during compensated polarized microscopy. These include *dust* and *dirt*, *paper fibers*, and *talc*. Careful evaluation of the optical sign of the crystal with the compensator, and comparison of polarized and conventional light microscopic appearance can help differentiate artifacts from true crystals. Artifacts may not be in the same plane of focus as cellular

material on the slide. When polarized, starch granules are brightly birefringent spherules and have a characteristic “Maltese cross” appearance (Fig. 27.23).

FIGURE 27.23. A: Starch observed under polarized light shows typical “Maltese cross” appearance. **B:** Starch viewed under polarized light with red compensator.

Chemical Examination

Chemical tests of synovial fluid generally contribute little to establishing a clinical diagnosis. Normal values for synovial fluid are shown in Table 27.5. Of the analytes, joint fluid protein, glucose, and uric acid may be requested in the usual clinical laboratory setting. Other tests, including lipids (89), enzymes, lactate, hyaluronidase, and ferritin, are more often performed in research studies. These are reviewed in several texts (21, 22, 23 and 24, 25). If viscosity presents a problem in clinical analyzers, the fluid can be sonicated or pretreated with hyaluronidase. Analyzers built to analyze serum may not perform as well with synovial fluid (90).

TABLE 27.5. NORMAL JOINT FLUID CHEMICAL VALUES[a]

Item	Range	Mean
pH	7.3–7.43	7.38
Total protein (g/dl)	1.2–3.0	1.8
Albumin (%)	56–63	60
Globulin (%)	37–44	40
Hyaluronate (g/dl)		0.3

[a] From McCarty DJ. Synovial fluid. In: McCarty DJ, ed. *Arthritis and allied conditions: a textbook of rheumatology*, 11th ed. Philadelphia: Lea & Febiger, 1989: 70.

Historically, low synovial fluid glucose has been used to differentiate infectious from inflammatory effusions. Synovial fluid and plasma specimens should be collected at the same time in tubes containing fluoride to minimize cellular metabolism of glucose *in vitro*. Joint fluid glucose equilibrates with serum levels after 6 to 8 hours; fasting levels are most reliable. In noninflammatory effusions, there is a glucose difference of less than 10 mg/dL between serum and joint fluid. With increasing inflammation, glucose levels in joint fluid fall. Infections should be considered when fluid glucose is less than 20 mg/dL (91). The large variation in glucose concentration in inflammatory and infectious disorders minimizes the usefulness of this test. Low glucose levels have been noted with CPPD in the absence of infection (92).

Measurement of synovial fluid protein does not distinguish transudates from exudates as in serous effusions. Increases in protein concentration above 2.5 g/dL are not normal, and concentrations above 4.5 g/dL often are associated with inflammation (27). Unfortunately, the protein concentration does not provide a specific diagnosis and is not of practical clinical value.

Urate is not concentrated within joints, and joint levels mirror serum levels (93). Therefore, there is little need to measure synovial fluid uric acid.

Synovial levels of pentosidine, ectonucleotide pyrophosphohydrolase, and YKL-40 are newer markers that can be measured in assessment of rheumatoid and osteoid arthritis and disease activity; however, their clinical utility has yet to be fully exploited (98, 99 and 100).

Immunologic Examination

There are few immunologic tests performed on synovial fluid that are helpful in establishing a clinical diagnosis. Many studies pertain to rheumatoid arthritis. Normal synovial fluid *immunoglobulin* concentrations are about 10% that of serum. Synovial fluid immunoglobulins in rheumatoid arthritis typically are increased and are derived both from peripheral blood and also local synthesis by plasma cells in the synovium. IgG, IgA, and IgM all are produced as part of the inflammatory response, and at least a portion of the IgM and IgG has *rheumatoid factor* activity. In most cases, the synovial rheumatoid factor titer is similar to the serum titer, but cases have been reported in which serum is negative in the presence of a positive synovial titer. Rheumatoid factors are unique to rheumatoid arthritis, and increased disease activity in RA may correlate with increased synovial fluid levels of pentosidine (25, 98).

Measurements of total hemolytic complement activity (CH_{50}) in synovial fluid need to be correlated with serum levels. The CH_{50} level is low in normal joints, and parallels serum CH_{50} in diseased joints with altered protein permeability. In the absence of local consumption, the CH_{50} is 33% to 50% of the serum value (25). A significant decrease in synovial fluid complement compared with peripheral blood levels correlates with activation by local immune complexes, and is a usual finding in RA. However, similar patterns may be seen in septic arthritis and gout (94). Decreases in both serum and synovial complement may be found in systemic lupus erythematosus. In general, because the levels do not permit separation of similar entities, complement levels are not of diagnostic assistance (95).

Microbiological Examination

Microbiological examination includes Gram stain of fluid smears and culture for microorganisms. While the Gram stain provides an immediate diagnosis when bacteria are visualized, its sensitivity is inadequate to detect all bacterial infections. Negative smears are not uncommon in gonococcal arthritis and in patients who have been treated with antibiotics before arthrocentesis (7). Thus, cultures should be initiated whenever a joint infection is a possibility.

Rapid analysis of fungal or bacterial antigens by latex agglutination or counterimmunoelectrophoresis, examination of metabolites by gas chromatography, and *Limulus* lysate assay for bacterial endotoxin are all applicable to synovial fluid samples.

Clinical Correlation

Based on the gross examination, WBC, and differential, most fluids can be classified into one of several categories (Table 27.6) Examples of disorders representative of each category are shown in Table 27.7. Synovial fluid findings often are nonspecific, and interpretation must be done in the appropriate clinical context. Algorithms for evaluation of synovial fluid microscopy have been proposed and evaluated (97).

TABLE 27.6. CLASSIFICATION OF SYNOVIAL EFFUSIONS

Gross Examination	Normal	Noninflammatory (Group I)	Inflammatory (Group II)	Septic (Group III)	Crystal (Group IV)	Hemorrhagic (Group V)
Volume (mL) (knee)	<3.5	Often >3.5	Often >3.5	Often >3.5	Often >3.5	Often >3.5
Viscosity	High	High	Low	Variable	Variable	Variable
Color	Colorless to straw	Straw to yellow	Yellow Cloudy	Yellow-white Cloudy	Yellow Cloudy	Red Xanthochromic
Routine laboratory examination						
WBC (mm^3)	<200	200–2,000	2,000–75,000	Often 100,000	2,000 to 75,000	50–10,000
PMN leukocytes (%)	<25	<25	>50 often	>75[b]	>50 often	<50
Crystals present	No	No	No	No	Yes	No
Culture	Negative	Negative	Negative	Often positive	Negative	Negative
Mucin clot	Firm	Firm	Friable	Friable	Friable	
Glucose (AM fasting)	Nearly equal to blood	Nearly equal to blood	<50 mg% lower to blood	>50 mg% lower to blood	>50 mg% lower to blood	Nearly equal to blood

Adapted from Schumacher HR. Synovial fluid analysis and synovial biopsy. In: Kelley WN, Harris ED, Ruddy S, Sledge CB, eds. *Textbook of rheumatology*. 3rd ed. Philadelphia: WB Saunders, 1989:638, and Kjeldsberg CR, Knight JA. *Body fluids: laboratory examination of amniotic, cerebrospinal, seminal serous, and synovial fluids*. 2nd ed. Chicago: ASCP Press, 1986:134.

WBC and %PMN leukocytes will be less if organism is less virulent or partially treated.

WBC, white blood cells; PMN, polymorphonuclear neutrophils.

TABLE 27.7. DIFFERENTIAL DIAGNOSIS BY JOINT FLUID GROUPS

Group I Noninflammatory	Group II Inflammatory	Group III Septic	Group IV Crystal-induced	Group V Hemorrhagic
Osteoarthrosis Traumatic arthritis Osteochondritis dissecans Osteochondromatosis Neuropathic osteoarthropathy Pigmented villonodular tenosynovitis	Rheumatoid arthritis Lupus erythematosus Reiter's syndrome Ankylosing spondylitis Regional eneritis Ulcerative colitis Psoriasis	Bacterial Mycobacterial Fungal	Gout CPPD crystal deposition disease Apatite-associated arthropathy	Traumatic arthritis Hemophiliac arthropathy Anticoagulation Pigmented villonodular tenosynovitis Neuropathic osteoarthropathy Synovial hemangioma

From Kjeldsberg CR, Knight JA. *Body fluids; laboratory examination of amniotic, cerebrospinal, seminal, serous, and synovial fluids*, 2nd ed. Chicago: ASCP Press, 1986:133

CPPD, calcium pyrophosphate dihydrate.

The highest WBCs and neutrophils count is found in septic arthritis. WBCs typically are in the 50,000/μL range, but vary from 150/μL to greater than 100,000/μL (27, 96). Neutrophils generally account for more than 90% of the differential count. Previous treatment or diagnostic procedures may modify the WBCs and percentage of neutrophils. Finding more than 80% neutrophils in a joint fluid should prompt a search for bacteria no matter what the WBC. Definitive diagnosis is established when an organism is cultured from the fluid or synovium. When organisms are identified on Gram stain, a presumptive diagnosis can be made. However, a negative Gram stain does not exclude infection. Gram-stained smears are positive in 75% of patients with staphylococcal infections and in 50% of those with Gram-negative bacilli. However, less than 25% of patients with gonococcal arthritis will have organisms on gram stain (6).

Cell counts and percentage of neutrophils in joint fluid in

septic arthritis overlap with those of crystal induced arthritis. WBCs in crystal-related arthritis range from less than 100/µL to greater than 100,000/µL with approximately 80% neutrophils (range: 2 to 100) (27). Identification of the characteristic crystal will establish a diagnosis. Both CPPD and MSU can occur in the same joint.

Inflammatory, noncrystal, nonseptic arthritis is characterized by more than 50% neutrophils and WBCs ranging from 2,000/µL to greater than 100,000/µL. This group is largely diagnosed by the exclusion of crystals and organisms. A specific diagnosis is not always possible from joint fluid analysis alone.

A WBC of less than 2,000/µL distinguishes noninflammatory from inflammatory disorders. This number is clearly not absolute, because there is much overlap between diagnostic categories. Most noninflammatory fluids will have variable numbers of neutrophils, lymphocytes, monocytes, and synovial cells. Fluid analysis alone usually will not provide a specific diagnosis.

Macroscopic examination of the fluid may help distinguish traumatic aspiration from hemarthrosis. A xanthochromic or dark red-brown supernatant after centrifugation of a bloody fluid favors hemarthrosis. Comparison of WBC/RBC ratios in fluid and peripheral blood also may be helpful.

PLEURAL FLUID

Part of "27 - Synovial, Pleural, and Peritoneal Fluids"

The pleural fluid is normally present in small amounts (1–10 mL) in the potential space between the visceral and parietal layers of the pleura. Several factors affect the production of pleural fluid from the capillaries of the parietal pleura. These factors include plasma oncotic pressure, capillary hydrostatic pressure, and capillary permeability. The pleural fluid is drained via vascular channels in the visceral pleura. Increased production or decreased removal will lead to accumulation of the pleural fluid (pleural effusion), which sometimes has significant clinical consequences.

Pleural effusions generally are classified into transudates or exudates. Examples of causes of transudative pleural effusions include congestive heart failure, nephrotic syndrome, and liver cirrhosis. Iatrogenic transudates may occur with misplacement of central venous lines into the pleural space. Exudative effusions usually are products of infectious or malignant processes.

Specimen

Specimens should be collected in heparinized containers (1,000 units of heparin/100 mL fluid). If cell counts are required, an EDTA tube should be added. Fresh specimens may be stored up to 48 hours in the refrigerator before processing. Optimal requirements for pH measurements include collection under anaerobic conditions in heparin, on ice. Exposure of the specimen to air causes elimination of carbon dioxide and, subsequently, a false increase in pH.

Laboratory Evaluation

The aim of evaluating the pleural fluid in the laboratory is to help in establishing a diagnosis. This generally can be achieved by differentiating transudative from exudative effusions, identifying infectious agents, and recognizing malignant involvement.

Transudates usually are clear and odorless, while exudates generally are turbid or cloudy. Infections also lead to characteristic odors of exudative pleural fluids. According to Light et al., pleural fluid is considered an exudate if at least one of the following criteria are established: pleural fluid protein greater than 0.5 of plasma protein, pleural fluid lactate dehydrogenase (LDH) greater than 0.6 of plasma LDH, or pleural fluid LDH greater than two thirds of the plasma upper limit of normal (>200 U/mL) (101). These criteria, however, were arbitrarily established, and pleural fluids with borderline values should be considered exudates.

Bloody pleural effusions usually result from trauma, malignancy, or infarction. A traumatic tap, however, should be excluded. Bloody or turbid fluids should be centrifuged. If the supernatant is still turbid, a chylous or pseudochylous effusion is present. Clear supernatant indicates that turbidity was a result of the presence of cellular elements.

Microscopic evaluation of the pleural fluid includes a cell count. Smears are prepared and stained with Romanowsky's stain for a differential count (102). Centrifugation or filtration should be performed before preparing the smears to warrant optimum cellular yield. Normal and most of the transudative pleural fluids typically contain less than 1,000 cells/µL of which less than 25% are neutrophils, and no red blood cells. The presence of more than 100,000 red blood cells/µL implicates an etiology of trauma, malignancy, or pulmonary infarct. White blood cell count should be interpreted with caution. Pleural fluids generally are considered exudative when more than 1,000 white blood cells/µL are present. This conclusion, however, should be examined carefully as exudates may include smaller numbers of white blood cells.

The different types of cells present in the pleural fluid should be identified, as the presence or increase of certain cellular population usually is associated with specific disease processes. For example, neutrophils predominate in acute inflammations such as bacterial pneumonia (parapneumonic effusions), lymphocytes in viral infection and tuberculosis, and eosinophils in hypersensitivity and drug reactions. It also is important to recognize the mesothelial cells, which are commonly present in inflammatory processes. Reactive mesothelial cells must be differentiated from malignant cells. Such a process sometimes necessitates the use of a panel of immunohistochemical stains (24). The two most common malignant tumors causing pleural effusion are lung and breast. If malignant effusion is suspected, cellblocks also must be prepared.

Lymphocytes may prevail in pleural effusions from patients with lymphoproliferative disorders. Morphologic differentiation of reactive and malignant lymphocytes can be significantly difficult and other ancillary studies must be undertaken, including flow cytometric analysis and immunohistochemical staining. The presence of blasts in the pleural fluid also may necessitate similar studies to establish and confirm the diagnosis.

Chemical and immunologic analyses of pleural fluids include measurements of pH, protein, triglycerides, cholesterol, glucose, lactate, enzymes, tumor markers, rheumatoid factor, and antinuclear antibody titer. The change in the normal, established levels

of some of these analytes sometimes is helpful to establish a diagnosis. For example, pH decreases in parapneumonic effusion (103). A decrease in glucose concentration (<60 mg/dL) is associated with purulent, parapneumonic effusions. Also, low levels of glucose may confirm a diagnosis of rheumatoid pleuritis. An increase in amylase of more than 1.5 times its serum levels occurs in pancreatitis and its complications (104). Levels of LDH are expected to increase in inflammatory conditions. Carcinoembryonic antigen in the pleural fluid increases with pulmonary adenocarcinoma (24). Mesotheliomas, on the other side, were found not to cause such an increase.

Efficient microbiologic evaluation of pleural fluids requires aerobic, anaerobic, and mycobacterial culturing. Gram's and acid fast smears often are desirable. *Staphylococcus aureus*, gram-negative bacilli, and anaerobic organisms are commonly identified in parapneumonic effusions. Diagnosis of tuberculous pleural effusion is important because tuberculous effusions tend to resolve spontaneously over a few weeks. Thereafter, patients frequently present with active pulmonary or extrapulmonary tuberculosis. The sensitivity, however, of pleural fluid culture for mycobacteria is only 30% (105). Therefore, a pleural biopsy is considered an essential diagnostic tool. In fact, histopathologic evaluation of the pleural biopsy is diagnostic in almost 80% of cases and cultures of the specimen are positive in 85% of the cases. Combining histopathologic evaluation and culture of pleural biopsy increases the sensitivity to more than 90%.

PERITONEAL FLUID

Part of "27 - Synovial, Pleural, and Peritoneal Fluids"

The space between the parietal and visceral peritoneal membranes normally is occupied by 10 mL of clear, pale yellow, serous fluid that provides lubrication. Ascites develops from a disruption of the homeostatic forces that control the movement of fluid across the peritoneal membranes and represents an excess state of total-body sodium and water. Ascitic fluid is an ultrafiltrate of plasma. Accumulations are categorized into transudates or exudates. Transudates are the result of increased capillary hydrostatic pressure or decreased plasma oncotic pressure and are commonly seen in congestive heart failure, hepatic cirrhosis, and hypoproteinemia. In contrast, exudates arise from conditions that cause increased capillary permeability or decreased lymphatic resorption, including infections, malignancies, trauma, and pancreatitis.

Specimen

A diagnostic paracentesis contains 50–100 mL of peritoneal fluid (106). Total RBC and WBC with differential, cellular, and microbiological examination and chemical analysis are conducted routinely (107). Peritoneal lavage (PL) is performed if an inadequate amount of fluid is isolated and abdominal trauma, acute pancreatitis, or acute peritonitis is suspected. The specimen should be aliquoted at the time of paracentesis or PL into several EDTA tubes for cell counts, differential, and cellular evaluation, and into heparinized containers for chemical analysis. Fluid should be placed into sterile containers for microbiological assessment. A sample should be collected in Roswell Park Memorial Institute medium (RPMI) for flow cytometric analysis if a hematopoietic neoplasm is suspected. The samples should be transported on ice and analyzed immediately; however, they may be stored in 4°C for 24 hours (108). When clinically indicated, blood specimens should be collected in serum and EDTA tubes for comparative analyses.

Laboratory Evaluation

Peritoneal fluid should appear pale, clear, and not exceed 100 mL in volume (107). Bloody, cloudy, greenish, or milky fluids all are abnormal and may be seen in traumatic, infectious, neoplastic, inflammatory, or chylous processes. Blood caused by a traumatic tap generally diminishes during fluid collection. Clear fluids that exceed 100 mL in volume also are abnormal.

Chemistry

Ascitic fluid is differentiated into transudate or exudate. Transudate is clear, colorless, and low in protein. There are few cellular elements and the specific gravity is low. Exudate, a protein-rich fluid, is composed of more cells and solid material. Exudate may clot if fibrinogen leaked through the vasculature. Depending on the pathogenic process, exudates can be categorized further into catarrhal, fibrinous, hemorrhagic, purulent, or serous. Traditionally, total protein, ascites/serum total protein ratio, ascites lactic dehydrogenase (LD) concentration, or ascites/serum LDH ratio should be measured to differentiate transudate from exudate. The diagnostic accuracy of traditional laboratory tests used to differentiate exudate from transudate has been challenged. Serum/ascites albumin gradient (SAAG), an alternative biochemical diagnostic criterion, is more accurate. SAAG accurately differentiated transudate from exudate in 98% of tested samples, compared to 52%–80% when tested by traditional methods (109). An ascitic SAAG >1.1 g/dL is characteristic of a transudate. In contrast, a SAAG <1.1 g/dL is indicative of an exudate (106).

The presence of any quantity of amylase in ascitic fluid is abnormal and reflects an increase in serum levels (107). Pancreatitis or a gastrointestinal tract defect (perforated peptic ulcer or necrotic bowel) may cause amylase to appear in ascitic fluid. Samples that are lipemic or collected in oxalate or citrate containers may exhibit falsely low amylase results (110).

In malignant effusions, levels of peritoneal fluid LD are greater than those of serum. However, LD activity in peritoneal fluid is less than its activity in plasma activity (107).

Lactate, ammonia, glucose, and lipid analyses lack sensitivity and specificity and should be correlated with additional tests.

Tuberculous peritonitis may be differentiated from sterile cirrhotic ascites by measurement of ascitic fluid pH, lactate, gamma interferon (γ-IFN), and adenosine deaminase activity (ADA) (111, 112). γ-IFN values of 3.2 U/mL detected tuberculous peritonitis with a sensitivity of 93% and a specificity of 98%. ADA activity showed a sensitivity and specificity of 93% at values > 30 U/L in tuberculous peritonitis (112); however, the sensitivity decreased to 30% in cirrhotic patients with tuberculous peritonitis in the United States where tuberculosis is uncommon and cirrhosis is common (113).

The sensitivity of tumor marker tests is low making them

most useful for the monitoring of oncologic disease. Increased carcinoembryonic antigen (CEA) values are seen in many primary gastrointestinal, lung, and breast cancers. Gynecologic malignancies (ovary, fallopian tube, or endometrial carcinomas) commonly have increased levels of CA 125. The CA 125 is most useful in identifying recurrence of a primary ovarian tumor (114).

A multivariate analysis model, using five measurements [total protein, LDH, tumor necrosis factor-α (TNF-α), C4, and haptoglobin], currently is being developed to discriminate between benign and malignant ascites. Cross validation showed that 70% of unknown cases were classified correctly with this model (115).

Microscopic Examination

The total cell count of normal peritoneal fluid is <500 cells/L. Mesothelial cells, red blood cells, and white blood cells, including monocytic phagocytes, polymorphonuclear leukocytes, and lymphocytes can be found in normal fluid. The percentage of polymorphonuclear leukocyte in normal peritoneal fluid should not exceed 25% of the total cell count (116).

Total erythrocyte count >100,000 cells/μL is virtually diagnostic of hemoperitoneum (117). The total leukocyte count in ascitic fluid varies with the disease state. A polymorphonuclear leukocyte count greater than 250 cells/μL is indicative of spontaneous bacterial peritonitis (118).

Cytospin slides of malignant ascites may demonstrate a monomorphous population of abnormal cells with an increased nuclear to cytoplasmic ratio, and hyperchromatic nuclei with prominent nucleoli.

Microbiologic Examination

Cloudy or blood stained fluids, or those exhibiting increased numbers of polymorphonuclear cells, should be stained for microorganisms. The sensitivity of gram- or acid-fast bacillus stains is low, and culture of ascitic fluid is routinely performed. Direct inoculation of ascitic fluid into blood culture bottles enhances identification of bacterial organisms by 50% compared to routine cultures (119).

ACKNOWLEDGMENTS

Part of "27 - Synovial, Pleural, and Peritoneal Fluids"

This chapter was based on the original version, “Body Fluid Analysis,” written by Dr. Karl S. Theil and published in the first edition of this book.

REFERENCES

1. Simkin PA. Physiology of normal and abnormal synovium. *Semin Arthritis Rheum* 1991;21:179–183.

2. Weinberger A, Simkin PA. Plasma proteins in synovial fluids of normal human joints. *Semin Arthritis Rheum* 1989;19:66–76.

3. Swann DA. Structure and function of lubricin, the glycoprotein responsible for the boundary lubrication of articular cartilage. In: Franchimont P, ed. *Articular Synovium: Anatomy, Physiology, Pathology, Pharmacology, and Therapy*. Basel: Karger, 1982:45–58.

4. Balazs EA, Watson D, Duff ID, et al. Hyaluronic acid in synovial fluid. I. Molecular parameters of hyaluronic acid in normal and arthritic human fluids. *Arthritis Rheum* 1967;10:357–376.

5. Shmerling RH, Delbanco TL, Tosteson ANA, Trentham DE. Synovial fluid tests: what should be ordered? *JAMA* 1990;264: 1009–1014.

6. Goldenberg DL, Reed JI. Bacterial arthritis. *N Engl J Med* 1985;312:764–771.

7. Ayers LW, Fritsche TR, Lancz G, Pindur A, Specter S. Skin, wound, and tissue specimens. In: Howanitz JH, Howanitz PJ, eds. *Laboratory Medicine: Test Selection and Interpretation*. New York: Churchill Livingstone, 1991:765–791.

8. Freeman R, Jones MR. Microbiology. *Clin Rheum Dis* 1983;9:3–26.

9. David CP, Cochran S, Lisse J, et al. Isolation of *Mycoplasma pneumoniae* from synovial fluid samples in a patient with pneumonia and polyarthritis. *Arch Intern Med* 1988;148:969–970.

10. Dijkmans BAC, Van Elsacker-Niele AMW, Salimans MMM, et al. Human parovirus B19 DNA in synovial fluid. *Arthritis Rheum* 1988;31:279–281.

11. Withrington RH, Cornes P, Harris JRW, et al. Isolation of human immunodeficiency virus from synovial fluid of patients with reactive arthritis. *Br Med J* 1987;294:484.

12. Forster SM, Seifert HJ, Keat AC, et al. Inflammatory joint disease and human immunodeficiency virus infection. *Br Med J* 1988;296: 1625–1627.

13. Madariaga L, Aramburu JM, Suarez MD, et al. Presence of HIV antigen p24 in the synovial fluid of a patient without antigenemia suffering from staphylococcal arthritis. *AIDS* 1991;5:337–348.

14. Kelley WN, Palella TD. Gout and other disorders of purine metabolism. In: Wilson JD, Braunwald E, Isselbacher KJ, et al., eds. *Principles of Internal Medicine*. 12th ed. New York: McGraw-Hill, 1991:1834–1843.

15. Pascual E. Persistence of monosodium urate crystals and low-grade inflammation in the synovial fluid of patients with untreated gout. *Arthritis Rheum* 1991;34:141–145.

16. Hoffman GS. Arthritis due to deposition of calcium crystals. In: Wilson JD, Braunwald E, Isselbacher KJ, et al., eds. *Principles of Internal Medicine*. 12th ed. New York: McGraw-Hill, 1991:1479–1482.

17. Lipsky PE. Rheumatoid arthritis. In: Wilson JD, Braunwald E, Isslebacher KJ, et al., eds. *Principles of Internal Medicine*. 12th ed. New York: McGraw-Hill, 1991:1437–1443.

18. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–324.

19. Owen DS. Aspiration and injection of joints and soft tissues. In: Kelley WN, Harris ED, Ruddy S, et al., eds. *Textbook of Rheumatology*. 3rd ed. Philadelphia: WB Saunders, 1989:621–636.

20. Samuelson CO, Cannon GW, Ward JR. Arthrocentesis. *J Fam Pract* 1985;20:179–184.

21. Gaffer RA, Schumacher HR. *A Practical Handbook of Joint Fluid Analysis*. 2nd ed. Philadelphia: Lea & Febiger, 1991:1–122.

22. Naib AM. Cytology of synovial fluids. *Acta Cytol* 1973;17: 299–309.

23. Schumacher HR. Intracellular crystals in synovial fluid anticoagulated with oxalate. *N Engl J Med* 1966;274:1372–1373.

24. Kjeldsberg CR, Knight JA. *Body Fluids: Laboratory Examination of Amniotic, Cerebrospinal, Seminal, Serous, and Synovial Fluids*. 2nd ed. Chicago: ASCP Press, 1986:129–165.

25. McCarty DJ. Synovial fluid. In: McCarty DJ, ed. *Arthritis and Allied Conditions: A Textbook of Rheumatology*. Philadelphia: Lea & Febiger, 1989:69–90.

26. Lally EV. High-dose corticosteroid therapy: association with noninflammatory synovial effusions. *Arthritis Rheum* 1983;26:1283–1287.

27. Cohen AS, Goldenberg D. Synovial fluid. In: Cohen AS, ed. *Laboratory Diagnostic Procedures in the Rheumatic Diseases*. 3rd ed. New York: Grune & Stratton, 1985:1–54.

28. Hogan DB, Pritzker KPH. Synovial fluid analysis-another look at the mucin clot test. *J Rheumatol* 1985;12:242–244.

29. Cho NH, Neuhaus OW. Absence of blood clotting substances from synovial fluid. *Thromb Diath Haemorrh* 1960;5:108–111.

30. Hunter T, Gordon DA, Ogryzlo MA. The ground pepper sign of synovial fluid: a new diagnostic feature of ochronosis. *J Rheum* 1974;1:45–53.

31. Kitridou R, Schumacher HR, Sbabaro JL, et al. Recurrent hemarthrosis after prosthetic knee arthroplasty: identification of metal particles in the synovial fluid. *Arthritis Rheum* 1969;12:520–528.

32. Reginato AJ, Feldman E, Rabinowitz JL. Traumatic chylous knee effusion. *Ann Rheum Dis* 1985;44:793–797.

33. Simkin PA, Brunzell JD, Wisner D, et al. Free fatty acids in the pancreatic arthritis syndrome. *Arthritis Rheum* 1983;26:127–132.

34. Roberts RD, Wong SW, Thiel GB. An unusual case of lead nephropathy. *Arthritis Rheum* 1983;26:1048–1051.

35. Coggeshall HC, Warren CF, Bauer W. The cytology of normal human synovial fluid. *Anat Rec* 1940;77:129–144.

36. Kerolus G, Clayburne GH, Schumacher HR. Is it mandatory to examine synovial fluid promptly after arthrocentesis? *Arthritis Rheum* 1989;32:271–278.

37. Palmer DG. Total leukocyte enumeration in pathologic synovial fluids. *Am J Clin Pathol* 1968;49:812–814.

38. Clayburne G, Baker DG, Schumacher HR. Estimated synovial fluid leukocyte numbers on wet drop preparations as a potential substitute for actual leukocyte counts. *J Rheumatol* 1992;19:60–62.

39. Vincent J, Korn JH, Podewell C, et al. Synovial fluid pseudoleukocytosis. *Arthritis Rheum* 1980;23:1399–1400.

40. Reginato AJ, Maldonado I, Reginato AM, et al. Supravital staining of synovial fluid with Testsimplets. *Diagn Cytopathol* 1992;8:147–152.

41. Schumacher HR, Holdsworth DE. Ochronotic arthropathy. I. Clinicopathologic studies. *Semin Arthritis Rheum* 1977;6:207–246.

42. Krieg AF, Kjeldsberg CR. Cerebrospinal fluid and other body fluids. In: Henry JB, ed. *Clinical Diagnosis and Management by Laboratory Methods*. 18th ed. Philadelphia: WB Saunders, 1991:459.

43. Baer AN, Wright EP. Lipid laden macrophages in synovial fluid: a late finding in traumatic arthritis. *J Rheumatol* 1987;14:848–851.

44. Reginato AJ, Schumacher HR, Allan DA, et al. Acute monoarthritis associated with lipid liquid crystals. *Ann Rheum Dis* 1985;44: 537–543.

45. Gardner GC, Terkeltaub RA. Acute monoarthritis associated with intracellular positively birefringent Maltese cross appearing spherules. *J Rheumatol* 1989;16:394–396.

46. Ugai K, Kurosaka M, Hirohata K. Lipid microspherules in synovial fluid of patients with pigmented villonodular synovitis. *Arthritis Rheum* 1988;31:1442–1446.

47. Astorga GP, Carvajal PR. Lipid spherule associated arthritis. *J Rheumatol* 1990;17:1720.

48. Kay J, Eichenfeld AH, Athreya BH, et al. Synovial fluid eosinophilia in lyme disease. *Arthritis Rheum* 1988;31:1384–1389.

49. Ménard HA, de Médicis R, Lussier A, et al. Charcot-Leyden crystals in synovial fluid. *Arthritis Rheum* 1981;24:1591–1593.

50. Del Blanco J, Valverde J, Mateo L, et al. Charcot Leyden crystals in synovial fluid. *J Rheumatol* 1991;18:1944.

51. Freemont AJ, Denton J. Disease distribution of synovial fluid mast cells and cytophagic mononuclear cells in inflammatory arthritis. *Ann Rheum Dis* 1985;44:312–315.

52. Malone DG. Mast cell numbers and histamine levels in synovial fluid from patients with diverse arthritides. *Arthritis Rheum* 1986;29:953–963.

53. Hasselbacher P. Sickled erythrocytes in synovial fluid. *Arthritis Rheum* 1980;23:127–128.

54. Glickstein SL, Melton JW, Katz P. Sickled cells in synovial fluid: clue to unsuspected hemoglobinopathy. *South Med J* 1989:769–771.

55. Yaron M, Djaldetti M. Platelets in synovial fluid. *Arthritis Rheum* 1978;21:607–608.

56. Endresen GKM. Investigation of blood platelets in synovial fluid from patients with rheumatoid arthritis. *Scand J Rheumatol* 1981;10: 204–208.

57. Lawrence C, Seife B. Bone marrow in joint fluid: a clue to fracture. *Ann Intern Med* 1971;74:740–742.

58. Dorfman HD, Siegel HL, Perry MC, et al. Non-Hodgkin's lymphoma of the synovium simulating rheumatoid arthritis. *Arthritis Rheum* 1987;30:155–161.

59. Barton A, Hickling P. Synovial involvement in Hodgkin's disease. *Br J Rheumatol* 1986;25:391–392.

60. Li CY, Yam LT. Blast transformation in chronic myeloid leukemia with synovial involvement. *Acta Cytol* 1991;35:543–545.

61. Fam AG, Voorneveld C, Robinson JB, et al. Synovial fluid immunocytology in the diagnosis of leukemic synovitis. *J Rheumatol* 1991;18: 293–296.

62. Harden EA, Moore JO, Haynes BF. Leukemia-associated arthritis: identification of leukemic cells in synovial fluid using monoclonal and polyclonal antibodies. *Arthritis Rheum* 1984;27:1306–1308.

63. Gramatzki M, Burmester GR, König KJ, et al. Synovial fluid involvement in null cell acute lymphoblastic leukemia diagnosed with monoclonal antibodies. *J Rheumatol* 1988;15:500–504.

64. Holdrinet RSG, Corstens F, VanHorn JR, et al. Leukemic synovitis. *Am J Med* 1989;86:123–126.

65. Marsh WL, Bylund DJ, Heath VC, et al. Osteoarticular and pulmonary manifestations of acute leukemia. *Cancer* 1986;57:385–390.

66. Gagnerie F, Taillan B, Euller-Ziegler L, et al. Arthritis of the knees in B cell chronic lymphocytic leukemia: a patient with immunologic evidence of B lymphocytic synovial inflammation. *Arthritis Rheum* 1988;31:815–816.

67. Fort JG, Fernandez C, Jacobs SR, et al. B cell surface marker analysis of synovial fluid cells in a patient with monoarthritis and chronic lymphocytic leukemia. *J Rheumatol* 1992;19:481–484.

68. Flint A, Remick DG. Metastatic squamous-cell carcinoma: diagnosis by synovial fluid aspiration. *Acta Cytol* 1984;28:776–777.

69. Villanueva T, Schumacher HR. Cytologic examination of synovial fluid. *Diagn Cytopathol* 1987;3:141–147.

70. Goldenberg DL, Kelley W, Gibbons RB. Metastatic adenocarcinoma of synovium presenting as an acute arthritis. *Arthritis Rheum* 1975;18: 107–110.

71. Fam AG, Kohn A, Lewis AJ. Metastatic carcinomatous arthritis and carcinoma of the lung. *J Rheumatol* 1980;7:98–104.

72. Muñoz-Gómez J, Gómez-Pérez R, Solé-Argues M, et al. Synovial fluid examination for the diagnosis of synovial amyloidosis in patients with chronic renal failure undergoing haemodialysis. *Ann Rheum Dis* 1987;46:324–326.

73. Lakhanpal S, Li CY, Gertz MA, et al. Synovial fluid analysis for diagnosis of amyloid arthropathy. *Arthritis Rheum* 1987;30:419–423.

74. Crugnola A, Schiller A, Radin E. Polymeric debris in synovium after total hip replacement: histological identification. *J Bone Joint Surg* 1977;59:860–862.

75. Bahremand M, Schumacher HR. Effect of medication on synovial fluid leukocyte differentials in patients with rheumatoid arthritis. *Arthritis Rheum* 1991;34:1173–1176.

76. Corbetti F, Malatesta V, Camposampiero A, et al. Knee arthrography: effects of various contrast media and epinephrine on synovial fluid. *Radiology* 1986;161:195–198.

77. Fagan TJ, Lidsky MD. Compensated polarized light microscopy using cellophane adhesive tape. *Arthritis Rheum* 1974:256–262.

78. Gatter RA. The compensated polarized light microscope in clinical rheumatology. *Arthritis Rheum* 1974:253–255.

79. Gatter RA. Use of the compensated polarizing microscope. *Clin Rheum Dis* 1977;3:91–103.

80. Gordon C, Swan A, Dieppe P. Detection of crystals in synovial fluids by light microscopy: sensitivity and reliability. *J Rheum Dis* 1989;48:737–742.

81. Halverson PB, Ryan LM. Triple crystal disease: monosodium urate monohydrate, calcium pyrophosphate dihydrate, and basic calcium phosphate in a single joint. *Ann Rheum Dis* 1988;47:864–865.

82. McCarty DJ. Crystal identification in human synovial fluids: methods and interpretation. *Rheum Dis Clin North Am* 1988:253–267.

83. Pascual E. The ordinary light microscope: an appropriate tool for provisional detection and identification of crystals in synovial fluid. *Ann Rheum Dis* 1989;48:983–985.

84. Fiechtner JJ, Simkin PA. Urate spherulites in gouty synovia. *JAMA* 1981;245:1533–1536.

85. Ettlinger RE, Hunder GC. Synovial effusions containing cholesterol crystals. *Mayo Clin Proc* 1979;54:366–374.

86. Langlands DR, Dawkins RL, Matz LR, et al. Arthritis associated with crystallizing cryoprecipitable IgG paraprotein. *Am J Med* 1980;68:461–465.

87. Netter P, Delongeas JL, Faure G, et al. Inflammatory effect of aluminum phosphate. *Ann Rheum Dis* 1983;42(supp l):114.

88. Tanphaichitr K, Spilberg I, Hahn B. Lithium heparin crystals simulating calcium pryophosphate dihydrate crystals in synovial fluid. *Arthritis Rheum* 1976;19:966–968.

89. Wise CM, White RE, Agudelo CA. Synovial fluid lipid abnormalities in various disease states: review and classification. *Semin Arthritis Rheum* 1987;16:222–230.

90. Georgewill DA, Graham GA, Schoen I. Applicability of the Ektachem 400 analyzer for assaying analytes in miscellaneous body fluids. *Clin Chem* 1988;34:2534–2539.

91. Owen DS. Synovial fluid glucose. *JAMA* 1978;239:193.

92. Wheeler AR, Graham BS. Case report: pseudogout presenting with low synovial fluid glucose: identification of crystals by Gram stain. *Am J Med Sci* 1985;289:68–69.

93. Seegmiller JE. Serum uric acid. In: Cohen AS, ed. *Laboratory Diagnostic Procedures in the Rheumatic Diseases*. Boston: Little, Brown and Company, 1975.

94. Schumacher HR. Synovial fluid analysis and synovial biopsy. In: Kelley WN, Harris ED, Ruddy S, et al., eds. *Textbook of Rheumatology*. 3rd ed. Philadelphia: WB Saunders, 1989:637–649.

95. Kim HJ, McCarty DJ, Kozin F, et al. Clinical significance of synovial fluid total hemolytic complement activity. *J Rheumatol* 1980;7: 143–152.

96. McCutchan HJ, Fisher RC. Synovial leukocytosis in infectious arthritis. *Clin Orthop Rel Res* 1990;257:226–230.

97. Freemont AJ, Denton J, Chuck A, et al. Diagnostic value of synovial fluid microscopy: a reassessment and rationalisation. *Ann Rheum Dis* 1991;50:101–107.

98. Chen JR, Takahashi M, Suzuki M, et al. Pentosidine in synovial fluid in osteoarthritis and rhematoid arthritis: relationship with disease activity in rheumatoid arthritis. *J Rheumatol* 1998;25:2440–2444.

99. Masuda I, Cardenal A, Ono W, et. al. Nucleotide pyrophosphohydrolase in human synovial fluid. *J Rheumatol* 1997;24:1588–1594.

100. Johansen JS, Hvolris J, Hansen M, et al. Serum YKL-40 levels in healthy children and adults. Comparison with serum and synovial fluid levels of YKL-40 in patients with osteoarthritis or trauma of the knee joint. *Br J Rheumatol* 1996;35:553–559.

101. Light RW, MacGregor I, Luchsinger PC, et al. Pleural effusions: the diagnostic separation of transudates and exudates. *Ann Intern Med* 1972;77:507–513.

102. Vernick MG, Sidawy MK. Cytologic examination of serous effusions. Processing techniques and optimal numbers of smears for routine preparation. *Am J Clin Pathol* 1993;99:182–186.

103. Sahn SA and Good JT, Jr. Pleural fluid pH in malignant effusions: diagnostic, prognostic, and therapeutic implications. *Ann Intern Med* 1988;108:345–349.

104. Light RW and Ball WC, Jr. Glucose and amylase in pleural effusions. *JAMA* 1973;225:257–259.

105. Kumar S, Seshardi MS, Koshi G, et al. Diagnosing tuberculous pleural effusion: comparative sensitivity of mycobacterial culture and histopathology. *Br Med J* 1981;283(6283):20.

106. Kaplan LM and Isselbaucher KJ. Abdominal swelling and ascites. In: Fauci SA, Braunwald E, Isselbacher KJ, et al., eds. *Harrison's Principles of Internal Medicine*, 14th edition. St. Louis, Mo: McGraw-Hill, 1998:255–256.

107. Grady HJ, Jacobs DS, Olsowka ES. Chemistry. In: Jacobs DS, Demott WR, Grady JH, et al., eds. *Laboratory Test Handbook*, 4th ed. Cleveland, OH: Lexi-Comp Inc., 1996:89–92.

108. Tolman KG and Rej R. Liver function. In: Burtis CA and Ashwood EF, eds. *Tietz, Textbook of Clinical Chemistry* 3rd ed. Philadelphia: WB Saunders Co., 1999:1125–1177.

109. Arkriviadis EA, Kapnias D, Hadjigavriel M. et al., Serum/ascites albumin gradient: its value as rational approach to the differential diagnosis of ascites. *Scand J Gastroenterol* 1996;31:814–817.

110. Dines DE. Studies on pleural fluid. *Mayo Clin Proc* 1981; 456–460.

111. Sanchez-Lombrana JL, de la Vega J, et al. Tuberculous peritonitis: diagnostic value of ascitic fluid pH and lactate. *Scand J Gastroenterol* 1995;30:87–91.

112. Sathor MA, Simjee AE, Coovadia YM, et al. Ascitic fluid gamma interferon concentation and adenosine deaminase activity in tuberculous peritonitis. *Gut* 1995;36:419–421.

113. Hillebrand DJ, Runyon BA, Yasmineh WG, et al. Ascitic fluid adenosine deaminase insensitivity in detecting tubercuolous peritonitis in the United States. *Hepatol* 1996;24:1408–1412.

114. Pinto MM, Bernstein LH, Brogan DA et al. Immunoradiometric assay of CA 125 in effusions: comparison with carcinoembryonic antigen. *Cancer* 1987;59:218–222.

115. Alexandrakis MG, Moschandrea JA, Koulocheri SA, et al. Discrimination between malignant and nonmalignant ascites using serum and ascitic fluid proteins in a multivariate analysis model. *Dig Dis Sci* 2000;45:500–508.

116. Kline MM, McCallum RW, Guth PH. The clinical value of ascitic fluid culture and leukocyte count studies in alcoholic cirrhosis. *Gastroenterol* 1976;70:408–412.

117. Gow KW, Haley, LP, Phang PT. Validity of visual inspection of diagnostic peritoneal lavage fluid. *Can J Surg* 1996;39:114–119.

118. Neighbor ML. Ascites. *Emerg Med Clin North Am*. 1989;7:683–697.

119. Koneman EW, Allen SD, Janda WM, et al. In: *Introduction to Microbiology: Part I: Color Atlas and Textbook of Diagnostic Microbiology*. 5th ed. Philadelphia: Lippincott, 1997:93–98.

28
Urine

Meryl H. Haber

Patrick C. J. Ward

Examination of urine as an aid in diagnosis has long been an important part of medical practice. Ancient physicians recognized various colors of urine and associated these changes with prognosis (1). With the advent of the microscope, microscopy of various substances, fluids, organs, tissues, and cells became a commonly applied procedure, with great reliance being placed on it for diagnosis. Frequently performed by well-educated physicians, “uroscopy” was also performed by untrained charlatans, called “pisse-mongers” or “pisse-prophets.” The unrestricted proliferation of these quacks finally led, in England in the 17th century, to the enactment of regulations and licensure laws, which protected the public from these individuals (2).

The modern microscope was first applied to the analysis of urine by Golding Bird in the mid-19th century. Bird, in his masterful textbook on urine published in 1842 (3), described certain elements in the sediment that are now commonly diagnosed; crystals and casts were correctly identified and named and were also accurately illustrated with engravings. Richard Bright, working at the same hospital as Bird (Guy's Hospital, London), in 1827, correctly recognized the presence of “red corpuscles” in the urine in a series of patients with severe renal disease, although apparently he did not identify them microscopically (4).

Present-day microscopy is very different from that used by Bird in the 19th century. Although the fundamentals of bright-field microscopy were in use then, the advent of the binocular microscope with better optics and the development of planar and apochromatic lenses has facilitated recognition of various elements. Analysis of urine exemplifies the use of modern microscopic techniques; here, phase-contrast microscopy, polarized microscopy, and interference-contrast microscopy have provided enhanced images that ultimately improve the accuracy of reporting. Staining the elements in a body fluid, whether urine or any other fluid, has become an accepted procedure, often with considerable enhancement of cell identification (5).

Aside from the microscopic techniques performed in the modern clinical laboratory, body fluids may also undergo a variety of physicochemical evaluations. Medical science has come a long way from the ancient and medieval “pisse-prophets” who used visual examination of the urine almost exclusively to diagnose and prognosticate. The advent of the science of chemistry introduced distillation, which was made part of the uroscopy ritual. Various substances were added to the distillation flask or “urine vesicle,” and these, in turn, often produced changes in urine color and precipitation of proteins, which caused lines or bands to form at various levels in the distillation flask. The flask was frequently molded into a human-like shape and contained an alembic “nose.” If a band formed, a marking on the vesicle flask could be established, which would then represent a portion of the human form of the flask (e.g., a band across the abdominal region of the flask). The location of the band would then be interpreted by the physician as indicating the seat of the patient's disease—in this example, the umbilicus (1). This crude application of chemistry to urinalysis was the beginning of modern urine chemistry. Rapid advances in chemistry have occurred during the past 300 years. Many substances were specifically identified in urine (sugar, uric acid), and these tests became available and were subsequently applied to other body fluids. Now chemical analysis is performed routinely, and the most advanced of techniques are utilized, including spectrophotometry, gas chromatography, and a variety of immunologic methods.

The urine reagent strip, a 20th-century invention, has all but replaced complicated individual chemical analyses for the determination of various bodily products in urine. Estimations of glucose, albumin, hemoglobin, and bile, for example, as well as the determination of physical properties such as pH and osmolality, can all be accomplished using dipstick methodology. Advances in the analysis of urine are continually being made, especially in the area of immunodiagnostic testing. Abnormal products of body metabolism may be expected to be found in the various fluids of the body, and these (especially urine) are more easily obtained than tissue. Medical microscopy and the physicochemical examination of body fluids will therefore continue to be an important source of information to assist in diagnosis and patient care.

QUALITY ASSURANCE IN URINALYSIS

Part of "28 - Urine"

Urinalysis, although more frequently performed than any other laboratory test, often does not receive the same quality assurance (QA) attention that is given to other common tests. When conducting a urinalysis, it is important to ensure that the results are accurate, reproducible, and precise (6). Standards of performance and QA procedures must be adopted and applied in the urinalysis laboratory; these measures may be followed with a minimal amount of inconvenience in the modern well-equipped

laboratory, but they should be described and documented appropriately (7,8). The procedure for the performance of a routine urinalysis has several parts; each is discussed separately to delineate major issues of QA. Ordinarily, the attending physician or house-staff officer writes an order for a urinalysis; a nurse, clerk, or other individual then completes a request form.

Urinalysis Request Form

The request form itself should be readily recognized and easy to complete. The use of color coding and bold labeling on the form is helpful. Space should be available for patient information, including name, gender, age, race, identification number (whether an outpatient or inpatient), the reason for ordering the test (i.e., chief complaint, symptoms and clinical impression, surgery), any medications, the time and date of collection, and how the specimen was obtained (e.g., midstream, catheter, aspiration). The form may also serve for reporting the results. If so, ample space should be included for this purpose. Conversely, if the report form is separate, it should have a design similar to that of the request form as far as the color coding and patient information are concerned.

Equipment and Supplies

An overriding consideration when performing a urinalysis is the possibility that the specimen contains an infectious organism that may be transmitted to the technologist. Thus, the test itself, as well as the design of all supplies and equipment used in its performance, must ensure protection for the person performing the test; universal precautions must be adhered to in the handling of all specimens.

Equipment needed for the performance of a routine urinalysis generally consists of a microscope, a centrifuge, and a refractometer. In higher volume laboratories, some of the physicochemical procedures (e.g., dipstick) may be automated or semiautomated, and appropriate instruments necessary for this automation must be available. High-quality work necessitates the use of modern equipment with known performance characteristics.

The microscope should be modern and binocular and have an adjustable condenser with a built-in light source (8). The lenses should be of high quality and should consist of at least three magnifications: low and high power and oil immersion. A lower power scanning lens can be added and often is of value in assessing the general characteristics of the urine sample and in determining whether the field being examined is representative of the entire specimen. The microscope should be kept clean and properly aligned at all times.

Bright-field microscopy is the microscopic technique most frequently used for examination of urine sediment. Tungsten-filament light is transmitted through an adjustable condenser to produce parallel rays of light, so-called Köhler illumination. It is important to be able to reduce the level of light to enhance contrast because many sediment elements have a low refractive index and are difficult to see. High contrast is provided by stopping down the iris diaphragm and lowering the condenser to the level at which the elements are best seen. Light is transmitted through the sediment, and whatever elements are present are observed through the eyepiece lenses (usually 10×). For the most accurate and reproducible results of a sediment examination, it is important that the same microscope be used daily; significant variations in element counts may occur if different microscopes are used. For example, if one observer uses a microscope with wide-field 10× eyepieces and another with ordinary 10× eyepieces, there will be a discrepancy in counts because the field of view is significantly greater with wide-field eyepieces.

The best microscopic lenses are plano to provide a “flat” field of view. Lenses without chromatic aberration are called apochromatic lenses. Some manufacturers produce a lens (the planapo lens) that gives both a flat field and essentially no chromatic abnormalities; this type of lens is usually more expensive and not essential for ordinary everyday work.

Polarized microscopy is a simple and inexpensive addition to the bright-field microscope. All specimens that demonstrate abnormalities, whether macroscopically, chemically, or microscopically, should be examined using two polarizing lenses, a polarizer and an analyzer. The polarizer is inserted between the condenser and the microscopic slide, and the analyzer between the slide and the eyepiece. Particular substances, such as crystals and some lipids, have the peculiar property of being *birefringent* or *anisotropic* (also known as doubly refractile). Through polarized microscopy, objects that have this property are easily observed; objects that do not show birefringence are totally obscured from the field of view. Cholesterol esters, present in the nephrotic syndrome when there is lipiduria, demonstrate a “Maltese cross” pattern in the urine; these same lipids may be entirely missed or misdiagnosed if polarization is not applied to the sediment. Enhancement of ordinary polarization with the use of retardation plates may be helpful in crystal identification. These plates are inserted into the light path to change the path of light vibration. Most commonly used in a first-order red plate, which is useful in distinguishing urate from calcium pyrophosphate crystals, especially in joint fluids. Plates made of mica may be similarly employed.

Phase-contrast microscopy is a technique that has proved to be very useful in examination of the urinary sediment. Phase-contrast microscopy utilizes a special microscope condenser that produces parallel and in-phase light rays, causing the peaks and troughs of the light to be in synchrony. As these waves pass through an object, the wave becomes asynchronous and out of phase. Elements viewed in this manner appear to have a white “halo” around them, greatly enhancing observation and facilitating diagnosis. Phase-contrast microscopy equipment is available from nearly every optical product manufacturer at relatively modest cost and should be standard equipment in the urinalysis laboratory.

Interference-contrast microscopy (ICM) is a valuable microscopic technique that provides a three-dimensional or “shadow-cast” image of the element under study (9). This is accomplished by inserting a pair of beam-splitting prisms in the light path (Nomarski interference), one to actually divide the light and reorient it into phased and perpendicular waves of light (a reference beam and an object beam), and the second to recombine the light *after* it has passed through the element. As the object beam courses through the specimen, both its amplitude and phase are altered;

however, the reference wave does not pass through the element and thus no change occurs in its amplitude or phase. When these two light waves are recombined by passing through the second Nomarski prism, the image of the object appears to be three-dimensional when viewed through the eyepieces of the microscope. ICM is of particular value in teaching because of the ease in demonstrating fine points of structure. The cost of ICM equipment is somewhat high, and therefore it is difficult to justify in most routine laboratories where the urinalysis microscope is used exclusively for patient care.

The *centrifuge*, necessary for sediment concentration for microscopic evaluation, must be calibrated to provide a constant relative centrifugal force (RCF) of 400 (8). Spinning the urine for 5 minutes will provide sediment in which there are no artifacts. Neither the exact time of centrifugation nor the RCF is “set in stone”; each laboratory must determine for itself which settings to use and be certain to check that exactly the same time and RCF are applied to each urine specimen every time. Regular periodic maintenance and calibration checks of the centrifuge are essential to ensure a standardized procedure that will eliminate variations in results.

Refractometers measure the total solids (TS) in a solution by comparing the refractive index of the specimen with water: solids dissolved in urine change the refractive index in proportion to their concentration (10). These instruments are operated with ease and provide a quantitative estimate of the specific gravity using a single drop of urine. Most commercially available refractometers are temperature corrected and have replaced the older (and far less accurate) urinometers (or hydrometers) for the measurement of specific gravity. Reference solutions with known quantities of dissolved substances, such as salt, are prepared to check the calibration of refractometers. Water is used as a standard and is given a value of 1.00. Calibration should be done frequently and at least daily in high-volume laboratories (8).

Osmolality, although not usually tested in the urinalysis laboratory when a routine urinalysis is requested, is a measurement of the number of particles in a solution. Instruments employed to measure osmolality are based on the principle that when a solution is frozen, or vaporized, the point at which this occurs is proportional to the number of particles present. This relationship may not always be a direct one because larger molecules, such as protein, affect the proportionality more than simple molecules, such as sodium chloride. The measurement of urine osmolality is a valuable test when estimating the kidney's concentrating ability, as in the case of possible tubular disease or in the calculation of free water clearance (11). Reference solutions, both high and low, should be utilized at frequent intervals in the testing procedure for instrument calibration to ensure precision and accuracy. Instruments that measure osmolality are complex and need to be treated with the utmost care, always making certain that the manufacturer's instructions are followed exactly.

The supplies used in routine urinalysis should be disposable (7,8). These include a specimen collection system consisting of a collection container and cover, a graduated centrifuge tube with a leakproof cap, labels, and disposable microscopic slides containing a constant-volume chamber for sediment quantitation. Reagent strips are available from a number of different suppliers and can be obtained with several configurations of tests. Most, however, include the following physicochemical tests: pH, hemoglobin, glucose, protein, bilirubin, urobilinogen, ketone bodies, and nitrites. Other commonly available pads test for leukocyte esterase and measure specific gravity (10). Each of these tests is discussed separately in a later section of this chapter.

It is important to conclude this section with a brief comment about reporting. No matter how accurate and precise a procedure may be, or how relevant the results, if one is not consistent in the method of reporting, misinterpretations may result (8). It is therefore incumbent on the laboratory to standardize reporting and to be certain that whoever is performing the test uses the same terminology and identical criteria for the identification of elements. A procedures manual should contain diagnostic criteria and specific definitions; all personnel performing the test should use these criteria. A definition of what constitutes normal urine should be included. Separate mention should be given to abnormal values, and those that should prompt a telephone call to the attending physician, the so-called “panic values,” should be defined.

MACROSCOPIC, PHYSICOCHEMICAL URINALYSIS

Part of "28 - Urine"

Physical Examination of Urine

An adult human produces 1 to 2 L of urine every 24 hours. The urine is formed in the kidney and contains predominantly water in which are admixed mostly urea and salts, especially sodium chloride (11). The amount of solute is directly related to the amount and kind of ingested foods and fluids. Urea is predominantly a byproduct of protein metabolites. Numerous other dissolved constituents are also found in urine, including sugars, bile, other electrolytes such as potassium, products of protein metabolism such as nucleic acids, hormones and their metabolites, vitamins, and some trace metals.

The routine macroscopic, physical, and chemical examination of urine includes appearance, dipstick testing, and measurement of specific gravity (8,10). Normally, urine is pale yellow and has an ammoniacal odor. This odor increases significantly when the urine has been allowed to stand for any length of time and is primarily owing to the release of nitrogenous compounds from bacterial action. The urine is more deeply pigmented as its concentration increases.

Physical Urinalysis

Appearance

When freshly voided, normal urine should be clear and pale yellow to amber in color. Many ingested foods, dyes, or pharmaceuticals may produce a color change. The normal yellow color is owing to the presence of urochrome pigment, a product of hemoglobin metabolism. Color intensity may vary and depends on the concentration of the urine; in highly concentrated urine such as in the early morning, the colors may be more vivid.

Abnormal color of urine occurs in the normal state and as well as in some diseases. Various foods, such as carrots and beets, produce abnormally colored orange to red urine. Although color by

itself is not a particularly sensitive indicator of abnormality, any variation from the normal should be investigated and the specific cause determined. Some conditions may be associated with abnormally colored urine, for example, hemoglobinuria (red), melanin (black), bile pigments (green), chyluria (white), and porphyria (red-purple). In addition, some dyes and pharmaceuticals and some food colorings are notorious for discoloring urine. Specific examples are phenolphthalein (red), methylene blue (blue), and pyridium (yellow-orange) (11).

Normal urine is not always clear, especially after refrigeration. Urate and phosphate crystals often form a cloudy precipitate when urine is subjected to low temperatures. Simply rewarming the urine to room temperature will frequently cause the crystals to dissolve and the urine to become clear again. Urine contaminated with seminal fluid or vaginal secretions and mucus may be cloudy. Clear urine that has been permitted to stand at room temperature for any length of time may become cloudy because of bacterial growth. Turbid urine is commonly associated with abnormal conditions, especially hematuria, leukocyturia, and bacteriuria. Chyluria, or the appearance of lymph in the urine causing cloudiness, is an extremely rare complication of blockage of the lymph passages with the production of a fistula and flow of the lymph into the urinary outflow tract.

Odor

Normal urine has a distinct and easily recognized odor, albeit ill defined. Upon standing, the same urine acquires a pungent ammoniacal odor caused by the action of bacteria and the release of ammonia. Food, such as asparagus, gives the urine a distinct odor promptly after eating. The odor of urine is given little importance in clinical diagnosis, but if an unusual odor is recognized during analysis, it too should be reported. Foul-smelling odors may be associated with urinary tract infections and “fruity” odors with diabetes mellitus when ketones are present.

Collection of Urine

An early-morning, freshly voided specimen is preferred. This should be a clean-catch midstream voided urine (8). The patient is instructed to cleanse the area surrounding his or her urethra using a mild soap and water solution and then to urinate into a clean (not necessarily sterile) collection container that is properly labeled, free of interfering chemicals, can be capped, and is leakproof.

The specimen should be transported to the urinalysis laboratory promptly and examined within 2 hours of voiding (12). Although refrigeration of the specimen is not necessary owing to frequent precipitation of urate or phosphate crystals, in some instances it is preferable to allow the urine to stand at room temperature. The addition of antibacterial agents such as formaldehyde or thymol crystals to the urine to prevent bacterial growth is not advised because these substances interfere with the chemical evaluation.

Once received in the laboratory, the urine is decanted into a graduated centrifuge tube and centrifuged to obtain sediment for evaluation. Transfer pipettes and microscopic slides, preferably with graduated, consistent-volume chambers, are desirable for use. None of the materials mentioned should be reused.

A most important concept in the analysis of urine must be that QA that ensures consistency and accuracy of results (6,7 and 8). Basically, this concept is one of careful control and consistency in every step of the urinalysis procedure to ensure that each and every urine sample is treated and examined in exactly the same manner, and the results obtained are compared with nonsamples to ensure that when and if a discrepancy occurs, corrective action will be taken.

Equipment such as microscopes, centrifuges, reagent strip readers, and refractometer all need to be clean, calibrated, and frequently checked to provide consistency and accuracy in evaluations. The reader is referred to NCCLS Document GP16-P for guidelines with regard to routine urinalysis (8).

Chemical Urinalysis

The performance of a routine physicochemical urinalysis has evolved into a sophisticated, but easily performed, procedure mainly owing to the advent of urine dipsticks (10). These are plastic strips on which are attached a series of chemically impregnated absorbent pads. Each pad contains chemicals that react with a substance in the urine, producing a color change in the pad. This color change is then compared with a series of known standards so that any reaction seen in the patient's urine may be evaluated and quantified. This methodology is simple and easily taught, is rapid, and provides a means of analysis that, although not without some drawbacks, offers an inexpensive and highly useful assessment of several urinary constituents. Most of the available dipsticks contain pads for determination of pH, hemoglobin, glucose, protein, urinary ketones, bilirubin, urobilinogen, and nitrites (13). Other available pad tests, depending on the manufacturer, measure specific gravity and leukocyte esterase (14).

Along with the bottle containing the dipsticks, each manufacturer provides an insert that describes in detail the proper use of the product. Care must be taken to follow these directions exactly. Before using the sticks, one must be certain that they are not outdated and that they have been stored in their original container at the recommended temperature. The bottle should remain unopened until the dipsticks are required for use. The strips should be removed from the container only a few at a time and, once opened, the container should be closed tightly for storage of the remaining dipsticks until further need (8).

In use, the dipstick is immersed into the urine by a rapid, deliberate motion and is then removed. During removal, the edge of the stick is guided over the lip of the urine container, usually a centrifuge tube, to allow excess urine to run off the stick and flow back into the container. Each manufacturer provides a time sequence for the interpretation of each test pad; it is important to follow these instructions because the color reactions on the pads change with prolonged exposure. As pad colors are compared with a colored comparison chart provided by the manufacturer, test results are recorded by the observer. For accurate interpretation, lighting conditions should be good and should simulate daylight. Visual reading of the pads is subject to variation from observer to observer; automated strip readers have improved this variability.

QA procedures need to be employed during the chemical analysis of urine (6). This is performed by utilizing solutions with known concentrations of the analytes being tested for on the dipstick. When spurious or unexpected results are obtained in any patient or if a trend toward abnormal results occurs, the source of potential error must be determined and corrected. Confirmatory testing is performed on the urine sample on occasion. In such instances, it may be desirable to obtain results that exceed those provided by the semiquantitative methodology of the dipstick (8). Confirmatory tests are available for essentially all urine dipstick pads.

Some substances may produce interference with the reaction on the dipstick pad. Many of these are well known and consist of pharmaceutical products, various foodstuffs, and dyes (10,11). In addition, particular physicochemical changes in the urine may interfere with the pad's color reaction and can result in inaccurate false-positive or false-negative readings. As an example, bilirubin, when exposed to light, breaks down and a false-negative reaction on the dipstick pad occurs if the urine has been exposed to direct light for long periods. An excellent example of an interfering substance is ascorbic acid (vitamin C), which is frequently taken in high concentrations in the belief that it affords better overall health and prevents the common cold. This substance is excreted by the kidneys and, when present in the urine, may obscure the dipstick pad for hemoglobin, giving a false-negative reaction. It is of primary importance to provide the laboratory with a good clinical history before the performance of a urinalysis to ensure useful clinical correlations and consistent and accurate interpretations.

Glucose

Glucose is one of a variety of reducing substances that may be present in the urine, but from a clinical standpoint, it is the most important one. Small amounts of glucose may normally be present in the urine, especially after ingesting foods or beverages that have a high concentration (11). Glucose is freely filtered by the glomerulus and almost entirely reabsorbed by the renal tubules, so that under normal circumstances little is found in the urine. *Glycosuria* is defined as the presence of detectable glucose in the urine. This generally occurs when blood levels of glucose exceed 180 mg/dL. At these levels, with similar concentrations occurring in the glomerular filtrate, the renal threshold for glucose is exceeded, and any glucose that exceeds this concentration cannot be reabsorbed by the renal tubules and is excreted in the urine. The renal threshold is not an absolute number and varies somewhat from individual to individual, in the range of 160 to 200 mg/dL.

Diabetes mellitus, a common clinical condition in which serum concentrations of glucose may reach very high levels, is frequently diagnosed and monitored by the detection of glucose in the urine. Elevated blood levels of glucose lead to similarly elevated levels in the glomerular filtrate, which then exceed the renal threshold for glucose, and glycosuria results. Glucose may be found in the urine in conditions other than diabetes; these affect the central nervous system (stroke, neoplasms), kidneys (uremia), endocrine system (overproduction of adrenocorticotropic hormone), and liver (glycogen storage disease) or may be related to general metabolic problems, such as starvation and obesity (10,11,15). Some pharmaceutical agents, such as diuretics and birth control pills, may also cause glycosuria.

Although glucose may be measured in various ways, the dipstick method employs an enzymatic methodology (10). This method takes advantage of the specific reaction of the enzyme glucose oxidase, which catalyzes the conversion of glucose to gluconic acid and hydrogen peroxide. A second enzyme, peroxidase, promotes the reaction of the H_2O_2 generated with a chromogen (iodine complex) to form a brown oxidized compound. The chromogen varies with the manufacturer, as does the sensitivity (10,13).

$$\text{Glucose} + O_2 \xrightarrow{\text{glucose oxidase}} \text{Gluconic acid} + H_2O_2$$

$$H_2O_2 + \text{Chromogen} \xrightarrow{\text{peroxidase}} \text{Oxidized colored chromogen} + H_2O$$

Color charts are provided with the dipsticks to enable semiquantitative determination of the amount of glucose present in the specimen. Reporting is commonly given in a 1+ to 4+ format, which may then be equated to concentrations ranging from approximately 100 to 2,000 mg/dL. Sensitivity of the various dipsticks varies with the manufacturer but is usually in the range of 50 to 100 mg/dL. Reagent strip testing for glucose has problems because at the higher concentrations (i.e., 1 to 2 g/dL), color differences in the pads are more difficult to interpret.

Reducing substances, especially other sugars, do not give a positive dipstick test because the method, which utilizes glucose oxidase, is specific for glucose. However, false-positive reactions may occur if the urine container used is not disposable and strong oxidizing detergents remain from prior washings (11,15). False-negative results may result when any part of the chemical reaction is interfered with, such as when high doses of aspirin or ascorbic acid are present in the urine and these compounds, both reducing agents, prevent the formation of the chromogen by preventing its oxidation by H_2O_2. One other relatively frequent cause of false-negative results is glucose being metabolized by bacteria when the urine sample has been allowed to stand for any length of time after voiding.

A second method for the measurement of reducing sugars, including glucose, in urine is by means of the copper reduction test (Benedict's reaction) (Table 28.1). This test is performed by

adding urine to a tablet composed of copper sulfate, sodium citrate, sodium carbonate, and sodium hydroxide (Clinitest, Bayer Corp., Consumer Division, Morristown, NJ, U.S.A.) (10). The cupric ion ($CuSO_4$), in the presence of a base (NaOH), reacts with reducing substances in the urine to produce cuprous ions plus heat. Depending on the amount of sugar present, a color change occurs that can then be compared with reference charts to quantitate the reaction. This test is performed by simply adding urine and water to a test tube containing a Clinitest tablet, waiting a short amount of time for the chemical reaction to occur (the solution boils with heat liberation and color development; the bottom of the test tube should not be touched), and then quantitating the amount of sugar present by comparing the color of the reactants with a color chart. The manufacturer's directions should be followed in their entirety; accuracy is related to adding the correct amounts of urine and water as well as to the timing of the reaction and the development of color.

TABLE 28.1. COMPARISON OF THE REACTIONS OF GLUCOSE OXIDASE COPPER SULFATE TESTS FOR URINE SUGARS

Substance	Dipstick	Copper Sulfate Tablet
Glucose	Positive	Positive
Other reducing sugars	Negative	Positive
Pharmaceutical agents	Negative	Some false positives
Detergent contaminants (peroxide or strong oxidants)	False positive	Negative

Reducing sugars other than glucose are infrequently seen in urine but are sometimes discovered in a pediatric population. They may be overlooked unless careful attention is paid to the possibility of their existence. This is accomplished by means of testing each pediatric patient's urine with Clinitest or by using another similar nonenzyme-specific methodology. This copper sulfate method will detect the presence of all reducing sugars, some of which may not be glucose. To verify the presence of a nonglucose-reducing sugar, comparisons should be made with dipstick results from the same patient; if the sugar is not glucose, the dipstick test is negative. Thin-layer chromatography may be utilized in the specific identification of nonglucose-reducing sugars when inherited inborn metabolic diseases of sugar metabolism are considered.

Galactosemia, an inborn error of metabolism, must be screened for using the copper sulfate methodology. The use of a dipstick will not discern galactose. Lactase deficiency in the infant results in lactosuria; early diagnosis is of importance because failure to thrive and intestinal damage may result if ordinary milk, which contains lactose, is not removed from the diet. Sucrase deficiency, less common than lactase deficiency, also results in mellituria (sucrosuria), which can be diagnosed using a Clinitest tablet. Symptoms are similar to those seen in lactosuria. Other sugars that may be present in the urine are fructose and pentose. Fructosuria and pentosuria are infrequent causes of nonglucose mellituria.

Protein

Relatively small quantities of protein are normally excreted in the urine, usually no more than 100 to 150 mg/day (16,17). Normal urine volume is between 1,000 and 1,500 mL/day; thus, the expected protein concentration in a normal sample is in the range of 8 to 10 mg/dL (10,15). The predominant protein present in the urine is albumin, which is filtered by the glomerulus and is not reabsorbed by the tubules; albumin accounts for less than 30 mg/day in normal urine. The other main source of urinary protein is the renal tubules, especially the distal tubules. These tubules account for the other two thirds of urine protein excretion, and of this portion, approximately one half is Tamm-Horsfall mucoprotein (THM), which is secreted by the distal tubular epithelium and collecting ducts and is of importance in the formation of urinary casts. Small quantities of other urinary proteins come from filtered low molecular weight proteins of the blood plasma that are not reabsorbed by the tubules, microglobulins from renal tubular epithelial (RTE) cells, and prostatic and vaginal secretions (16).

Proteinuria, defined as increased protein in the urine, is a frequent finding in intrinsic diseases of the kidney and urinary tract. Proteinuria, often associated with urine sediment abnormalities, frequently serves as a basis for the diagnosis of intrinsic renal disease (16). Albumin is most often the predominant protein excreted in excess in urine in disease states and may account for as much as 80% to 90% of the total protein present. In multiple myeloma, a neoplastic condition affecting plasma cells, large quantities of a low molecular weight globulin (Bence Jones protein) may be present in the urine, replacing albumin as the predominant protein. The presence of substantial amounts of albumin in the urine of a patient with known myeloma may be indicative of evolving amyloidosis. Proteinuria may accompany generalized non-urinary tract conditions such as hemorrhage, systemic infections, heart failure, dehydration, and starvation.

Normal persons may excrete abnormal quantities of protein, mainly albumin, in some situations such as during strenuous physical exercise (18), pregnancy, exposure to extreme cold, and psychological stress.

Orthostatic proteinuria is a condition mainly found in adolescents and young adults in which increased quantities of protein are found in the urine when the patient is in a vertical position; the proteinuria disappears when the patient is resting horizontally (15). Pressure on the renal veins when standing is thought to be the cause of this condition, and it frequently disappears with increasing age. Testing the first urine specimen of the morning for protein and comparing the result with that obtained from a specimen collected later in the day after physical activity will usually provide the correct diagnosis and may eliminate the need for a more elaborate workup.

Dipstick screening for protein is designed to differentiate between what is considered a random normal amount (i.e., less than 10 mg/dL) in a given urine sample, preferably an early-morning specimen, and quantities that exceed these norms. One reason for the early-morning collection of urine is that samples of urine obtained after eating or drinking may produce false-negative dipstick tests for protein because of the dilutional effects of food and water. Whenever unsuspected proteinuria is found, confirmatory tests for protein are essential because a diagnosis of proteinuria carries with it considerable clinical importance (8). Any diagnosis of proteinuria should be verified and substantiated. Clinically, proteinuria is considered mild (less than 0.5 g/day), moderate (0.5 to 4 g/day), or severe (more than 4 g/day) (10). Most patients with intrinsic renal disease such as acute glomerulonephritis or pyelonephritis present with moderate proteinuria, but patients with the nephrotic syndrome commonly excrete more than 4 g of protein daily in the urine.

Quantitative protein concentrations can be determined by setting up standards containing measured amounts of protein and comparing the precipitation obtained on an unknown urine

sample with these standards using a nephelometer or spectrophotometer. Immunoassays have recently been employed for the accurate determination of urinary protein content (16). In practice, a qualitative estimation of protein is initially made by comparing the degree of precipitation (i.e., turbidity) with a set of standard tubes containing known concentrations of protein. This reaction may be classified as 1 to 4+ or may be quantified as milligrams per deciliter. A trace reaction is one in which there is perceptible, but minimal turbidity (approximately 20 mg/dL); a 1+ reaction (approximately 50 mg/dL) shows definite turbidity, but no white precipitation granules; in a 2+ reaction (approximately 250 mg/dL), granules and turbidity are seen; a 3+ reaction (approximately 500 mg/dL) demonstrates granules and flocculation; in a 4+ reaction (more than 1 g/dL), large clumps of white precipitate are seen.

Present-day dipsticks for the determination of urine protein use a colorimetric methodology based on the concept of “the protein error of indicators.” This method uses a physicochemical property of protein that causes particular indicators (e.g., tetrabromophenol blue) to change color in the presence of varying concentrations of protein when the pH is kept constant (the dipstick pad is impregnated with buffers kept at a pH of approximately 3). This method provides a rapid means of testing for protein, but one that is somewhat difficult to interpret when low concentrations of protein are encountered, generally in the trace to 1+ range. This difficulty is caused by subtle color changes that occur in the indicators in this range (i.e., yellowish tan). As protein levels increase, the colors become more intense and easier to interpret (i.e., green and blue). Results are read as trace and 1+ to 4+. A dipstick trace reaction corresponds to approximately 15 to 30 mg/dL of protein, whereas a 4+ reaction corresponds to more than 2 g/dL.

Standard dipsticks are most sensitive to albumin. They may completely fail to detect other proteins such as globulins, hemoglobin, myoglobin, and Bence Jones protein. The presence of such proteins must be confirmed by other methods, notably (and most simply) by the sulfosalicylic acid test (11).

Bence Jones protein is found in the urine in approximately one half of patients with multiple myeloma; it may also be present in malignant lymphoma and macroglobulinemia (15,19). Reagent strip testing for Bence Jones protein often is not reflective of the type or amount of protein present in the urine because they are globulins, and the protein pad may even show a completely negative result (10) unless there is accompanying albuminuria. Urine electrophoresis or immunoelectrophoresis is the procedure of choice for recognizing this protein (16), but other tests may also prove helpful, especially as screening tests, because of the peculiar solubility properties of Bence Jones protein.

When gradual heat in a water bath is applied to a test tube of urine containing Bence Jones protein, precipitation occurs between 40° and 60°C. However, if the heat is continued by putting the tube in boiling water, precipitated Bence Jones protein will redissolve at approximately 100°C after a few minutes. Cooling the tube reverses the process, so a precipitate will reform at 40° to 60°C and will redissolve when the temperature falls below 40°C (13). Severe proteinuria in myeloma, if left untreated, often produces renal tubule injury (myeloma kidney), which may lead to clinical manifestations of the nephrotic syndrome and to renal failure. If large amounts of protein other than Bence Jones protein are present in urine, they may interfere with the heat test. Filtering the urine after boiling it and heating it to 100°C and then cooling it to room temperature will remove these interfering substances.

Clinically, it is often valuable to know the total amount of protein excreted in the urine in 24 hours (16). A single urine sample is not sufficient to provide this information; therefore, a 24-hour sample must be collected. Chemicals that cause no interference with the tests for protein are added to the urine collection bag to preserve the specimen. At the end of the timed collection period, the urine is analyzed for its protein content using a precipitation test (i.e., sulfosalicylic acid turbidity). The results are reported as milligrams of protein per deciliter and, depending on the total volume of urine voided, the amount of protein excreted by the kidneys in a given day, can easily be calculated.

Microalbuminuria

The normal excretion of albumin in urine is less than 30 mg/day. Albumin levels greater than 30 mg/day may be present while total urinary protein levels remain normal, i.e., less than 150 mg/day. The commonly used dipstick is negative until total protein excretion is greater than 300 to 500 mg/day. The term microalbuminuria describes levels ranging from 30 to 299 mg/day (20). Hence, microalbuminuria may be present in a dipstick-silent range. Detection of small increments in albumin excretion is of critical importance in that it portends irreversible glomerular injury both in insulin and non-insulin-dependent diabetes. The sulfosalicylic acid test may exhibit trace positivity in this dipstick-silent range of albuminuria but is unreliable for this purpose.

Trace positivity with the sulfosalicylic acid test should then lead to confirmatory testing for microalbuminuria. This is usually performed by nephelometry, an immunoturbidimetric method that provides an accurate measurement of albumin. Because the rate of albumin excretion is in question, a timed collection of urine is required.

An alternative to this confirmatory procedure is the measurement of a urinary albumin:creatinine ratio in a first-voided morning urine (21). Microalbuminuria by this method is interpreted as 30 to 300 mg/g of creatine. The test is also commonly performed on a random “grab” urine specimen. A recently described reagent strip device (dipstick) for quantitation of the urinary albumin: creatinine ratio (Clinitek, Bayer, Newbury, UK) appears promising (22).

It is generally accepted that diabetics should be screened for microalbuminuria 8 to12 years postdiagnosis and yearly thereafter. Detection of microalbuminuria in diabetics, whether insulin dependent or not, should provoke intervention with a view to preventing or slowing down the progression of glomerular disease (21,23).

Microalbuminuria is also encountered in middle-aged nondiabetic patients with major cardiovascular risk factors such as hypertension, weight gain, and increased plasma cholesterol. These factors may be associated with microalbuminuria independently of each other (20).

Blood

Hematuria, defined as increased numbers of red blood cells (RBCs) in the urine, is an important indicator of diseases involving the genitourinary tract. Normal urine may contain as many as five RBCs/μL. Increased numbers of erythrocytes may come from a variety of sources, including the kidneys, ureters, bladder, urethra, prostate gland, and uterus and vagina. Erythrocytes may be found intact or they may lyse and release their contained hemoglobin into the urine, resulting in *hemoglobinuria*. Frank bleeding into the urine causes a reddish discoloration, but fewer RBCs or less hemoglobin in the urine (i.e., occult bleeding) have been described by some patients as a "smoky" urine. Normally, there are small numbers of RBCs in the urine sediment, but these are not detectable using dipstick methodology.

There are many causes of hematuria (19). Among the more common are bladder and kidney tumors, trauma to the kidneys, glomerulonephritis, pyelonephritis, renal calculi, and bleeding disorders related to anticoagulant use. Hemoglobinuria, however, is far less common. It may be associated with trauma or transfusion reactions and is seen in severe burns or poisonings. Myoglobinuria, the presence of myoglobin in the urine, is found on occasion after severe physical exercise or in trauma where muscle fiber necrosis has occurred. Myoglobin gives a positive dipstick reaction for hemoglobin, and therefore the distinction between myoglobinuria, which indicates rhabdomyolysis, and hemoglobinuria must be made using other tests (15). Myoglobinuria should be considered when the dipstick pad is positive for hemoglobin, but microscopic analysis of the sediment does not show RBCs.

Dipstick testing for hemoglobin is based on the property of hemoglobin to catalyze the oxidation of a chromogen because of the release of the RBC enzyme hemoglobin peroxidase (10). Various manufacturers use different chromogens that, when catalyzed by hemoglobin peroxide in the presence of an organic peroxidase buffer, form a green to dark blue chromogen.

$$H_2O_2 + \text{chromogen} \xrightarrow{\text{Hemoglobin peroxidase}} \text{colored chromogen} + H_2O$$

Reagent strip pads for hemoglobin are impregnated with the benzidine chromogen and buffered organic peroxide. If hemoglobin is present in the urine, the hemoglobin peroxidase catalyzes the reaction shown above so that a color develops on the pad. This colored pad can then be compared with color charts for a semiquantitative determination of the amount of hemoglobin present. The readings are 1+ to 3+ and indicate trace, small, moderate, and large amounts of blood. Intact RBCs are lysed when they come in contact with the pad, resulting in a speckled pattern. In most urine containing blood, intact RBCs, as well as free hemoglobin from lysed RBCs, combine to provide the positive reaction on the strip.

Hemoglobin dipstick testing, especially for microscopic hematuria, has been very effective (24). However, there are well-known interfering substances that must be considered whenever the clinical history or the sediment examination does not substantiate the results of the dipstick test. Interference by ascorbic acid (vitamin C) is a prime example of this. False-negative reaction errors can be avoided if a careful clinical history is taken and a check is made for the presence of vitamin C in the urine (13). Other methods used to detect hemoglobin when interfering substances are present include direct chemical analysis using a spectrophotometer or by use of another dipstick with a hemoglobin pad that does not produce false-negative results for hemoglobin in the presence of ascorbic acid.

pH

The kidneys, along with the lungs, are prime regulators of acid-base balance. Depending on the foods ingested, the state of health or disease, lung and renal function, medications being used, and a host of other issues, urine pH may vary from 4.5 to 8. pH is a measure of the hydrogen ion content of the urine. Acid urine, by definition, has a pH less than 7, whereas the pH of alkaline urine is greater than 7. The usual finding in freshly voided samples from normal patients is acidic urine with pH between 5 and 6. If fresh urine is allowed to stand at room temperature before analysis, the pH will change from acid to alkaline owing to the growth of bacteria and the breakdown of urea in the urine with the liberation of ammonia. The kidneys constantly regulate acid-base balance and tend to conserve various ions, especially sodium. This process is accomplished primarily by sodium reabsorption in the renal tubules with concomitant excretion of ammonium and hydrogenic ions (25). In metabolic acidosis such as occurs in diabetes, for example, the urine is acidic; in metabolic alkalosis, the urine is alkaline. Control of the urine pH becomes important in the treatment of certain diseases, such as renal stone formation and bacterial infections of the urinary tract. Maintaining alkaline urine provides a more effective environment for some antibiotic agents, and drugs that keep the urine alkaline therefore assist in the use of these agents. Renal calculi may form from soluble inorganic ions, and by controlling pH, the formation may be slowed or prevented (i.e., uric acid stones tend to develop in acid urine and calcium carbonate stones in alkaline urine).

Measurement of urine pH is facilitated by dipstick methodology and comparison with a color chart (26). Confirmatory testing may be performed with a pH meter, but this is usually not necessary. The pH pad is impregnated with two indicators, methyl red and bromthymol blue, which change color when immersed in urine. Two indicators are usually used to provide a broad range of color change, orange-green-blue. QA checks of pH dipsticks should be made regularly with known solutions (7,8).

Ketone Bodies

Ketonuria, or the presence of ketone bodies in the urine, occurs in conditions in which there is incomplete metabolism of fat. In normal metabolism, the body completely metabolizes fat to carbon dioxide and water. However, in disease conditions such as diabetes, especially when poorly controlled, the supply of carbohydrate

is ineffectually utilized, more energy from fat is required, and fat is frequently broken down as an energy source, with the products of fat metabolism appearing in the blood (ketosis) and urine (ketonuria). These intermediary products, which are formed in the liver, are called ketone bodies; they are acetoacetic acid, acetone, and betahydroxybutyric acid (15). The latter two are derivatives of acetoacetic acid. Nearly 80% of the ketone bodies found in the urine are betahydroxybutyric acid, with acetone accounting for only 2% and acetoacetic acid making up the remainder (10). Ketones are excreted in the urine along with basic ions, depleting the body of its anions and producing acidosis.

Diabetes is the most common disease in which ketosis and ketoacidosis are found, but there are many other conditions in which ketosis also occurs. Among the most important on a worldwide basis is starvation, but in the United States, ketosis is also produced in acute dieting. Under this condition, the body begins to break down its own fat surplus, which results in the production of ketone bodies. Other conditions in which ketones are found in the urine are anorexia nervosa, prolonged vomiting, and in several gastrointestinal diseases.

Dipstick testing for ketone bodies is based on the principle that both acetone and acetoacetic acid react with sodium nitroprusside in an alkaline buffer medium to produce a purple complex (Legal's test). Betahydroxybutyric acid does not react. When the color is fully developed, it is compared with a color chart to determine the amount present; the test is interpreted as trace (5 mg/dL) to 4+ (more than 100 mg/dL) (purple color) (10). Acetoacetic acid is the most sensitive reactant, and because the proportions of the three ketone bodies are similar in ketotic states, it is not clinically important that betahydroxybutyric acid is not measured.

Certain dyes, such as phthaleins, may produce a red color on the test pad. However, this color is different from the violet-purple of acetoacetic acid and should not be confused with it. Similarly, phenylketones may give an orange-red color that may be misinterpreted as a positive dipstick reaction. If the urine stands too long before analysis, there is a potential for ketones to be broken down by bacterial action and give a false-negative reaction. Specific chemical tests are available to measure individual ketone bodies, but these are rarely used today because they have little clinical importance.

Reagent tablet tests exist for the determination of urinary ketones and can also be used on serum samples (10). These tablets are sometimes used to follow the course of strict dieting or when the concentration of acetoacetic acid is so high that dilution is required for accurate estimation. In using this product (Acetest Reagent, Bayer), lactose is added to the other reactants on the test strip to provide enhanced color differentiation.

Bilirubin

Bilirubinuria, the presence of bilirubin in the urine, is an important early indicator of liver dysfunction and of intrinsic or extrinsic biliary obstruction. Bilirubinuria may be diagnosed before clinical jaundice appears. Bilirubin is formed from the breakdown of RBC hemoglobin. It attaches to albumin (indirect or unconjugated bilirubin) and is carried to the liver to be processed by the hepatocyte into a water-soluble form (direct or conjugated bilirubin). Unconjugated bilirubin is water insoluble and cannot be excreted by the kidneys because of its albumin binding and the fact that albumin is not generally filtered by the glomerulus. Conjugated bilirubin, on the other hand, is excreted in the urine because of its water solubility and ease of transport across the glomerular basement membrane (GBM). In normal circumstances, conjugated bilirubin never reaches the kidneys; it is entirely secreted into the bile after conjugation as bilirubin diglucuronide and then into the intestines, where it is broken down by bacteria into urobilinogen. In pathologic conditions, however, such as bile duct obstruction or hepatitis, conjugated bilirubin enters the bloodstream owing to very high concentrations of processed bilirubin and is then found in the urine because of its water solubility and ease of passage across the GBM (15).

Small quantities of conjugated bilirubin are normally present in urine (approximately 0.02 mg/dL), but concentrations this low are not detectable using ordinary dipstick methodology. Unconjugated bilirubin is not normally found in the urine. A crude but rapid method for the determination of bilirubin in urine is often performed by house-staff officers when they want a quick answer to the question of whether a patient is jaundiced. Because conjugated bilirubin is water soluble, vigorous shaking of the test tube containing the suspect urine will produce a yellow foam if increased quantities of bilirubin are present, whereas if the urine contains normal amounts of bilirubin, the foam will be white.

Methodology for the semiquantitative determination of conjugated bilirubin in the urine depends on the color produced by the reaction of a diazotized aniline dye and bilirubin; a purplish azobilirubin compound develops. Ictotest Reagent tablets (Bayer), utilizing similar diazotized dyes, are available for use to confirm dipstick results and are more sensitive than the dipstick pads and less affected by interfering substances (10). The exact aniline dye used in dipsticks varies with the manufacturer. Results are interpreted as negative to 1+ (small amounts of bilirubin) to 3+ (large amounts of bilirubin).

It is of considerable importance that the urine be tested for bilirubin immediately after voiding because bilirubin is unstable and breaks down rapidly, especially after exposure to light; prolonged exposure to room air will also cause the dipstick pad reaction to become negative owing to oxidation of the bilirubin to biliverdin, a substance that is nonreactive to the diazo dyes. False-negative results are obtained on the bilirubin pad when the tested urine contains large quantities of ascorbic acid (vitamin C) or nitrites (released because of bacterial growth).

Urobilinogen

Urobilinogen is also a product of hemoglobin metabolism and is formed with the degradation of bilirubin in bacteria in the intestine. Approximately one half of the urobilinogen is excreted in the feces and the other half is reabsorbed into the circulation from the intestines. This reabsorbed urobilinogen is then filtered by the kidneys and appears in the urine if increased in amount. In normal persons, only small amounts of urobilirubin are excreted in the urine daily, usually less than 4 mg; a random urine

sample contains 0.1 to 1 mg/dL. Serum and urine urobilinogen levels increase when there is increased bilirubin production. This may occur in severe liver disease such as cirrhosis or hepatitis (the liver cannot remove reabsorbed urobilinogen from the blood) and in hemolytic disorders (excess urobilinogen is produced in the intestines because of excessive bilirubin production and the liver cannot handle it all). Blockage of the biliary system, such as occurs in ampullary carcinoma or with gallstones in the common bile duct, does not lead to an increase in urine urobilinogen levels because bilirubin never reaches the intestinal tract where it can be broken down into urobilinogen.

The dipstick method for the detection of urobilinogen (Multistix, Bayer, Newbury, UK) makes use of the Ehrlich reaction; when urobilinogen is mixed with dimethylaminobenzaldehyde in an acid buffer, a tan to orange colored compound is formed (10)]. Using the Ehrlich reactants, some compounds other than urobilinogen (i.e., porphobilinogen, sulfonamides, and aminosalicylic acid) also produce a colored reaction and, if present, cause false-positive results. A relatively new and more specific method for the detection of urobilinogen utilizes a stable diazonium salt (4-methoxybenzenediazonium fluoborate) in an acid buffer. With the Chemstrip dipstick (Roche Diagnostics, Indianapolis, IN), a red azo dye color is produced when urobilinogen is present in excess (27). After the dipstick is immersed in urine, the color produced is compared with a color chart to interpret the urobilinogen concentration. No color development on the pad suggests that a normal amount of urobilinogen is present in the specimen (i.e., less than 1 mg/dL); development of significant color implies an abnormal concentration of urobilinogen in the urine. Compounds other than urobilinogen do not produce color on the pad.

False-negative results for urobilinogen can be obtained whenever the urine is allowed to remain in daylight for extended periods after voiding, as the urobilinogen breaks down. If formaldehyde is used as a urine preservative, it will inhibit the reaction and produce a false-negative result. Pharmaceuticals that contain azo dyes may mask the colors of the urobilinogen pad and lead to erroneous interpretations. Finally, the Ehrlich reaction may be inhibited by high concentrations of nitrites, a product of bacterial metabolism (11,28).

Determination of the presence of excess urobilinogen in urine in a jaundiced patient may be of primary importance in differentiating between obstruction of the biliary tract or upper intestines and hepatic or hemolytic disease. If the upper bowel or common bile duct is obstructed, bilirubin does not reach the intestines and urobilinogen is not produced and therefore is not found in the urine. Conversely, in the same conditions, bilirubin is greatly elevated in both serum and urine, and the bilirubin test pad will show a positive result.

Nitrite

This test provides an effective and rapid method for screening urine for the presence of a bacterial infection (29). The test is based on the principle that most bacteria found in urine have the ability to reduce urine nitrate, a plentiful constituent of urine, to nitrite. These organisms include *Escherichia coli*, *Klebsiella*, *Proteus*, *Staphylococcus*, and *Pseudomonas*, to mention a few that are frequent causes of urinary tract infections. The determination of clinical bacteriuria, is of major significance when culture shows the presence of more than 105 colonies per milliliter. Bacterial infections of the urinary tract are usually accompanied by the presence of large numbers of white blood cells (WBCs) in the urine, most often neutrophils. Occasionally, urinary tract infection may demonstrate no clinical symptoms; in these asymptomatic individuals, the urine examination is of prime importance in diagnosis and in the prevention of long-term complications, especially in the kidneys. Urine nitrite determinations are not designed to replace culture as a means of identification of specific organisms or to replace sediment examination to determine whether the bacterium is a contaminant. However, the nitrite test does serve as a valuable screening tool, along with the leukocyte esterase dipstick test, in defining whether a urinary tract infection is present (30). Concomitant negative dipstick results from both tests provide a predictive value of greater than 95% that culture for urinary tract infection will be negative (10).

The chemical basis of the nitrite reaction is that in an acid environment, nitrite reacts with an aromatic amine (sulfanilamide or *p*-arsanilic acid) to form a colored diazonium salt that in turn reacts with a hydroxybenzoquinoline to provide a pink color (Griess reaction). A positive reaction occurs when the bacterial content is greater than 10^5/mL. It is imperative that fresh urine samples are used when this test is interpreted to eliminate the possibility that the nitrate has been produced by bacterial contamination. First-voided morning specimens are best for detecting bacteruria. If bacterial infection is present in the urinary tract, a single dipstick test will be positive in 80% to 90% of patients when the first morning specimen is promptly examined (10). The nitrite test is qualitative, and any shade of pink produced is considered a positive result and indicative of bacterial infection of the urinary tract.

Negative test results for nitrite are found when the particular species of bacteria does not reduce nitrate in the urine. Therefore, a negative dipstick result when the patient is suspected of having an infection of the urinary tract should always be verified by urine culture and/or microscopy of the sediment. Fortunately, most bacteria that cause infections of the urinary tract are gram negative and are nitrite forming. However, yeasts and a significant number of non-nitrite-producing, gram-positive cocci may cause infections of the urinary tract and produce a negative nitrite dipstick reaction. False-negative results can also be caused by the presence of high levels of ascorbic acid and urobilinogen in the urine (27). Adequacy of diet, especially in vegetables, to allow sufficient nitrates to be formed is essential to the proper interpretation of this test. A diet deficient in vegetables may lead to erroneously false-negative dipstick reactions in urinary tract infections. Antibiotic agents may inhibit the growth of bacteria, even if present, and a false-negative test result is reported. False-positive results are obtained when various dyes, such as pyridium, discolor the strip test pad pink or red.

Leukocyte Esterase

Leukocyturia or *pyuria* is a valuable indicator of the presence of urinary tract inflammation (31,32). The relatively recent advent of an enzymatic test for the presence of leukocyte esterase has

provided a simple means for the detection of increased numbers of WBCs in the urine (27). Neutrophils are nearly always present in infections and inflammations of the kidneys and lower urinary tract. They are nearly always associated with bacterial infections and are thus found in the urine in these conditions in elevated numbers. WBCs are normally present in a centrifuged urinary sediment in relatively small numbers, usually fewer than five per high-power field (hpf). Counts of WBCs of more than 10 per hpf may be considered abnormal and should be investigated further (25). Pyelonephritis, cystitis, urethritis, renal calculi, interstitial nephritis, and glomerulonephritis all give rise to leukocyturia. On close examination, the WBCs are predominantly polymorphonuclear neutrophils (PMNs), although a scattering of other WBC types may be seen. Until the initiation of the leukocyte esterase test, the presence of leukocyturia was verified by microscopic analysis of the sediment (33).

The leukocyte esterase test is based on the release of the enzyme leukocyte esterase from neutrophils, on the ability of this enzyme to split an ester and form a pyrrole. Depending on the test, the pyrrole either reacts with a diazo compound to form a colored azo dye (Multistix) or becomes oxidized to form an indigo dye that then reacts with a diazo salt to form a colored reaction within 2 minutes (Chemstrip) (10,27). The color produced depends on the number of PMNs present in the urine. The sensitivity of the dipstick test has been adjusted to provide a negative reaction when fewer than five WBCs per hpf are present and a trace when there are from five to 15 WBCs per hpf. False-positive results are usually not seen, except when oxidizing agents have been mistakenly added to the urine container. Cells other than leukocytes that may be present in the urine, such as RBCs, epithelial cells, and bacteria, do not contain the leukocyte esterase enzyme in their cytoplasm and cannot produce a positive strip reaction. However, large numbers of eosinophils or a heavy infestation of the vaginal parasite *Trichomonas vaginalis* may provide sufficient esterase to give a false-positive result (10). False-negative results are found when there are high levels of vitamin C, oxidizing agents, formalin, or albumin in the urine (27). Mention should be made of the fact that the leukocyte esterase test may occasionally be positive even when neutrophils cannot be demonstrated in the urinary sediment. This may be owing to lysis of neutrophils in the urine before microscopic examination and release of their cytoplasmic contents. If such a circumstance occurs, a repeat of the urinalysis using a freshly voided specimen will usually reveal many intact WBCs in the sediment.

Specific Gravity

Waste product excretion and reabsorption of vital chemicals are two of the most important functions of the kidneys (25). A measure of their reabsorptive capacity provides a valuable test of renal function, and estimation of urine specific gravity provides just such a measurement. The specific gravity of a liquid provides an estimation of the density of the solid substances dissolved in it. As a measure of density, specific gravity varies according to the size and number of the particles present in the liquid. The specific gravity of urine is a useful indicator of the ability of the renal tubules to concentrate or dilute the urine. Normally functioning kidneys are able to alter the specific gravity of urine in a range of 1.003 to 1.035; early-morning urine specimens are usually more concentrated than random samples taken at various times during the day and tend to have specific gravity readings higher than 1.020, whereas urine voided at other times during the day, especially after ingestion of water and food, are lower. A specific gravity of more than 1.025 in a random urine sample serves as reasonably good evidence that the concentrating ability is preserved. Specific gravity may be at the upper end of the normal scale (i.e., greater than 1.025) in patients with congestive heart failure, severe liver disease, dehydration and vomiting, and diarrhea who have experienced severe water loss. Kidneys lose their concentrating abilities in severe end-stage renal disease; in these cases, the urine may have a “fixed” specific gravity near 1.010. Low values are also obtained in diabetes insipidus, in which antidiuretic hormone excretion is compromised.

The refractometer, which is temperature corrected, is the most accurate and widely used instrument for the estimation of urine specific gravity (6,7 and 8). It measures the amount of dissolved substances in a liquid by measuring the refractive index of the liquid. This task is performed by comparing the velocity of light in air with the velocity of light in the urine. Increases in the concentration of dissolved substances in a solution change the velocity of light penetrating the solution and therefore the angle of the light as it enters a prism. This angle is then converted to specific gravity by the use of standard scales. The most important advantages of the refractometer over the urinometer is that it can provide an accurate measurement of specific gravity with only a few drops of urine and that the instruments in current use are temperature corrected, facilitating standardization. When large amounts of protein or glucose are present in the urine, certain corrections are necessary to ensure accuracy because of the significantly increased density of the solution. Maintaining the refractometer requires little more than cleaning the glass prism and cover after each use (quality control procedures are considered elsewhere in this chapter).

Specific gravity can be estimated with a dipstick available from one manufacturer (Multistix, Ames Company, Elkhart, IN, U.S.A.) (10). This colorimetric method provides a fast and convenient way to estimate the specific gravity of urine without using a refractometer (34). Specific gravity estimation by dipstick is an indirect method in that it measures only ionized particles in urine. It is based on the principle that a change in ionic concentration of a solution initiates a change in the dissociation constant (pK) of pretreated polyelectrolytes (polymethylvinyl ether/maleic acid) embedded in the dipstick pad. This change is then detectable by the use of a pH indicator dye (bromthymol blue). The polyelectrolyte is sensitive to the electrolyte ions in solution in the urine; when these ions increase in number, the pK of the dipstick pad is decreased, thus reducing the pH. This pH change is reflected by a change in the color of the indicator on the dipstick pad, i.e., the indicator, bromthymol blue, changes from blue to yellowish-green (more acidic), and this transformation is directly related to specific gravity.

Another method to assess specific gravity is by the use of the “falling drop” procedure. This method is useful for equipment that automates the reading of dipsticks (e.g., Yellow IRIS, International Remote Imaging Systems, Chatsworth, CA); it is a more specific measure of specific gravity than the one provided

by refractometry. A water-immiscible silicon oil that has a predetermined constant specific gravity is utilized (10). As the urine drops in a carefully designed column filled with this oil, the time it takes for a urine drop to fall a given distance is determined; the falling time is accurately measured by electronic means and is then directly converted to specific gravity units.

Ascorbic Acid

Vitamin C is a frequently administered pharmaceutical that often interferes with several of the usual dipstick reactions (e.g., glucose, bilirubin, leukocyte esterase, hemoglobin, and nitrite). For this reason, it became important to develop a convenient means of testing for its presence in the urine so that if interference was suspected, alternate testing for specific substances could be initiated. Erroneous results should be suspected when a microscopic sediment evaluation shows RBCs but the dipstick is negative for hemoglobin. The various dipstick reactions are based on the powerful reducing property of vitamin C and the change in indicator dye color when this reduction occurs (10).

Phenylpyruvic Acid

It is clinically important to determine whether phenylpyruvic acid is present in the urine of very young infants because its presence is diagnostic of phenylketonuria (PKU) (15). PKU is an inherited metabolic disease in which there is absence of an enzyme, phenylalanine hydroxylase, that is necessary for the conversion of phenylalanine, present in milk, to tyrosine. In the absence of this enzyme, metabolites of phenylalanine, such as phenylpyruvic acid, accumulate and cause brain damage. Treatment is dietary with removal of foods and milk containing phenylalanine. If undiagnosed and untreated, PKU results in permanent brain damage and mental insufficiency, thus early diagnosis by screening urine for phenylpyruvic acid is an important preventive measure. Strip testing is available for the detection of phenylpyruvic acid (Phenstix) and when present, phenylpyruvic acid reacts with a buffered ferric ion solution on the strip to form a bluish-green reaction; the color develops in 30 seconds and is compared with a color chart for an estimate of the concentration of phenylpyruvic acid (10).

Miscellaneous Tests

Several other tests may be performed on urine samples using various technologies (10,11,15). Melanin may be present in the urine in metastatic malignant melanoma, and several different methods for its detection are available. Serotonin, a metabolite of tryptophan, is found in the urine of some patients with tumors of argentaffin cells (carcinoid tumors). Testing urine for its metabolite, 5-hydroxyindoleacetic acid, is a valuable means of diagnosing this condition. In porphyria, a group of inherited enzyme-deficient diseases of hemoglobin metabolism, intermediary metabolic products may appear in the urine. Specific methods for the detection of these products (i.e., porphobilinogen, uroporphyrin, and coproporphyrin) are available and should be utilized when a clinical diagnosis of one of the porphyrias is suspected because of the presence of such symptoms as skin lesions, photosensitivity, recurrent acute abdominal pain, neurologic symptoms, or urine that changes color in light.

MICROSCOPIC URINALYSIS

Part of "28 - Urine"

A microscopic evaluation of the urinary sediment often produces valuable information that provides the clinician with a more specific diagnosis or an assessment of therapy that could not otherwise be obtained if only a physicochemical examination of the urine were performed (35). Recent controversy concerning the cost efficiency of the microscopic part of the urinalysis procedure has centered on the relatively low rate of positive results from sediments in which the chemical examination was negative (33,36,37,38). Without debating this issue, it should be the responsibility of each laboratory to establish its own guidelines for urine analysis that include when to perform a microscopic evaluation; these guidelines should be based on sound scientific evidence (8). As a general “well-population” screening procedure, urine microscopics are not particularly cost effective; conversely, in a hospitalized population, especially in patients on urology or general medical wards, this part of the examination has been shown to be very valuable.

The microscopic urinalysis procedure is quite simple and requires little additional equipment other than what is ordinarily found in any well-equipped laboratory, i.e., a calibrated centrifuge, a modern, clean binocular microscope, and a means of ensuring that strict QA procedures are followed (7). There is no substitute for experience and training on the part of the technologist or physician for accurate sediment evaluation (35). Constituents in the sediment can be varied, and accurate interpretation often depends on previous experience. Some have advocated not centrifuging the urine when doing a microscopic examination (a common practice in England), this author does not. Centrifugation of 10 or 12 mL of urine at a constant time of 5 minutes and a relative centrifugal force (RCF) of 400 to 500 is standard practice in the United States (11,25). The sediment is produced at the bottom of the centrifuge tube. Several commercial products offer a system for processing urine to obtain a sediment; these all yield a constant volume of sediment admixed with urine so that an aliquot may be decanted and then viewed with a microscope (25). As an example, if the initial volume of urine is 12 mL and the volume of the sediment remaining after decanting the supernatant urine is 1 mL, the resulting concentration of the sediment is 12:1. If a specific system is utilized, the use of one in which the concentrated sediment is pipetted into a constant-volume chamber on a microscopic slide is highly recommended. In such instances, this procedure ensures that the volume of the concentrated urine observed in the chamber will be constant from patient to patient and from day to day—an important consideration in QA (6). Because constant volumes of urine are used, when sediment elements are recognized, these can be quantitated on a volume basis (i.e., numbers per milliliter) rather than as numbers per microscopic field. Before the advent of constant-volume slide chambers, the usual procedure for microscopic examination of the urine was to place a drop of the concentrated centrifuged urine sediment onto a glass microscopic slide and observe it with a microscope. Great discrepancies

could be found in the quantitation of the elements because of the variability of the size of the urine drop and the weight of the coverslip. The use of a standardized system for this examination permits much greater consistency in the reporting of results.

Centrifugation at an RCF of 400 to 500 for 5 minutes produces a concentrated sediment in which all elements present may be easily found and are not artifactually distorted by compacting during centrifugation (8). Because most modern centrifuges can be adjusted for revolutions per minutes (rpm) but not for RCF, the following formula, which takes into consideration the radius of the centrifuge head, is used.

$RCF = 1.118 \times 10^{-3} \times \text{radius of centrifuge head (in cm} \times \text{rpm}^2)$

Normal Urinary Sediment

Recognition of sediment elements depends on having a "good eye," knowing what should be present in normal urine, and being able to define accurately and contrast the normal with the potentially abnormal. The appearance of some particles or elements in the urine may be normal. These can be blood cells, cells lining the urinary tract, mucous secretions of glands, cylindrical proteinaceous particles that have formed in the nephron (casts), crystals that have formed in the urine, and foreign cells (e.g., spermatozoa in a woman), microorganisms, or contaminants (Table 28.2) (13,25,35). Each of these constituents is discussed separately.

TABLE 28.2. CONSTITUENTS OF NORMAL URINE SEDIMENT

Cells	Crystals	Casts	Other
Blood cells	Acid urine	Hyaline	Mucus
Red	Amorphous	Granular	Sperm
White	Uric acid		Microorganisms
Epithelial cells	Neutral urine		Bacteria
Squamous	Calcium oxalate		Fungi
Urothelial	Hippuric acid		Contaminants
Renal tubular	Alkaline urine		Fibers
	Triple phosphate		Pollen
	Ammonium biurate		
	Calcium carbonate		

Blood Cells

Erythrocytes (RBCs) and leukocytes (WBCs) may be found in small numbers in the normal sediment. Both of these cells may pass through the glomerulus and enter the urinary flow. Quantitation of these cells over a period of time, such as 12 hours, is rarely performed today because of the great variation in cellular excretion from person to person and the difficulties associated with the collection of the urine and the counting technique (Addis count; a hemocytometer counting chamber is utilized) (39). A disease-free individual may eliminate as many as 750,000 RBCs and 1,750,000 WBCs in 12 hours.

RBCs

In urine sediment obtained from a disease-free person, small quantities of erythrocytes may be present (Fig. 28.1); these vary in number from none per hpf to as many as five per hpf (35). The usual normal sediment will contain an occasional RBC per every other field; counts greater than five per hpf should be thoroughly investigated and an explanation for the hematuria sought. Microscopically, the RBCs appear similar to those found in the peripheral blood, i.e., biconcave disks that have a pale faint orange color confirming their hemoglobin content (Fig. 28.2). In hypertonic urine, RBCs may be crenated and in hypotonic urine they may swell, become spherical, and, in time, burst, leaving only the cell membrane or "ghost" visible. Small globules of oil may simulate the appearance of RBCs. The oil is frequently from the skin of the microscopist. Oil droplets may be differentiated from RBCs based on varying size, absence of hemoglobin, and a spherical rather than biconcave disk shape.

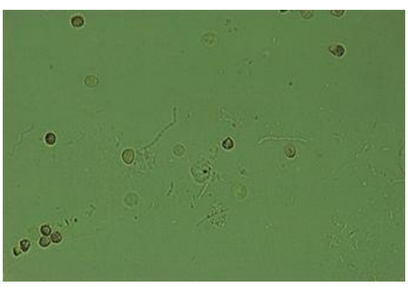


FIGURE 28.1. Red blood cells (RBCs) and bacteria in urine sediment. This field contains a scattering of RBCs and bacillary forms. This is slightly more than the usual number of RBCs one would expect to find in a high-power field and would generally warrant further investigation. Two leukocytes are also present in the center of the field. (Bright-field microscopy, ×160.)

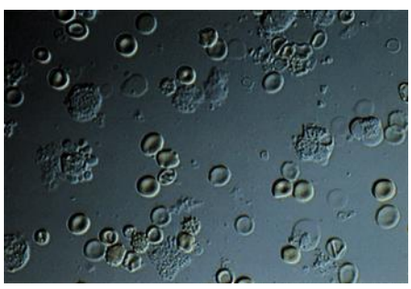


FIGURE 28.2. Polymorphonuclear neutrophils and red blood cells (RBCs) in urine. This unstained photomicrograph vividly demonstrates the biconcavity of RBCs and the multilobed nuclei and granular cytoplasm of the neutrophils. A few RBCs are crenated. (Interference-contrast microscopy, ×200.)

WBCs

WBCs are frequently found in the normal urine sediment, but are few and should not exceed five per hpf (13,25). Although all

the WBC types that appear in the peripheral blood may also be found in urine (i.e., lymphocyte, monocyte, eosinophil), the most common cell present is the PMN. In fact, it is unusual to find any leukocyte other than a PMN in a normal person's sediment. PMNs have a phagocytic function, are actively motile, and move by means of ameboid action with the extension of pseudopods. Leukocytes measure 10 to 20 µm in diameter, with lymphocytes being on the lower end of that range and PMNs at the upper. PMNs are readily recognized in urine because of their multisegmented nuclei and their granular cytoplasm in which brownian movement is often noted.

Staining the sediment enables the observer to identify PMNs more readily because the multilobed nucleus is clearly apparent and confusion with nonleukocytic cells, such as RTE cells, is less likely. Differential staining with Wright's or Giemsa stain provides an accurate means of identifying the various other leukocytes, such as lymphocytes and eosinophils (35). This becomes important when there is an increase in WBCs but is unnecessary in the normal sediment evaluation.

Epithelial Cells

Normal urine contains three main varieties of epithelial cells: renal tubular, transitional (urothelial), and squamous (35). These cells provide the lining for the urinary tract and comprise the tubules of the nephron. Table 28.3 lists some differentiating features of each type of epithelial cell.

TABLE 28.3. EPITHELIAL CELLS OF URINE

	Renal Tubular	Urothelial	Squamous
Origin	Nephron	Renal pelvis, ureter, bladder, proximal urethra	Terminal urethra vagina
Size (µm)	15–25	20–30	30–50
Shape	Polyhedral	Polyhedral, “tadpole”, spherical	Flattened
Other	Microvilli if from proximal tubule		

Renal Tubular Epithelial Cells

RTE cells are infrequently present in the urine sediment in the normal person (zero to one per five hpf). When observed, they are usually single but may also be found in pairs. If a microvillus border is present, their origin can be presumed to be the proximal tubule. Specific immunohistochemical identification may be made, when warranted, by means of acid phosphatase staining, because the RTE cells have a high intracellular content of this enzyme (35). Their shape is most often polyhedral, but they may also be somewhat flat, indicating that they originate from the loop of Henle. Their nucleus is usually eccentric but may be central; it is spherical with nucleoli readily apparent if there are no autolytic changes. In some disease states, primarily those that affect the tubular portions of the nephron, the number of RTE cells in the sediment may be significantly increased.

RTE cells are normally found in the urine because of the constant process of tubular cell renewal and regeneration. In a renal biopsy, tubular lining cells show frequent mitotic activity; the older cells are sloughed into the urine flow and may then be seen in the sediment after having been replaced by the newly fashioned cell. This type of cell regeneration occurs in the most proximal, rather than the distal, nephron.

Transitional Epithelial Cells

These cells (also called urothelial cells) constitute the lining epithelium for a large portion of the urinary tract and are frequently present in the sediment (zero to one per hpf) owing to their high turnover rate. Transitional epithelium is stratified and ordinarily several cell layers thick. Only the most superficial cells are found in the urine and these take three main shapes: spherical (Fig. 28.3), polyhedral, and “tadpole.” Transitional cells have the peculiar characteristic of being able to absorb water and thus swell to as much as two times their original size when they acquire a spherical contour (25). Polyhedrally shaped transitional cells may present difficulties in differentiation from RTE cells if they have no microvillus surface and have a central rather than an eccentric nucleus. The cytoplasm of transitional cells does not contain large quantities of acid phosphatases. Tadpole-shaped urothelial cells are frequently observed in urine. They probably derive from the mid-stratification layers of the transitional epithelium (13). Tadpole transitional cells appear in groups or pairs, as well as singly. Their nucleus is usually central, and they have a fusiform-shaped cytoplasm. Transitional cells increase in number in the urine in certain inflammatory conditions affecting the urinary tract.

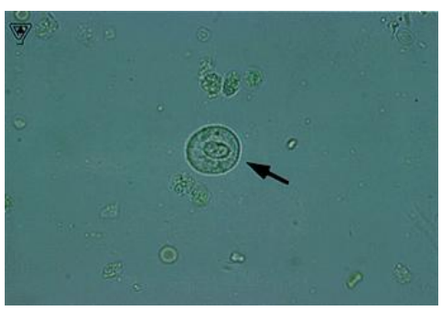


FIGURE 28.3. Transitional cell (*arrow*) and occasional red and white blood cells in urine. Note the spherical shape and central nucleus of this cell. (Bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

Squamous Epithelial Cells

Squamous epithelial cells are the easiest of all the epithelial cells to recognize in urine because they are large, flat, and frequently encountered (Fig. 28.4). In a well-collected midstream “clean-catch” specimen, there should be few of them (i.e., fewer than one per hpf). Numerous squamous cells in urine from a female patient usually indicate vaginal contamination. Squamous epithelium normally lines the terminal third of the urethra in both men and women, and it is also the epithelium that lines the vagina. Clean-catch urine collection is more difficult to achieve in women than in men, and contamination with vaginal fluid

and cells is common, especially when the patient has not been instructed previously on how to collect the specimen properly (8).

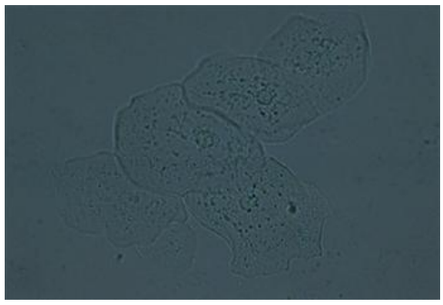


FIGURE 28.4. A group of squamous epithelial cells in urine. The cells are large and flat and have some granules in their cytoplasm. The central nucleus is approximately the size of a large lymphocyte. (Bright-field microscopy, ×160.)

Crystals

Finding crystalline forms in urine sediment is common if the urine has been allowed to stand after voiding, but this finding is infrequent in fresh urine. The formation of crystals is related to the concentration of various salts in the urine, which, in turn, is related to the patient's metabolic state, diet, and fluid intake as well as to the effects of the changes that occur in the urine after voiding (i.e., changes in pH and temperature, which change the solubility of salts in the urine and promote crystal formation). Because the kidneys play a major role in metabolite excretion and the maintenance of homeostasis, the end products of metabolism are found in high concentrations in the urine, and these tend to precipitate and form crystals (10). The pH of normal urine varies from acid to alkaline, and some crystals are associated with an acid pH, whereas others may be more commonly found in alkaline urine (25). There are relatively few crystals associated with a normal state, and the student is well advised to become aware of their various morphologic forms and characteristics. Most normally occurring crystals have little if any pathologic significance. The more common varieties are described here. In a later section of this chapter, several types of abnormal crystals are considered.

Uric Acid Crystals

Uric acid, a metabolic product of protein breakdown, is present in the urine in high concentrations and commonly produces a wide assortment of crystalline structures. Amorphous urates may be described as granular, birefringent, colorless to pale yellow crystals without a particular microscopic structure; they appear as fine granules when observed with the 10× or 40× microscopic objectives (Fig. 28.5). These crystals frequently develop when urine has been refrigerated, and they may be so plentiful that they obscure any other sediment elements present. When urine is prepared for microscopic examination, these crystals are noted as the compacted “pink button” of sediment at the bottom of the centrifuge tube. Most amorphous urates dissolve when alkali is added to the sediment or when the urine is warmed (to at least room temperature) after refrigeration.

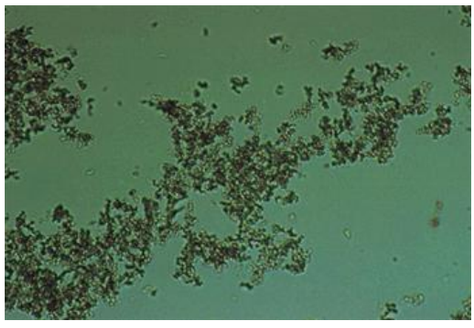


FIGURE 28.5. Amorphous urate crystals in urine. (Bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

Uric acid crystals are among the most pleomorphic of all urine crystals; they assume a variety of forms, including rods, cubes (Fig. 28.6), rosettes, six-sided plates, rhombi, and a whetstone shape. They are strongly birefringent and vary greatly in size. Uric acid crystals are soluble in alkaline solutions and insoluble in acid (15). They are colorless to pale yellow to pinkish brown. Uric acid crystals have been associated with renal stones, but their presence in the urine in the normal person is so common that this association has little if any clinical significance.

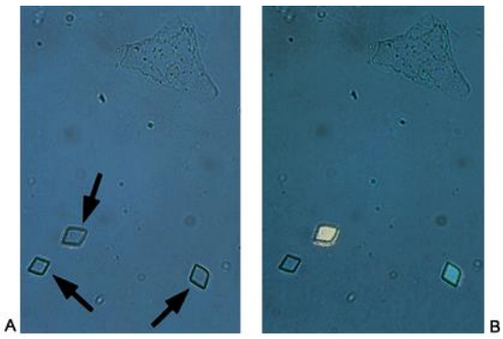


FIGURE 28.6. Uric acid crystals (*arrows*) and squamous cell. In this view, the urate crystals are rhomboid **(a)** and display anisotropism under polarized light **(B)**. (Bright-field microscopy and polarization, ×80.) (From the College of American Pathologists, with permission.)

Other urate salts may form crystals in the urine, i.e., sodium and potassium urates. These can be seen as colorless, needle-shaped crystals and brownish spherules, respectively. Addition of a drop of glacial acetic acid to a wet mount exhibiting the latter spheroids results in their transformation into more usual rhombic and whetstone plates.

Calcium Oxalate

Calcium oxalate crystals are the next most frequently observed in both acid and neutral urine (Fig. 28.7). The most common variant of this crystal is the dihydrate form, a colorless, octahedral crystal with the appearance of an envelope (35). Crystals of this type are found in abundance in the urine of otherwise normal individuals, especially after ingestion of ascorbic acid in large doses or foods rich in oxalic acid such as tomatoes or asparagus. Also intermittently encountered in otherwise normal urine are monohydrate forms, which appear elliptical or dumbbell shaped depending on whether viewed flat or on edge (Fig. 28.8).

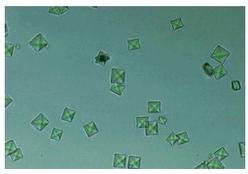


FIGURE 28.7. Calcium oxalate crystals, dihydrate forms. Their square shape and "star," "envelopelike" appearance are characteristic. (Bright-field microscopy, ×160.)

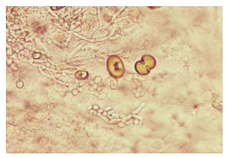


FIGURE 28.8. Bile-stained, calcium oxalate crystals, monohydrate forms. Note oval appearance when lying flat, dumbbell shape when on edge. From urine of severely jaundice patient. (Bright-field microscopy, ×160.)

Hippuric Acid Crystals

Although they are not often diagnosed in urine, hippuric acid crystals are also associated with a neutral pH. These crystals are usually seen as colorless, elongate prisms with pyramidal ends; they may also be thin and needlelike. They are birefringent and associated with diets high in fruits and vegetables that contain large quantities of benzoic acid (10,11).

Amorphous Phosphate Crystals

Crystals derived from phosphates are the most frequently observed crystals associated with alkaline urine. The most common of these is the amorphous phosphate crystal. Microscopically, these cannot be distinguished from amorphous urate crystals present in acid urine. They are fine, granular, birefringent crystals that produce a white precipitate in the bottom of the centrifuge tube. They may also obscure other sediment elements because they are often formed in large quantities.

Triple Phosphate Crystals

Triple phosphate (ammonium-magnesium phosphate) crystals are birefringent crystals that assume a "coffin-lid" appearance (Fig. 28.9). They are usually colorless and birefringent and vary greatly in size. The crystals may also be found in neutral urine and are soluble in acetic acid.

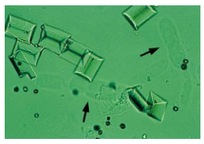


FIGURE 28.9. Triple phosphate crystals in urine. Hyaline casts (*arrows*) are seen in the background. (Bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

Occasionally found in otherwise typically alkaline urine are "star" forms or large polyhedral clear plates resembling the shapes of U.S. states (40).

Ammonium Biurate Crystals

Ammonium biurate crystals have a peculiar crablike or "thorn apple" shape (Fig. 28.10). They are yellowish brown and often demonstrate radial or concentric striations in central spheres with radiating "arms" or spicules. They are usually found in urine with a neutral pH and are soluble in sodium hydroxide. They are rarely encountered in normal urine.

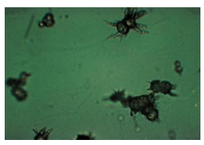


FIGURE 28.10. Ammonium biurate crystals in urine. These "crab-shaped," spiculated crystals are characteristic and are associated with alkaline urine. (Bright-field microscopy, ×400.) (From the College of American Pathologists, with permission.)

Calcium Carbonate Crystals

Calcium carbonate crystals are small spherules of dumbbell-shaped birefringent crystals found in alkaline urine (Fig. 28.11). Because of their small size, they are often mistaken for bacteria. A ready means of differentiation is by the use of polarization, since bacteria are not birefringent. These crystals dissolve in acetic acid and, in the process, effervesce.

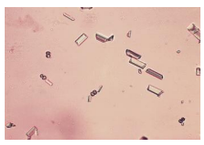


FIGURE 28.11. Dumbbell-shaped calcium carbonate crystals shown here with small triple phosphate crystals (Bright-field microscopy, ×160).

Casts

Urine casts are defined as cylindrical microscopic structures that form in the distal nephron and occur in the urine in both health and disease (41). These "cylinders" are composed primarily of a

kidney-specific protein that is produced only in the distal tubule and collecting duct of the nephron (THM) (42,43). As THM is excreted by the cells lining the distal tubules and collecting ducts, events occur in the glomerular filtrate in the lumen of the nephron that induce this soluble protein to become insoluble and to form thin, proteinaceous bands that may then coalesce or be molded into a cast. The cast formed represents an exact model of the shape of the lumen of the renal tubule in which it originated, much as a mold of gelatin assumes the exact shape of the container in which the liquid gelatin is placed. In experimental studies, changes in particular physical factors are necessary to induce soluble THM to become fibrillar. These include changes in urine flow, ionic strength, concentration, salinity, and pH of the glomerular filtrate. It is likely that in both normal and disease states, similar changes occur in human urine, especially in physiologic stress, which induces previously soluble THM to become insoluble and to form fibrils, which then coalesce into casts.

The classification of renal casts is simple and based on morphologic appearance (Table 28.4) (41). In the normal state, only two varieties of casts appear in the urinary sediment: hyaline casts and granular casts (18). Any additional cast forms must be considered “abnormal” and are most often associated with generalized metabolic or intrinsic renal diseases. Each cast type is discussed separately in detail. Casts, in contrast to cells, are quantitated in the sediment as number per low-power microscopic field (lpf).

TABLE 28.4. CLASSIFICATION OF URINARY CASTS

Acellular Casts	Cellular Casts
Normal	Normal
Hyaline	None
Granular	None
Abnormal	Abnormal
Hyaline	RBCs
Granular	WBCs
Waxy	Epithelial (RTE)
Pigment	Fatty/oval fat bodies
Fatty	Bacterial/fungal

RBC, red blood cells; WBC, white blood cells; RTE, renal tubular epithelial.

In the normal person, the presence of small numbers of hyaline or granular casts in the urine is a frequent occurrence and does not ordinarily present diagnostic difficulty or necessarily imply renal disease. The microscopist usually scans the sediment using the 10× objective, and occasionally casts are found, somewhere on the order of one or two per 10 lpf. These two forms of casts have a low refractive index and are therefore somewhat difficult to see with the ordinary light microscope unless the contrast is enhanced. Stopping down the iris diaphragm on the condenser while lowering it and keeping the light intensity the same will produce the optimal contrast for this observation. A quick scan of the entire microscopic slide will reveal whether any casts are present and their numbers; subsequent specific identification can be made using the 10× or 40× lens.

Hyaline Cast

This is the most frequently observed cast in the urine. It is typically translucent and somewhat difficult to see because of its low refractive index. Careful inspection reveals a smooth outer perimeter and a matrix that is most often smooth but may be wavy (Fig. 28.12) (44). Occasional granular inclusions may be present in the matrix, and sometimes a cell or two may also be seen. The cast may have one end that, instead of being curved, is drawn out into a “tail” or point. In the past, these casts with tails were called *cylindroids*; this term is considered antiquated and is not commonly used today (Fig. 28.13).

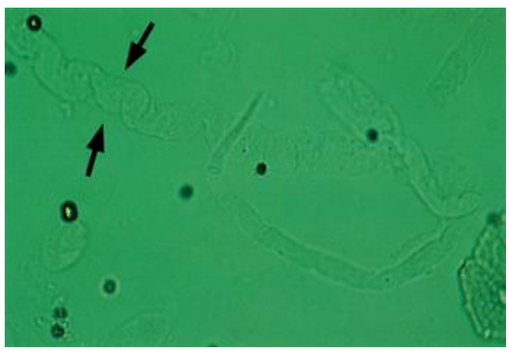


FIGURE 28.12. Hyaline casts. These translucent, colorless proteinaceous structures (*arrows*) are frequently found in normal urine sediment. Their low refractive index makes them difficult to discern using ordinary bright-field microscopy (× 100). (From the College of American Pathologists, with permission.)

FIGURE 28.13. Urine cylindroid, a hyaline cast with a tail. The term cylindroid is antiquated. (Interference-contrast microscopy, × 160.)

Staining the sediment or using phase microscopy (45) is very helpful in the identification and quantitation of all casts but is especially useful for hyaline casts in particular; with ordinary bright-field microscopy and an unstained sediment sample, hyaline casts are somewhat difficult to find.

The effects of stress and strenuous physical exercise on the production and appearance of hyaline casts (also granular casts) in the urine have been well documented (18). Running a mile race or participating in a vigorous physical sport may induce the production of a large number of these casts (46). Stressful emotional situations, such as having to take an examination, can invoke the same phenomenon. Despite obsolescence of the term

cylindroids, the term *cylindruria* lives on. It simply connotes increased numbers of casts in urine. When a patient has undergone physical or emotional stress within the previous 24 hours, cylindruria should not be considered pathologic. In such instances, the urine reverts to a normal state within 24 to 48 hours if the stressful situation or the physical exercise has stopped (18).

Granular Cast

This cast may also be observed in increased numbers in the urine if the patient has been involved in an emotionally stressful situation or has undergone recent strenuous physical exercise (18,35). Compared with hyaline casts, granular casts are found in a ratio of approximately four hyaline to each granular. On removal of the stress or ending the exercise, the number of granular casts in the urine returns to normal within 24 to 48 hours. The reasons for the increased production of these casts in stress or exercise are unknown. Also unknown is the reason why granular casts sometimes appear in the urine of patients on diets rich in carbohydrates.

Granular casts have a higher refractive index than do hyaline casts and are therefore easier to find. They are also cylindrical, although some may have "tails," and have a smooth, well-defined perimeter. Generally, in the normal person, granules covering the cast surface are small and regular (Fig. 28.14). The origin of these granules in the normal person appears to be, at least in part, intracellular lysosomal particles that are expelled into the urine as metabolic products of the renal tubular epithelium (44,47). When in the urinary flow, the lysosomal granules are incorporated into a preformed hyaline cast matrix and thereby transform a previously smooth-surfaced hyaline cast into a granular cast.

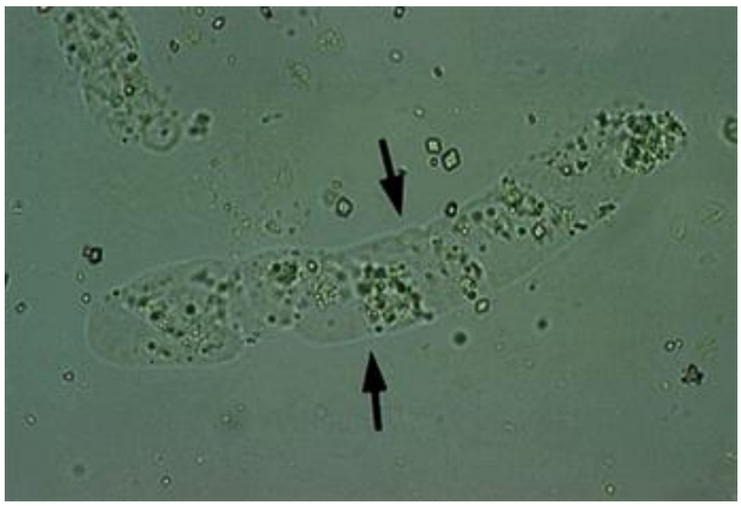


FIGURE 28.14. Granular cast. This type of cast is fundamentally a hyaline cast to which granules have attached along its surface. In the example shown here (*arrows*), the granules do not cover the entire surface of the cast but are relatively evenly dispersed. (Bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

Mucus

It is thought that various glands that line the genitourinary tract, such as those in the prostatic urethra and urinary bladder, secrete this mucopolysaccharide into the urine. In addition, recent immunologic studies showed that at least some of the mucus present in the urine is in fact THP, a specific immunoprotein secreted exclusively into the nephron by the distal tubule and collecting duct lining cells (43). The clinical significance of THP in the urine is unknown. When large quantities of mucus are present in the sediment, vaginal secretions may contaminate the specimen.

Mucus is a low-density substance that is often difficult to see unless sufficient contrast has been obtained with the microscope. It appears as thin, fibrillar, wavy, fiberlike bands or lines across the field that have no particular form or shape (Fig. 28.15). The inexperienced observer may sometimes confuse mucus with a hyaline cast because of coalescence of the individual ribbons into what may at first glance appear as a cylindrical object. Mucus has a low refractive index and is not birefringent. Sometimes cells or microorganisms may get caught in it.

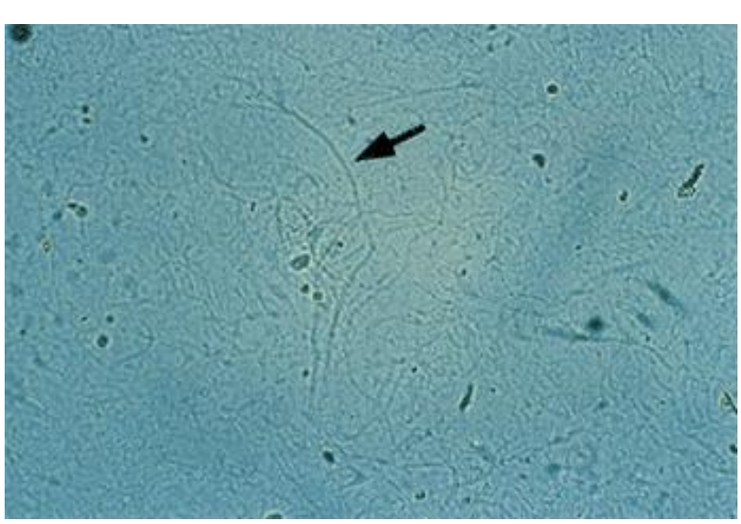


FIGURE 28.15. Mucus threads in urine. Thin, fibrillar mucus is a common feature of urine sediment and should not be confused with cast structures because these threads do not acquire a cylindrical form. (Bright-field microscopy, ×112.) (From the College of American Pathologists, with permission.)

Microorganisms

Bacteria (Fig. 28.1) and occasionally yeast (Fig. 28.16) are often found in what is considered normal urine (48). Theoretically, at

least, microorganisms should not be part of the normal sediment. However, they are commonly found in normal urine because collection of the sample is performed without the use of adequate or effective precautionary techniques that would ensure that no (or few) organisms contaminate the specimen. Additionally, collection systems frequently used to collect and transport the urine for examination are not sterile and may also serve as a source of microbiological contamination.

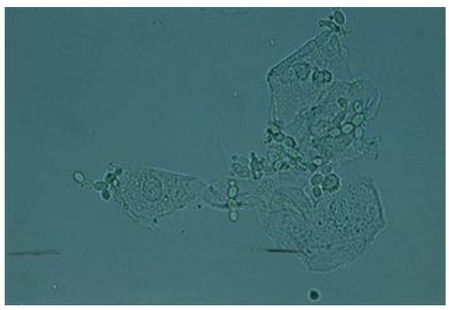


FIGURE 28.16. Budding yeast forms and squames in urine. When not accompanied by leukocytes, yeast forms are most likely growing contaminants, rather than pathogenic. They appear as spherical to elliptical transparent microorganisms approximately the same size as a red blood cell, with which they may be confused. (Bright-field microscopy, ×160.)

Urine is an excellent culture medium in which microorganisms grow with great rapidity. Therefore, when bacteria or yeast forms are present in the sediment and are unaccompanied by PMNs, they are most likely the result of contamination of the specimen. The bacteria may be either cocci or bacilli; yeasts often show budding but no hyphae.

Having stated that normal urine should ordinarily be sterile, one qualifying statement may be made: bacteria normally inhabit the distal urethra. Therefore, should a sample consist of the first part of the voided specimen, rather than a midstream collection, one might expect to find a few bacterial forms in the sediment. This occurrence is more common in women than men because a urine sample obtained from a female patient is more likely to be contaminated than one received from a male patient.

Spermatozoa

Spermatozoa are frequently found in male urine but may also be present in female urine (Fig. 28.17). Because urine is naturally spermicidal, the sperm are usually nonmotile when found, except when ejaculation and/or intercourse occurred very recently. The presence or absence of spermatozoa in the urine of a woman has been used as evidence in rape cases but is not a reliable determinant of the actual time of the rape.

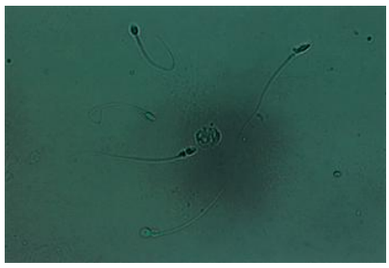


FIGURE 28.17. Spermatozoa. Several sperm are present around the neutrophil in the center of the field. Sperm can normally be found in urine from both men and women. (Bright-field microscopy, ×200.)

Contaminants and Artifacts

Contaminants, especially fibers, appear in abundance in urine sediments. The importance of the correct recognition of these substances, fibers (Fig. 28.18), pollen grains, plant cells, starch granules (Fig. 28.19), or other substances, is that they must not be mistaken for actual constituents of the sediment and result in an incorrect diagnosis. To avoid these errors, it is important to keep in mind the following: (a) most fibers and other contaminants have a high refractive index and are easily seen, whereas most actual constituents of the sediment are much more difficult to detect; (b) many artifacts and contaminants are far larger than any sediment element and can be differentiated by size alone; and (c) plant and synthetic fibers nearly always show birefringence with polarized light (Fig. 28.18), whereas urine casts do not (except fatty casts, which do not look like fibers). In general, the best and most effective means of avoiding a misdiagnosis owing to a contaminant is to be certain about the appearance of and the ways to identify normal sediment elements. If what is seen does not fit the normal pattern, the structure is either abnormal or a contaminant.

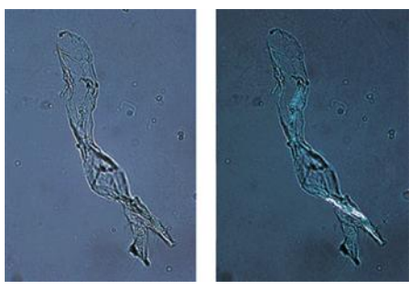


FIGURE 28.18. Fiber in urine. Fiber artifacts are frequently present in urine sediment and should not be confused with “real” elements that may be pathologic. Polarization will usually reveal their true characteristic as they display anisotropism. (Bright-field microscopy and polarization, ×80.) (Reproduced by permission from the College of American Pathologists.)

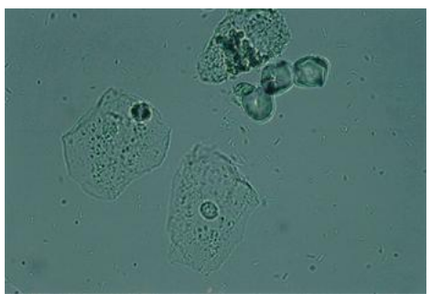


FIGURE 28.19. Starch granules in urine. These highly refractile, irregularly shaped, dense granules are easily identified in urine and are commonly found because starch is the material used to powder latex gloves and to dry the skin (i.e., baby powder). They are birefringent. In this field, three squamous cells and bacteria are also present. (Bright-field microscopy, ×160.)

Abnormal Urinary Sediment

When the normal contents of urinary sediment can be recognized, it is not particularly difficult to identify deviation from normal and make a diagnosis of abnormal. Most frequently, abnormal sediment is defined by an *increase* in the number of cells or casts present, not by an alteration in the type of cell or cast. However, different sediment elements are found in the urine in various disease states, and each of these is discussed in this section (Table 28.5). It should be noted again that there are often subtle differences between normal and abnormal, and when these differences depend on quantitative or semiquantitative measurements, as in cell and cast counting, the examiner must be especially careful to produce results that are consistent, accurate, and reproducible.

TABLE 28.5. CONSTITUENTS OF ABNORMAL URINE SEDIMENT

Cells	Crystals	Casts	Other
Blood cells	Acid urine	Acellular	Parasites
Red	Bilirubin	Hyaline	
White	Cholesterol	Granular	
Epithelial cells	Cystine	Waxy	
Renal tubular	Leucine	Fatty	
Oval fat bodies	Tyrosine	Cellular	
	Hemoglobin	RBCs	
	Drug-related	WBCs	
		Bacterial	
		Epithelial	
		Crystal	

RBCs, red blood cells; WBCs, white blood cells.

Blood Cells

Erythrocytes and leukocytes are increased in number in urine in many diseases. The patient may report having “smoky” urine if hematuria is present or of having a foul-smelling, “white” urine if pyuria is the cause. In each instance, a careful examination of the sediment is needed with a quantitation of the number and type of blood cells present. RBCs or WBCs in quantities greater than five per hpf should be considered abnormal (35).

RBCs

Hematuria accompanies many clinical conditions (Fig. 28.20). However, the most frequent cause is not pathologic but represents contamination of the urine with menstrual blood. One of the first questions to ask when hematuria is encountered in an adult premenopausal female is whether she is having her period. A subsequent urinalysis, when the urine has been collected after appropriate patient instruction and preparation, will produce a normal number of red blood cells in the sediment if menstrual contamination is the cause of the hematuria.

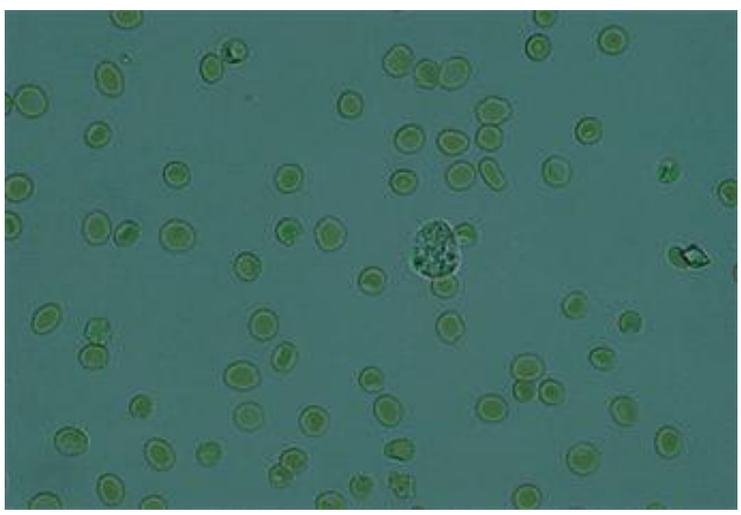


FIGURE 28.20. Hematuria. This field shows a large number of red blood cells (RBCs) accompanied by one polymorphonuclear neutrophil in the sediment. Note that the RBCs appear to be regularly shaped, biconcave disks. (Bright-field microscopy, ×160.)

Another common cause of hematuria is related to exercise. Contact sports such as football, boxing, wrestling, and ice hockey may cause trauma to the kidney or bladder and induce

hematuria. This is usually self-limited and disappears after a day or 2. The terms *pseudoathletic nephritis* and *sports anemia*, coined in the 1950s, have been used to describe the findings in patients with hematuria, and sometimes hemoglobinuria, after strenuous physical exercise (49,50). Before it was recognized that this was an unusual, but not pathologic, response to exercise, many dollars were spent on medical workups of patients presenting with these symptoms. It has been shown that the hematuria in these patients is induced by the stress of the exercise and/or the trauma associated with the constant pounding of bladder urine against the bladder wall, irritating and traumatizing the lining to the point that bleeding into the urine occurs. In addition, the continual pounding of the feet of runners against the pavement causes the RBCs in the foot capillaries to be crushed and rupture, thus releasing their hemoglobin into the circulation where it is then cleared by the kidneys and found in the urine. Abatement of the exercise in these cases produces normal urine within 48 hours.

Medical causes of hematuria include glomerulonephritis, pyelonephritis, cystitis, and neoplasms of the kidney and urinary tract, among others. The hematuria may be microscopic, usually less than 20 RBCs per hpf to gross or macroscopic, more than 50 RBCs per hpf. In bladder or renal cell carcinoma, microscopic hematuria is often an early and important finding that is manifested during a routine laboratory workup in an asymptomatic patient.

Dysmorphic erythrocytes have been described when the hematuria originates in the kidneys (51). These deformed RBCs become an important indicator of intraglomerular hemorrhage and are often overlooked by the inexperienced observer. When dysmorphic RBCs are seen, they are easily recognized. Dysmorphic RBCs have been bent out of shape and into peculiar angles but are not disrupted. In other instances, dysmorphic red cells exhibit one or more budding excrescences such that some manifest a “Mickey Mouse” configuration (Fig. 28.21a and Fig. 28.22). The peculiar contortions of the red cells are thought to be caused by squeezing through a damaged GBM.

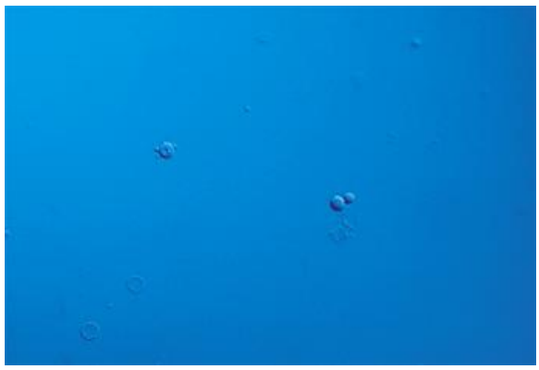


FIGURE 28.21. Dysmorphic erythrocytes in urine sediment, interference-contrast microscopy (×160). Two red blood cells (RBCs) are shown in the center of the field. Each is markedly deformed with buds projecting from their surfaces. In one of the RBCs, the buds are small, whereas in the other RBC, a single bud can be seen.

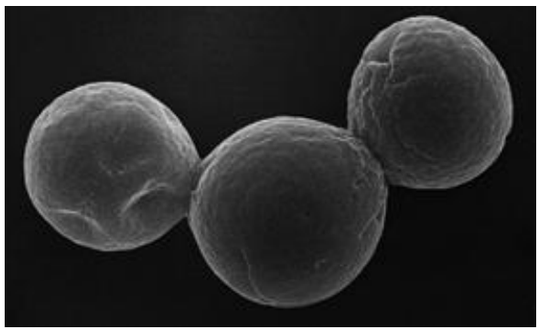


FIGURE 28.22. Scanning electron micrograph of a dysmorphic erythrocyte in urine (“Mickey Mouse” cell) (×7,200).

WBCs

Leukocytes in the sediment in numbers greater than five per hpf are considered pathologic. This number is not absolute and may vary from laboratory to laboratory, depending on the established normal values. Norms for WBCs in urine may be slightly higher for females. WBCs appear in increased numbers in the urine primarily when an inflammatory process affects the urinary tract. Pyuria is a common finding in cystitis, urethritis, and pyelonephritis. The white cells are predominantly PMNs, but careful study has shown that lymphocytes, monocytes, and eosinophils may also be present (Fig. 28.23). PMNs are actively motile. They enter the urine in the kidney or in the remainder of the urinary tract by actively migrating to the site of the inflammatory process moving through the GBM or tubular basement membrane (TBM) in the kidney or across the lining epithelium of the renal pelvis, ureter, bladder, or urethra (35).

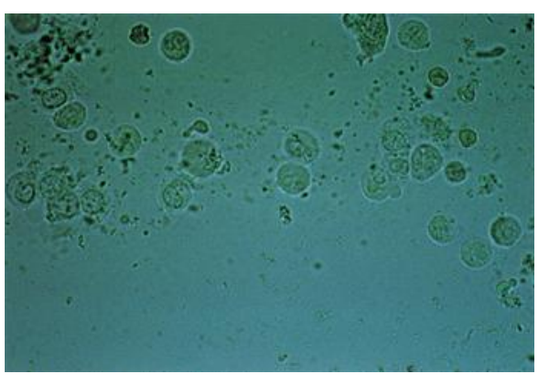


FIGURE 28.23. Polymorphonuclear neutrophils (PMNs) in urine. Across the center of the photomicrograph are many white blood cells; most are PMNs, but there are occasional mononuclear forms as well. Bacteria and a few erythrocytes are present in the background. (Bright-field microscopy, ×160.)

Phase microscopy greatly facilitates the recognition of PMNs because their characteristic spherical shape, granular cytoplasm in which brownian motion is often evident, and multilobed nuclei are readily seen (Fig. 28.24) (38). PMNs are actively phagocytic and serve the body well in protective functions as destroyers of microorganisms and as scavengers and destroyers of microorganisms (Fig. 28.25). “Glitter” cells are PMNs in which

the intracytoplasmic vacuoles are large and in motion and appear to be twinkling when observed in the fresh state. Supravital staining of urine (e.g., Sternheimer-Malbin) provides a means of specific identification of PMNs because their granules and individual nuclear lobes are prominently displayed (5).

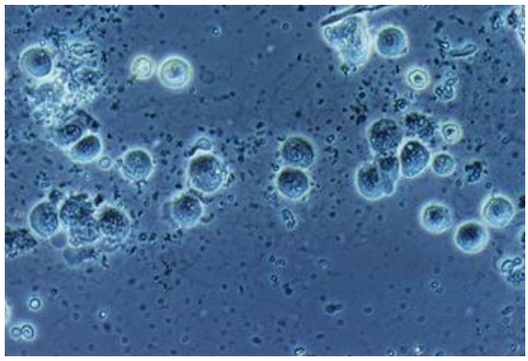


FIGURE 28.24. Phase microscopy (×160) of leukocytes in urine. This photomicrograph represents the same field as seen in Fig. 28.23. Here, the cytoplasmic granules in the polymorphonuclear neutrophils and the multilobed nuclei are prominent.

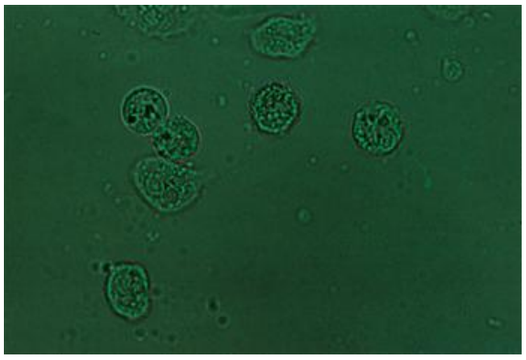


FIGURE 28.25. Glitter cells. These polymorphonuclear neutrophils (PMNs) display amoeboid motion. PMNs may show active motility, as is seen in the large cell near the center of the field. They do this by extending pseudopods from their cytoplasm. (Bright-field microscopy, ×250.)

In interstitial nephritis, an infrequent condition associated with allergic reactions to therapeutic agents, eosinophils are often present in the sediment in greatly increased numbers (52,53). Eosinophils may not be readily identified without the use of special stains, such as Wright's or Giemsa, or a specific stain for eosinophils (Hansel's stain). Eosinophiluria was recently reported to be associated with atheroembolic renal disease, a condition in which renal emboli occur in patients with severe atherosclerosis of the aorta (54).

Epithelial Cells

Epithelial cells may occur in the urine sediment in increased numbers in pathologic conditions. For practical purposes, only the presence of increased numbers of RTE cells is of clinical significance. However, it is important to recognize that large numbers of squamous cells are generally a sign of contamination of the specimen by vaginal fluid, whereas an increased number of transitional cells, if not accompanied by PMNs, most often indicates that some type of medical manipulative procedure involving the urethra and/or bladder was recently performed, such as catheterization.

RTE Cells

RTE cells are increased in the urine when there has been damage to the renal tubules and these cells are sloughed. This increased sloughing is most often accompanied by cylindruria, especially the presence of RTE casts in the urine. The specific etiologies for increased numbers of RTE cells in the urine are varied and multiple; renal tubular necrosis (55) heavy metal poisoning, cytomegaloviral, renal transplantation (56), aminoglycoside nephrotoxicity (57), and renal vein thrombosis are among the more common causes. Tubular lining cells are 15 to 25 µm in diameter and are usually polyhedral and contain an eccentric single nucleus that is approximately the size of a mature lymphocyte; they must be differentiated from similar-appearing transitional cells, which can be a difficult task.

Generalized viral diseases, such as viral hepatitis and cytomegaloviral disease, tend to affect the RTE cells and cause their sloughing (35,58). Common diseases, such as the common cold and measles, produce similar damage to the RTE cells, resulting in large numbers being sloughed into the urine. Tubular lining cells are active metabolically and may show bile staining because bilirubin is actively excreted by the kidney when the patient is jaundiced and has severe liver disease. Heavy metals such as lead, bismuth, antimony, and mercury, when ingested in toxic amounts, produce tubular cell death, resulting in increased numbers of these cells in the sediment. Lead toxicity induces an acid-fast inclusion body to form in some RTE cells, which, if recognized in the sediment, is of great diagnostic assistance in the identification of the clinical problem. Viruses may also induce the formation of inclusion bodies in RTE cells, best demonstrated in herpes, cytomegaloviral disease, and measles.

Oval Fat Bodies

Oval fat bodies (OFBs) (Fig. 28.26) are RTE cells in which globules of lipid have become visible (35). These deformed cells are frequently associated with a nephrotic syndrome but may be seen in various diseases that affect the tubule portion of the nephron. The lipid globules vary in size, are usually spherical, and, because cholesterol is a major component (rather than neutral fat), demonstrate anisotropic properties (Fig. 28.26) by showing a Maltese cross pattern (59). OFBs may occur singly or in pairs or groups.

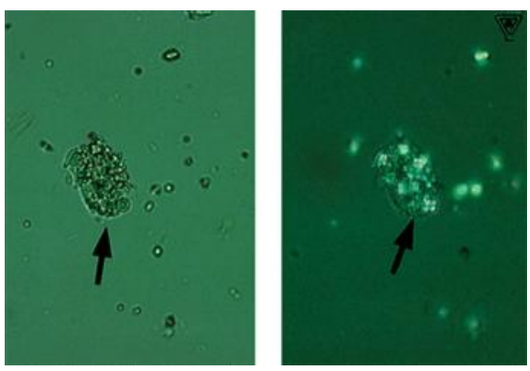


FIGURE 28.26. Oval fat bodies in urine. This renal tubule epithelial cell contains spherical globules of lipid (*arrow*), which, when polarized **(right)**, show a Maltese cross pattern. (Bright-field microscopy and polarization, ×160.) (From the College of American Pathologists, with permission.)

Transitional Epithelial Cells

These cells are rarely of pathologic significance unless there are large numbers present; even then, the usual etiology is that the patient has undergone a recent urinary tract procedure, such as a cystoscopy. In these circumstances, the cells may be observed in groups because of traumatic removal of the epithelium during the procedure. Inflammatory conditions of the bladder and urethra,

in particular, may give rise to increased numbers of urothelial cells in the sediment. However, they are usually accompanied by large numbers of PMNs because inflammation causes a degeneration and sloughing of transitional lining cells of the bladder or urethra. The morphology of these cells is as previously described, and most assume either spherical or caudate shapes. On occasion, irregular forms of transitional cells are observed in the sediment. Some of these cells may be larger than normal and irregularly shaped and have huge and irregular nuclei. Because bladder cancer is relatively common and may be suspected, an additional sediment sample should be obtained for submission to the cytology laboratory for analysis (60,61).

Squamous Epithelial Cells

The presence of large numbers of squamous cells in the urine usually represents contamination from vaginal contents. If accompanied by many bacteria, this is especially true. In patients with rectovesical or rectovaginal fistulas, many squamous cells may be seen in conjunction with fecal matter and bacteria. One form of squamous cell that has recently been identified is the “clue” cell. It has been shown that this cell, a squamous cell with bacillary forms adhering to its surface, is a frequent finding in bacterial vaginosis, an inflammatory condition of the vagina caused by the bacillus *Gardnerella vaginalis* (25). Clue cells should be diagnosed and reported when present (Fig. 28.27).

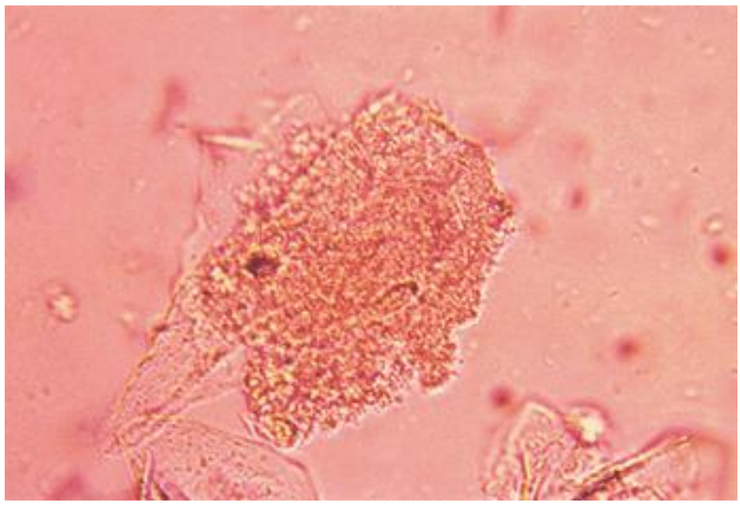


FIGURE 28.27. Clue cell. Note adherent surface bacteria (×160).

Abnormal Crystals

Of primary importance in the recognition of abnormal crystals is the fact that they are found only in urine of acid or neutral pH. For definitive crystal diagnosis, chemical analysis is performed. A few of the more frequently seen varieties are discussed in detail.

Cystine

These crystals are usually flat, six-sided plates that are clear and colorless (Fig. 28.28) (35). The sides of the hexagon are not necessarily equal in length. Polarization demonstrates a negative birefringent pattern (62). Sometimes, in very fresh urine, the hexagonal shapes are incomplete, and aberrations of these forms may be observed. The crystals may be layered or joined. They may be distinguished from six-sided uric acid crystals in most instances by the fact that in the latter, two of the opposing sides are clearly elongated. Although both cystine and uric acid crystals are soluble in ammonia water, only cystine crystals are soluble in dilute hydrochloric acid.

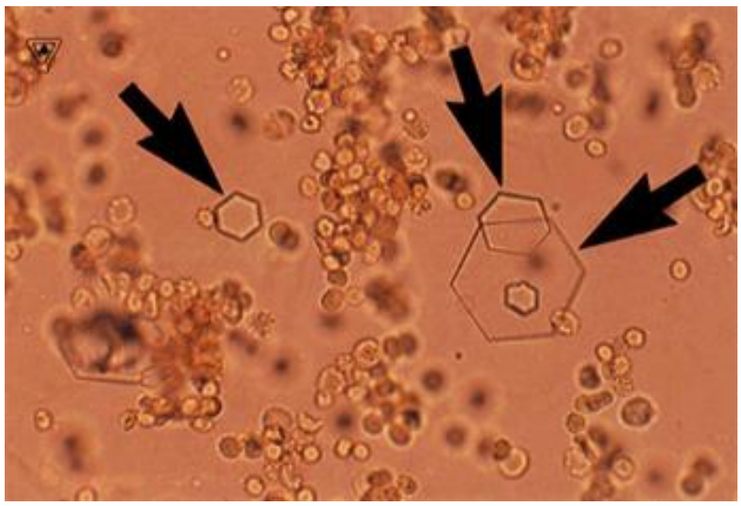


FIGURE 28.28. Cystine crystals in urine (*arrows*). These are flat, colorless, hexagonal crystals. Numerous red blood cells are present in the background. (Wright's stain, bright-field microscopy, ×400.) (From the College of American Pathologists, with permission.)

Cystine crystals appear in the urine in an inherited disorder that prevents the reabsorption of the amino acid cystine by the epithelial cells of the nephron and causes cystinuria. This disease is one of the more common inherited diseases and is expressed equally in each gender as an autosomal-recessive trait. Cystinuria also occurs in diseases affecting the liver (Wilson's disease) and in other diseases of the renal tubules.

Cholesterol

These crystals are most frequently found after refrigeration of the urine; lipid appears in globular form in urine at body temperature but crystallizes on refrigeration. Cholesterol crystals are large, flat, colorless thin plates that assume rectangular shapes of various diameters (Fig. 28.29) (35). One peculiar feature is that

the crystals often show a “staircase” pattern or contain an indentation of one or more of the corners of the rectangular plate. They are birefringent and often overlap or appear together (Fig. 28.30).

FIGURE 28.29. Cholesterol crystals in urine. These colorless, flat plates are characteristic because of their rectangular shape and “staircase” patterns. In the background are many red blood cells and an occasional leukocyte. Some of the red cells demonstrate dysmorphic shapes. (Bright-field microscopy, ×160.)

FIGURE 28.30. Cholesterol crystals, polarized (same as Fig. 28.29). These crystals are anisotropic and often overlap. (Polarization, ×160.)

Cholesterol crystals are occasionally found in the urine sediment of patients with lipoid nephrosis, membranous glomerulonephritis, polycystic renal disease, or the nephrotic syndrome (edema, lipiduria, lipidemia, and hypercholesterolemia). In lymphatic obstruction caused by neoplasms or parasitic disease (filariasis), chyluria may be a consequence if the dilated lymphatics rupture into the kidney or urinary tract.

Leucine

Leucine crystals are rarely found in the urine sediment. They appear as spheroids with a central nidus surrounded by concentric bands with apparent radiations originating from the center and crossing the bands (35) (Fig. 28.31). They are brownish yellow and highly refractile and show birefringence. Patients with severe liver disease, such as fulminant hepatitis and advanced cirrhosis, as well as patients with congenital metabolic conditions, such as maple sugar urine disease, excrete these crystals in their urine. They may occur in association with tyrosine crystals. Microscopically, they should not be confused with prostatic concretions or pollen grains, and if there is doubt, chemical testing of the urine for leucine is warranted.

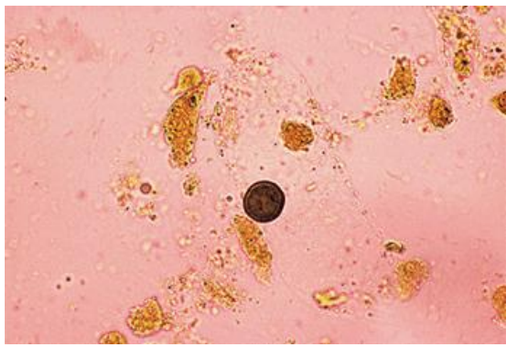


FIGURE 28.31. Leucine crystal in urine of patient with advanced liver disease. Note both radial striations and concentric rings. (×160).

Tyrosine

This crystal is one of the most difficult to see because of its relatively low refractive index; it is birefringent. Tyrosine crystals occur as fine, delicate, colorless or pale yellow elongated needles that may appear singly or in clusters or sheaves (Fig. 28.32). Microscopically, the sheaves look like they have a black central focus, which is a useful point in diagnosis. Tyrosine crystals in the urine are uncommon, even in severe liver disease.

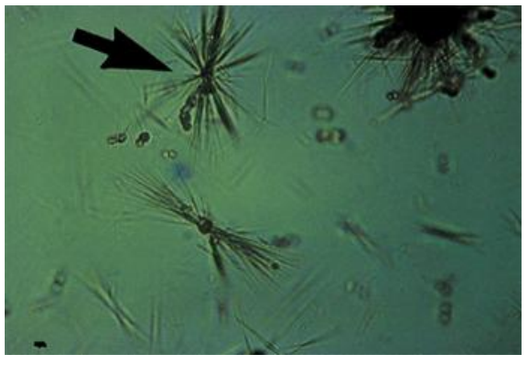


FIGURE 28.32. Tyrosine crystals in urine. These fine needlelike crystals commonly occur in clusters or sheaves. The black central focus depicted in the photomicrograph (*arrow*) is a useful characteristic for diagnosis. (Bright-field microscopy, ×320.) (From the College of American Pathologists, with permission.)

Bilirubin

This brownish red crystal appears in the sediment as amorphous granular material or as short needles. It may be difficult to distinguish from amorphous urates or phosphates without confirmatory chemical testing. For these crystals to form, high levels of

direct-reacting bilirubin need to be reached in the blood and urine, usually as a result of obstructive jaundice (Fig. 28.33).

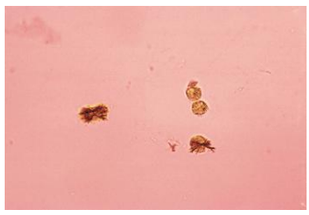


FIGURE 28.33. Brownish yellow bilirubin crystals attached to free renal tubular cells or neutrophils. Urine of patient with advanced metastatic liver disease. Crystals form a sheaf in left cluster (×200).

Hemoglobin

Granules of hemoglobin and hemosiderin may be found in the urine sediment on occasion. When concentrations of hemoglobin in the urine are excessive, as in massive hematuria with RBC lysis or in severe hemolytic anemia and crush syndrome after trauma, hemoglobin crystals form and appear as tiny, refractile, reddish brown, amorphous granules. These may be difficult to distinguish microscopically from other granular crystals, and chemical testing along with staining for hemoglobin (e.g., Prussian blue) may be required.

Pharmaceutical Agents

Many pharmaceutical products may, on occasion, produce urinary crystals. Perhaps the most common of these are the dyes associated with radiologic visualization of the urinary tract (i.e., meglumine diatrizoate, Hypaque, and Renografin). These crystals produced by these dyes are colorless crystals and highly refractile and assume needlelike or elongated rectangular shapes with indented or notched ends. The urine specific gravity is frequently more than 1.040 in patients with these crystals (10). A careful history will reveal the true origin of the crystal.

Ampicillin

The crystal produced in the urine by this drug is an elongated, needlelike, colorless crystal that is found in acid urine. It may be difficult to distinguish from radiographic dye crystals, and a history aids in making a correct diagnosis.

Sulfa Drugs

Now that most sulfanilamide drugs manufactured are water soluble, finding their crystalline products in the urine is uncommon but does occur infrequently. Sulfadiazine is the most commonly found. The urine is acid, with the pH usually less than 6. The sulfa crystals each assume various needlelike shapes, are birefringent, and are most often seen as bundles of brownish yellow “sheaves of wheat” formations with eccentric binding (Fig. 28.34). A lignin test confirms their presence.

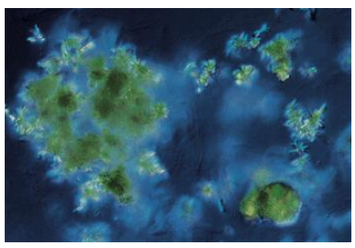


FIGURE 28.34. Sulfadiazine crystals. Sulfa drugs manufactured today are water soluble, and finding sulfa crystals in urine is unusual. (Interference-contrast microscopy, ×250.)

Indinavir

Among the growing numbers of crystallized drugs or drug metabolites detected in urine are those occurring in some patients treated with antiviral agents such as indinavir (63,64). These crystals are readily identified by bright-field, polarized, phase, and interference microscopy (Fig. 28.35A,Fig. 28.35B). Crystalluria tend to be accompanied by significant numbers of neutrophils and by fewer mononuclear cells, particularly lymphocytes and possibly histiocytes. Crystals are elongated with squared or slightly angular ends. They may be intracellular in as many as 5% to 10% of neutrophils and mononuclear cells.

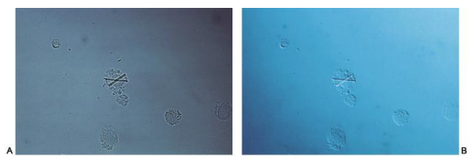


FIGURE 28.35. Indinavir crystals in leukocyte in the urine sediment. Patient with acquired immunodeficiency syndrome. **A:** Bright-field microscopy; **B:** interference-contrast microscopy (×160). These crystals are birefringent and can easily be seen using polarized light. They occur extracellularly as well as intracellularly in the urine.

Finding indinavir crystals in the urinary sediment is not an unexpected complication of therapy with this compound. The solubility of indinavir depends in part on the acidity of the urine, i.e., the lower the pH, the more soluble the medication. From the physical standpoint, it is reasonable to assume that the sharp angular ends of these crystals irritate RTE cells, resulting in an inflammatory reaction. Cells participating in this reaction, mostly neutrophils, migrate into the urine as it flows through tubular lumina, thus producing pyuria without bacteruria.

The birefringent properties of indinavir crystals are distinct from those of ordinary uric acid crystals, even though there is some morphologic similarity. Indinavir crystals are weakly birefringent and do not have the properties of urate crystals, which appear blue when a red retardant filter is used. The crystals may be found inside PMNs and appear to cause their demise when ingested. In addition, many crystals are extracellular.

Abnormal Casts

Both hyaline and granular casts are commonly found in the urine sediment of patients with renal disease. Both of these cast forms were discussed previously and were associated with the “normal” state, but there are some morphologic variations that are worthy of comment. All other casts are abnormal, for all practical purposes, and each is discussed in detail.

Casts originate in the distal tubules and collecting ducts of the kidney and are molds of the tubular lumen; therefore, a damaged nephron may have an abnormal tubular lining epithelium that

may be atrophic, degenerating, necrotic, or absent, which results in a dilated lumen (41). In such instances of nephron damage, casts form that represent this pathologic dilation by being wider than normal, i.e., broad casts (Fig. 28.36). Although the evaluation is subjective, the uroscopist is quickly able to recognize broad casts because they may be as much as five or six times wider than ordinary casts. A large number of broad casts present in a sediment is representative of considerable tubular alteration and nephron injury, and for this reason, they have also been called “renal failure casts.”

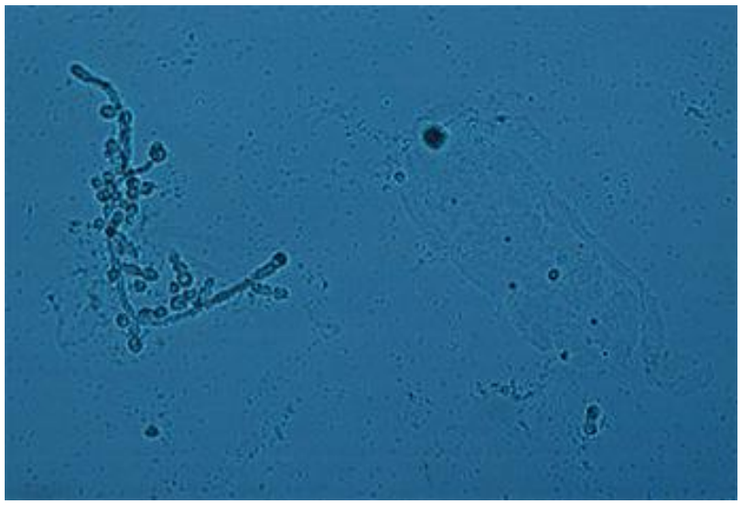


FIGURE 28.36. Broad hyaline cast. When looking at this cast, it is easy to determine that it has a greater breadth than any cast associated with a “normal” state (cf. Fig. 28.12 and Fig. 28.13). The cast depicted here shows some disruption at one end. The background contains budding yeast and hyphae, as well as a bacterium. (Bright-field microscopy, ×100.)

Damage to either the glomerular or tubular portions of the nephron may give rise to products in the sediment not normally found there, some of which may then become incorporated into casts (35). These products may be present in the glomerular filtrate (e.g., fat, hemoglobin, myoglobin, RBCs, WBCs) or may be directly related to tubular damage or destruction (e.g., epithelial cells or cell products). Conversely, casts are produced only when the environment in the renal tubule is appropriate for their formation. Although this is a little understood phenomenon, it is known that alterations in urine flow, pH, and ionic strength of the glomerular filtrate and sufficient concentration of THP are all necessary for the production of these casts (25).

Hyaline Casts

In disease states, hyaline casts may increase significantly in number in the sediment, but their morphologic appearances are not different than what is associated with the normal, except where broad forms are noted (Fig. 28.36). Broad hyaline cast forms are usually an indication of some serious intrinsic renal abnormality and require further clinical and laboratory investigation. Generally, in disease states, increased numbers of hyaline casts are found in conjunction with other sediment or chemical abnormalities, such as the presence of additional cast types and/or proteinuria. It is uncommon to find only large numbers of hyaline casts in the urine in a disease state.

Granular Casts

Large numbers of granular casts are rarely seen in the sediment in health, but this is not an infrequent finding in various diseases. These casts may be small or large or broad. The granules may be diffusely scattered over the entire cast surface (Fig. 28.37) or may be concentrated in one or more sections of the cast. Granule size has long been a topic of discussion. In the normal person, the granules are most often small and regular, and the casts are considered to be finely granular (44). Finely granular casts are also found in diseased urine. Conversely, in diseases in which there has been considerable cellular damage and death (e.g., acute tubular necrosis), the granules may be large and vary greatly in shape and size, and the resulting casts are called coarsely granular (Fig. 28.38). The difference between fine and coarse granules is somewhat subjective, and there is little reason to make this distinction in reports. Cellular destruction leads to the formation of granules that vary in size; when these granules are incorporated into a cast, the resulting granular cast is readily recognized. In

some granular casts, cell wall remnants are observed. Because of this fact, the concept arose that some preexisting cellular casts degenerate during their transit through the nephron and evolve into granular casts. In turn, some granular casts continue this evolution and become waxy casts.

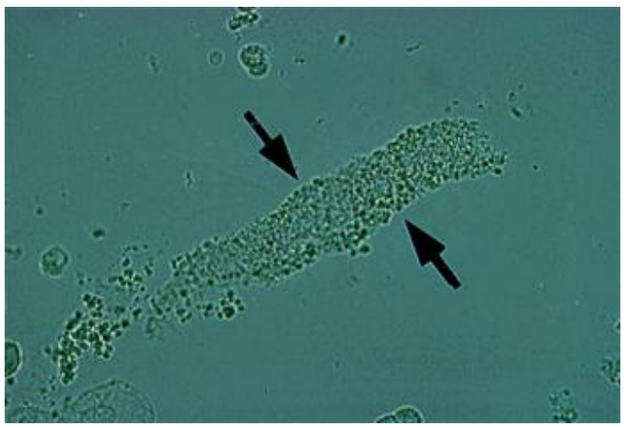


FIGURE 28.37. Granular cast. This cast (*arrows*) is composed of a diffuse scattering of fine granules over its entire surface (bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

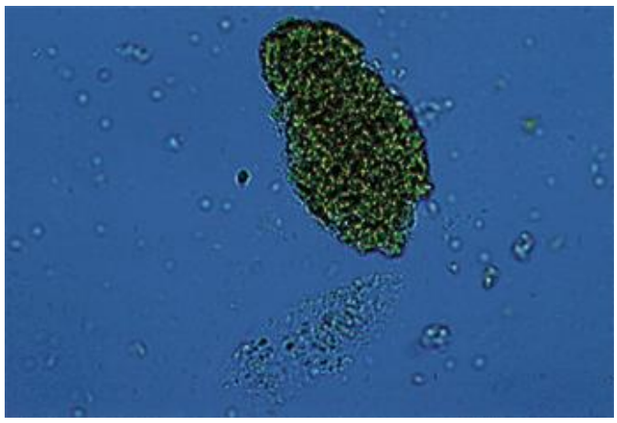


FIGURE 28.38. Broad coarsely granular cast. This cast from a jaundiced patient has adsorbed the bile pigment in the urine. The granules composing it are irregularly shaped and far larger than those seen in the cast in Fig. 28.29. A second, less obvious granular cast is also present in this photomicrograph, just below the broad granular cast initially described. (Bright-field microscopy, ×100.)

Waxy Casts

Morphologically, waxy casts are the most easily recognized of all urinary casts. This is owing to their high refractility, characteristic shape, and peculiar “tallow-wax” surface appearance. Classically, waxy casts appear as cylinders with ends that appear to have been sharply cut or broken off (not rounded) and whose sides are notched (Fig. 28.39) (11,35). Their surface is smooth, for the most part, but may contain a few granules or an occasional cell or cell membrane.

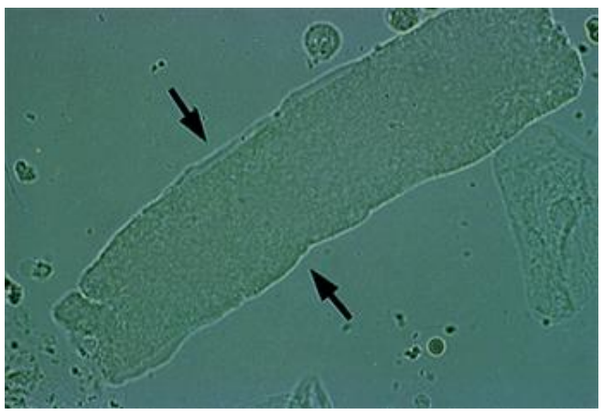


FIGURE 28.39. Waxy cast (*arrows*). This cast presents the typical appearance of a waxy cast in the urine. It has sharp, angular, squared-off ends and an irregularly notched margin. Its surface is relatively smooth and looks like it has a “tallow-wax” consistency. (Bright-field microscopy, ×160.) (From the College of American Pathologists, with permission.)

The origin of waxy casts is not clear, and scanning electron microscopy provides little in the way of solution (44). On electron microscopic study, the surface of a waxy cast is composed of what probably is a typical proteinaceous material (most likely derived from cells) that coats the cast. Waxy casts often appear as broad casts; finding large numbers of waxy casts in the sediment portends a poor prognosis and is often associated with severe nephron destruction, oliguria, and renal failure.

Pigment Casts

Mention should be made of a variety of cast that contains one of the pigments or abnormal products excreted in the urine. These pigments are most commonly bilirubin, hemoglobin, or myoglobin. Melanin pigment casts also occur but are rarely present in melanuria, a complication of malignant melanoma that has metastasized. Pigment casts are usually hyaline or waxy but appear as though stained when in fact no sediment stain has been applied (Fig. 28.40). Some pigments can be differentially stained (i.e., hemoglobin). In multiple myeloma, a neoplastic disease of plasma cells, Bence Jones protein may be excreted by the kidneys

in large quantities (see Protein section under Chemical Urinalysis section, earlier in this chapter), and casts composed of this low molecular weight protein may form and be found in the urine.

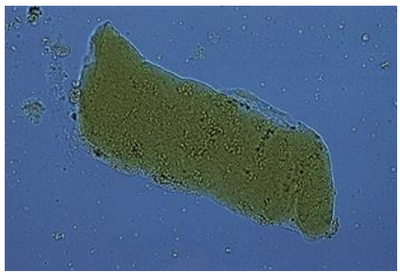


FIGURE 28.40. Waxy cast, bile-stained (pigment cast). This broad cast has a relatively smooth-appearing surface with blunt ends and notched sides. Bile has been absorbed by its matrix to give it a greenish surface appearance. (Bright-field microscopy, ×160.)

Fatty Casts

Lipid may be present in a cast in two forms: either as free globules or within cells. Cholesterol and cholesterol esters demonstrate anisotropism and can be seen with the use of a polarizing microscope (Fig. 28.41). It is observed as having a Maltese cross pattern (35,59). Conversely, neutral fats and triglycerides do not have this anisotropic property but can be stained with Sudan III or oil red O dye.

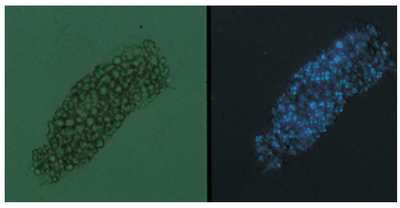


FIGURE 28.41. Fatty cast. This cast contains multiple different-sized lipid globules that, when subjected to polarized light, are anisotropic and display a Maltese cross pattern (*right*). (Bright-field microscopy and polarization, ×160.) (From the College of American Pathologists, with permission.)

Whether the fat is contained within cells (OFBs) or is present as free lipid spherules on the cast surface, accurate recognition is essential. Fatty casts are associated with renal diseases that affect both tubules and glomeruli. These casts are most commonly associated with the nephrotic syndrome but are also present in diseases in which there is severe tubular damage, such as heavy metal poisoning. Fatty casts may form as a result of glomerular injury, which causes a leakage of serum lipids into the glomerular filtrate. The newly formed fatty globules are either absorbed by the tubular epithelium with the formation of OFBs or become a component of the cast by adhering to its surface.

Fatty casts are easily confused with RBC casts because of the globular nature of the free lipid. Careful observation assists in distinguishing them because the lipid globules are of varying size, spherical, and have a rather pale golden brown color. Simple polarization reveals a Maltese cross pattern of the globules. RBC casts, however, are not anisotropic, contain reddish brown RBCs of uniform size, and are in the typical nonspherical, biconcave disk shape.

Crystal Casts

These casts are infrequently present in the urine sediment and may be a helpful indicator of renal obstruction. In obstruction when casts form, precipitation of various crystals occurs in the nephron, and these crystals become incorporated into the cast matrix (Fig. 28.42). The basic crystal cast has a hyaline matrix. Crystal casts might contain uric acid, calcium oxalate, and, rarely, sulfadiazine. Because all these crystals are anisotropic, simple polarization of the cast should demonstrate their true nature if there is any doubt about their diagnosis (Fig. 28.43). It is important to distinguish crystal casts from crystalline pseudocasts (compacted amorphous urate or phosphate crystals assuming a cylindrical castlike shape) because crystal casts may be associated with pathologic conditions, whereas pseudocasts have no clinical significance.

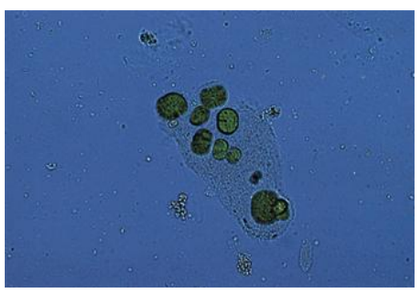


FIGURE 28.42. Crystal casts in urine. Note that the crystals are irregularly scattered over the hyaline cast matrix, which shows fine granularity. This urine sediment sample was from a liver transplant patient who was severely jaundiced. (Bright-field microscopy, ×100.)

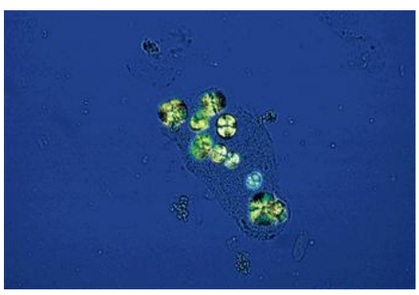


FIGURE 28.43. Crystal cast, polarized. The same cast shown in Fig. 28.42 is depicted under polarized light. The crystals in the cast show anisotropic properties that enable them to be distinguished. These casts may be associated with renal obstruction. (Polarization, ×1,003.)

RBC Casts

RBC casts are among the most diagnostically important of all elements in the urine sediment; their presence usually indicates glomerular injury (i.e., acute glomerulonephritis, IgA nephropathy, lupus erythematosus nephropathy) (65,66). RBC casts can also be found, but far less frequently, in disease states that do not primarily involve glomeruli (i.e., pyelonephritis and renal infarction). RBCs enter the urinary flow via two major mechanisms: they either pass through the GBM, a transglomerular route, or they cross the TBM, a transtubular route (41). Both mechanisms are of importance in the formation of RBC casts. When serious glomerular injury occurs, as in acute poststreptococcal glomerulonephritis, the GBM is damaged and increased numbers of RBCs transgress it; some of these RBCs become incorporated

into a preexisting cast matrix and form a RBC cast. A similar mechanism occurs in instances of transtubular bleeding. Here, the TBM is damaged, as in renal infarction or pyelonephritis, and RBCs leak across it and into the urine flow, to become incorporated into the cast (Fig. 28.44).

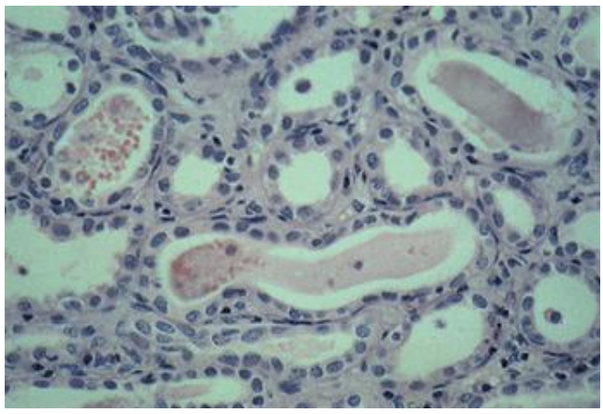


FIGURE 28.44. Red blood cell casts forming in the distal nephron. This histologic section shows three casts in tubular lumina. Two of the casts contain red blood cells, whereas the third is purely hyaline. (Hematoxylin and eosin, ×100.)

A RBC cast is basically a hyaline cast to which RBCs adhere. Scanning electron micrography of this cast type demonstrates that the RBCs are attached to the matrix by means of thin fibrils, presumably in part THP, which serve as a kind of glue to attach the cells to the cast surface. The RBCs apparently attach after the cast has been formed in the distal tubule or collecting duct.

Recognition of RBC casts requires careful microscopic observation because they may be confused with other cast forms such as fatty casts. The RBCs in the cast may be sparse, with only a few present in any one area, or they may be closely packed so that the underlying cast matrix can hardly be seen (Fig. 28.45). The RBCs themselves may be intact, thereby facilitating a diagnosis, but they frequently show degeneration and disruption. If the disruption is severe enough, one may be left with only RBC membranes and brownish red hemoglobin pigment in the cast; these are named blood or hemoglobin casts. To diagnose a RBC cast, the cast surface should contain sufficient numbers of RBCs, or their disrupted membranes, so that no doubt exists regarding the specific diagnosis. Often, RBCs within the cast will show dysmorphism, and appear in twisted and irregular shapes (51). In most patients with RBC casts, proteinuria and hematuria are seen concurrently in the chemical evaluation and the urine sediment surrounding the casts contains free-floating RBCs.

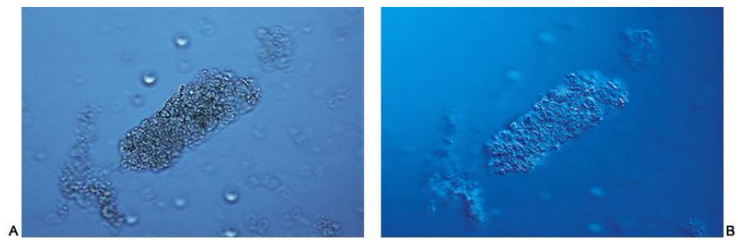


FIGURE 28.45. Erythrocyte casts in urine sediment. **A:** Bright-field microscopy; **B:** interference-contrast microscopy. The red blood cell cast is composed almost entirely of red blood cells. These are mostly intact, but a few have indistinct cell membranes. A few others show dysmorphic alterations (×160).

WBC Casts

WBC casts in the urine usually indicate that an inflammatory process is occurring in the kidney, most frequently involving the interstitium in pyelonephritis and interstitial nephritis. WBCs migrate to the affected region of the kidney and then across the TBM, where they enter the urine flow and become incorporated into the cast matrix. These WBCs are secured to the hyaline matrix by means of fibrillar protein bands, similar to those found in RBC casts (44). Glomerular damage that induces an immune response and the subsequent chemotactic attraction of leukocytes to the site of injury also cause WBC casts to form (i.e., lupus nephritis).

WBC casts may be difficult to recognize accurately because of the possibility of confusion with RTE cell casts. The white cells occupying the cast surface are most often neutrophils (PMNs) (Fig. 28.46). These cells measure approximately 20 μm across and can be identified by their multilobed nuclei and granular cytoplasm. However, neutrophils in these casts are frequently disrupted, fragmented, or undergoing degeneration, and the individual nuclear lobes may be difficult to distinguish. Phase-contrast or interference-contrast microscopy or differential staining may assist in making the distinction between leukocytes and RTE cells.

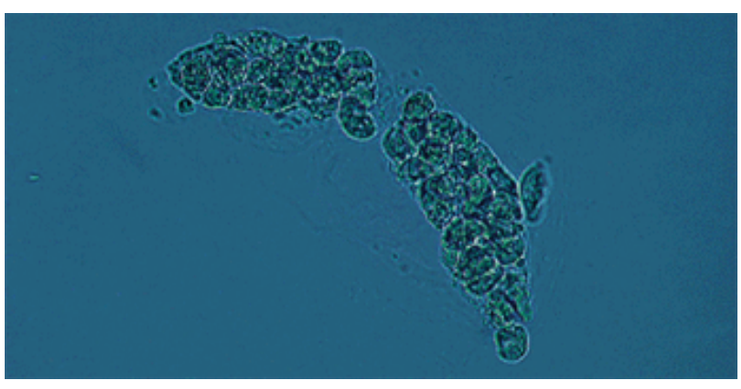


FIGURE 28.46. White blood cell cast. The cast depicted in the photomicrograph is a hyaline cast to which polymorphonuclear neutrophils have adhered. The leukocytes on the cast surface show a granular cytoplasm and multilobed nuclei. (Bright-field microscopy, ×160.)

Lymphocytes and eosinophils may occasionally be found in WBC casts. Lymphocytes, being much smaller cells than neutrophils (12 μm) and having a large central nucleus, are easily recognized. Eosinophils, conversely, may present diagnostic difficulties unless Giemsa or Hansel stain is applied. The authors have never observed a pure eosinophil cast in interstitial nephritis but have noted occasional eosinophils in WBC casts in this disease.

RTE Cell Casts

RTE cell casts are cellular casts composed of RTE that has been sloughed off. Tubular lining cells are easily affected by various poisons, especially the heavy metals such as lead and mercury, ethylene glycol (antifreeze), and viruses (cytomegalovirus and hepatitis) and are also sensitive to alterations in blood flow in instances of prolonged shock and in renal transplantation. In these examples, the RTE cells are caused to exfoliate. If hyaline casts are forming in the distal portions of the nephron, the cells become attached to the cast matrix and form a RTE cell cast (44).

RTE cell casts are sometimes difficult to distinguish from WBC casts because the two cell types are of approximately the same size (20 μm). One means of accurate identification is to look for a single large central nucleus in the cell (Fig. 28.47). In contrast, neutrophils have a smaller nucleus with multiple lobes. Both cell types have intracytoplasmic granules, although those in RTE cells are usually less prominent than those in neutrophils. When the RTE cells begin to degenerate, their cell and nuclear membranes become permeable and the cytoplasmic granules may leak into the surrounding cast matrix, giving a granular appearance. The tendency for microscopists to incorrectly diagnose RTE cell casts as WBC casts is pervasive because WBC casts are far more common than RTE cell casts. This mistake should and can be avoided by careful observation of the cells that make up the cast.

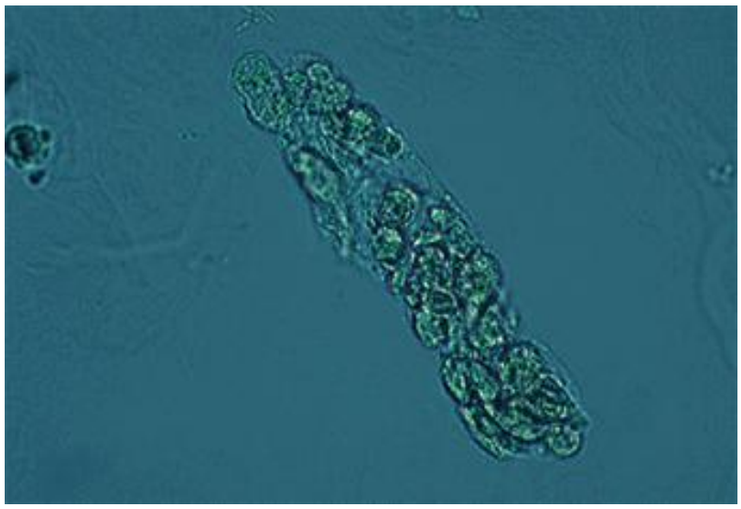


FIGURE 28.47. Renal tubular epithelial (RTE) cell cast. The basic matrix of this cast is hyaline and RTE cells coat its surface. The cells are large (approximately 20 μm) and have a large central nucleus. Their cytoplasm contains far fewer granules than do PMNs and the single nucleus distinguishes their character. They may present difficulty in differentiation from white blood cell casts. Mucous strands are present in the background. (Bright-field microscopy, ×160.)

OFB Casts

In diseases that cause severe degeneration and necrosis of RTE cells, such as heavy metal poisoning or acute tubular necrosis, the intracellular lipids become visible and the resulting sloughed OFBs are incorporated into a cast. These are truly RTE cell casts but have been designated OFB casts because of the peculiar appearance of the damaged epithelial cells and their visible and anisotropic intracytoplasmic lipid droplets. The lipid may be cholesterol or cholesterol esters, in which case it demonstrates anisotropism in a typical Maltese cross pattern, or it may consist of neutral fats and triglycerides, which can be differentially stained with Sudan III or oil red O dye. The fat globules within the RTE cells are of varying size and are usually spherical but may be somewhat irregular. If the RTE cell has a disrupted membrane, the fat droplets may leak out into the surrounding cast matrix.

Bacterial/Fungal Casts

The recognition that microorganisms could be included within the cast matrix is an observation that has often been overlooked in routine urinalysis (67). Bacterial forms, being as small as they are, are frequently not diagnosed within a cast because they appear as granules, and a granular cast is diagnosed when ordinary bright-field microscopy is used in analysis. In studies of women with acute pyelonephritis, WBC casts were often present in the urine and it became apparent that some of the granules within these WBC casts were in fact bacteria. Scanning electron microscopy studies verified this conclusion (Fig. 28.48). It is therefore

important to analyze carefully all abnormal sediments that contain granular or WBC casts and to look for bacterial and/or fungal forms in these casts. This analysis is significantly aided by using a phase-contrast or interference-contrast microscope or by staining with Sternheimer-Malbin Sedi-Stain; in either instance, bacilli will be readily identifiable.

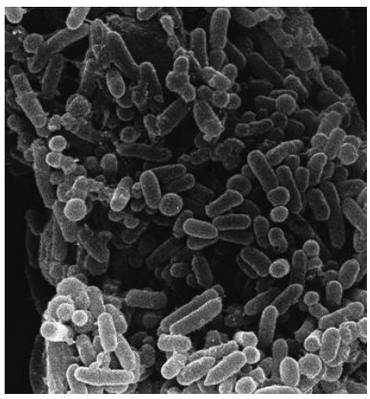


FIGURE 28.48. Bacterial cast. Numerous bacillary forms completely occupy the cast surface. (Scanning electron microscopy, ×5,000.)

It is important to recognize bacterial and fungal casts because their presence is pathognomonic for acute kidney infections (67). Bacterial forms are generally scattered throughout the entire cast matrix and are bound to its surface by bands of fibrillar protein, probably THP. WBCs are often admixed with the bacteria or fungi. Pure bacterial casts are uncommon. Fungal casts are being diagnosed more frequently today because of the recent increase in transplantation surgery and the extensive use of chemotherapeutic and immunosuppressive agents in these patients and in the many patients with malignant neoplasms. Patients with a decreased immune response are more susceptible to various infectious diseases, especially bacterial and fungal (e.g., acquired immunodeficiency syndrome). Fungal forms are somewhat easier to recognize in casts than are bacteria because of their larger size. Yeasts are often budding and, on occasion, even hyphae may be found in the cast.

Parasites

Parasitic organisms are rarely present in the urinary sediment (35). In the United States, the eggs of *Enterobius vermicularis* (pinworm) are occasionally found, especially in children. They result from fecal contamination (Fig. 28.49). The female worm lays her eggs in the perirectal region, and during urine collection, they may be carried into the urine collection bottle, especially if the collection is not performed under supervised and controlled conditions and the genital area has been inadequately cleaned. Other parasites, such as pubic lice and fleas, may be present in the sediment as contaminants when proper collection procedures have not been followed.

FIGURE 28.49. Pinworm egg from urine of 11-year-old girl. One side of the long axis is always more curved than the other.

The single parasite that inhabits the urinary tract (bladder) as part of its life cycle is *Schistosoma hematobium*. The egg of this parasite is difficult to see with ordinary microscopic techniques, as it is clear and colorless. It has a characteristic terminal spine or spike and an elliptical shape. Schistosoma eggs measure approximately 150 µm in length and 60 µm in diameter. A particular species of snail is necessary for the maturation of this organism, which, fortunately, is not present in the United States. In other parts of the world, especially Africa and Egypt, however, schistosomiasis or genitourinary bilharziasis is a significant public health problem, and the finding of these parasitic eggs in urine is not unusual.

Trichomonas vaginalis is the most frequently encountered parasite in urine sediment in the United States and is usually a contaminant from the vagina. This organism typically inhabits the male and female distal urethra, the vagina, and, rarely, the prostate gland. These flagellated protozoans are actively motile in freshly voided urine and have a characteristic appearance. The body measures 10 to 30 µm in length and typically has an undulating membrane running approximately half its length; there are four anterior flagella (Fig. 28.50). The inexperienced observer may easily misdiagnose these organisms as neutrophils because they usually occur in the presence of PMNs and are approximately the same size. Careful scrutiny must be given to the sediment to detect their motility, elliptical shape, and flagella.

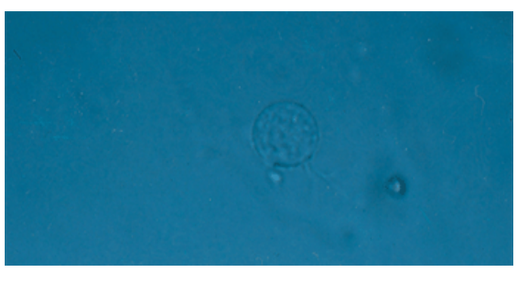


FIGURE 28.50. *Trichomonas vaginalis* in urine. Note the flagella and the spherical character of this parasite. These organisms are easily misdiagnosed as white blood cells. (Bright-field microscopy, ×400.)

Artifacts

As previously discussed, the ability to distinguish between actual urinary constituents and artifactual or contaminant material that may appear in the sediment is of primary importance in sediment

analysis. The experienced observer will have little difficulty in making this distinction and will use the microscopic tools of polarization, phase contrast, and staining to this end. Fibers, precipitated stains, oil droplets, and scratches will frequently be observed and should not be confused with, or misdiagnosed as, abnormal elements (Fig. 28.51 and Fig. 28.52).

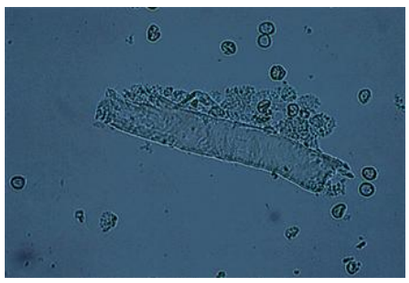


FIGURE 28.51. Fiber artifact in urine. This fiber from a disposable diaper may be easily confused with a waxy cast as it has some of the same morphologic features. A clear distinction may be made using polarized microscopy (cf. Fig. 28.52). Red and occasional white blood cells are present in the background. (Bright-field microscopy, ×160.)

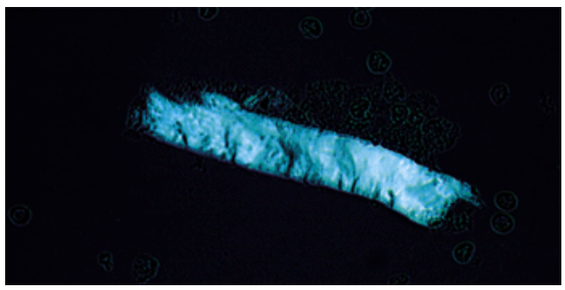


FIGURE 28.52. Fiber artifact in urine, polarized. Note the birefringence of the fiber. Waxy casts do not display anisotropism, and therefore any misdiagnosis of this fiber as a waxy cast should be prevented. (Polarization, ×160.)

SUMMARY

Part of "28 - Urine"

Microscopic evaluation of urine sediment is often a valuable diagnostic part of “routine” urinalysis. It provides a means of evaluating the urine or confirming an abnormal physicochemical finding, which is clinically useful. For this part of the urine examination to be most effective, however, a regularly employed standardized approach must be followed. Whether a laboratory purchases one of the many commercially available systems or develops one of its own, the results of performing the microscopic examination by doing the same thing to each urine sample every time a urinalysis is ordered will undoubtedly pay dividends in the long run by producing more accurate and reliable results. Table 28.6 presents an approach to the performance of urine microscopic examination after the urine has been prepared for sediment evaluation.

TABLE 28.6. METHODOLOGY IN PERFORMING A URINE MICROSCOPIC EXAMINATION

View concentrated sediment with low-power objective (100 ×)
Observe sediment for cells, casts, and crystals
Count casts as number per low-power field
Search for any other type of abnormality (i.e., yeast, parasites)
Polarize sediment to look for lipid (i.e., cholesterol)
If any part of physicochemical examination is abnormal, or
If urine is cloudy, or
If increased numbers of cells or casts are present, then
View sediment under high power (400 ×)
Identify *all types* of cells present
Count blood cells as numbers per high-power field
Count renal tubular epithelial cells as number per high-power field
Identify all types of casts present; specify type
If casts are increased in number then:
Normal vs. abnormal
If abnormal, acellular vs. cellular
Identify crystals present; determine whether
Normal or
Abnormal (only in acid or neutral urine)
Identify any microorganisms, bacteria, fungi, or parasites
Produce report and sign

REFERENCES

1. Haber MH. Pisse prophecy: a brief history of urinalysis. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:415–430.

2. Brian T. The Pisse-prophet or, certain pisse-pot lectures. London, 1637.

3. Bird G. *Urinary deposits*. London: John Churchill, 1844.

4. Bright R. *Reports of medical cases selected with a view of illustrating the symptoms and cure of diseases by a reference to morbid anatomy*. London: Longman, 1827:1.

5. Sternheimer R. A supravital cytodiagnostic stain for urinary sediments. *JAMA* 1975;231:826–832.

6. Haber MH. Quality assurance in urinalysis. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:431–448.

7. Fisher A. Quality assurance program for urinalysis. In: Haber MH, ed. *A primer of microscopic urinalysis*. Garden Grove, CA: Hycor Biomedical, 1991:12–18.

8. Free HM, Baker J, Haber MH, et al., eds. *Routine urinalysis. Proposed guidelines*, NCCLS Document GP 16-P. Villanova, PA: National Committee for Clinical Laboratory Standards, 1991.

9. Haber MH. Interference contrast microscopy for identification of urinary sediments. *Am J Clin Pathol* 1972;57:316–319.

10. Free HM. *Modern urine chemistry*. Elkhart, IN: Miles Laboratories, 1987.

11. Schumann GB, Schweitzer SC. Examination of urine. In: Henry JB, ed. *Clinical diagnosis and management by laboratory methods*, 18th ed. Philadelphia: Saunders, 1991:387–444.

12. Howanitz PJ, Saladino AJ, Dale JC. Timeliness of urinalysis: a College of American Pathologists Q-probes study of 346 small hospitals. *Arch Pathol Lab Med* 1997;121:669–672.

13. Strasinger SK. *Urinalysis and body fluids. A self-instructional text*, 2nd ed. Philadelphia: Davis, 1991.

14. Gillenwater JY. Detection of urinary leukocytes by Chemstrip-L. *J Urol* 1981;125:383–384.

15. Freeman JA, Haber MH. Urinalysis. In: Freeman JA, Beeler MF, eds. *Laboratory medicine/urinalysis and medical microscopy*, 2nd ed. Philadelphia: Lea & Febiger, 1983:1–260.

16. Kim MS. Proteinuria. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:527–540.

17. Hindmarsh JT. Microalbuminuria. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:611–616.

18. Haber MH, Lindner LE, Ciofalo LN. Urinary casts after stress. *Lab Med* 1979;10:351–355.

19. Schoolwerth AC. Hematuria and proteinuria: their causes and consequences. *Hosp Pract* 1987;22:45–62.

20. Cirillo M, Senigalliesi L, Laurenzi M, et al. Microalbuminuria in nondiabetic adults. *Arch Intern Med* 1998;158:1933–1939.

21. Mogenson CE, Keane WF, Bennett G, et al. Prevention of diabetic renal disease with special reference to microalbuminuria. *Lancet* 1995;346:1080–1084.

22. Parsons M, Newman DJ, Pugia M, et al. Performance of a reagent strip device for quantitation of the urine albumin: creatinine ratio in a point of care setting. *Clin Nephrol* 1999;51:200–227.

23. Mitchell TH, Nolan B, Henry M, et al. Microalbuminuria in patients with non-insulin-dependent diabetes mellitus relates to nocturnal systolic blood pressure. *Am J Med* 1997;102:531–535.

24. Corwin HL, Silverstein MD. Microscopic hematuria. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:601–610.

25. Haber MH. *A primer of microscopic urinalysis*, 2nd ed. Garden Grove, CA: Hycor Biomedical, 1991.

26. James GP, Bee DE, Fuller JB. Accuracy and precision of urinary pH determinations using two commercially available dipsticks. *Am J Clin Pathol* 1978;70:368–374.

27. Anonymous. *Urinalysis today*. Indianapolis: Boehringer Mannheim Diagnostics, 1987.

28. Binder L, Smith S, Kupka T, et al. Failure of prediction of liver function test abnormalities with the urine urobilinogen and urine bilirubin assays. *Arch Pathol Lab Med* 1989;113:73–76.

29. Monte-Verde D, Nosanchuk JS. The sensitivity and specificity of nitrite testing for bacteriuria. *Lab Med* 1981;12:755–757.

30. With T. *Bile pigments*. New York: Academic Press, 1968.

31. Kusumo RK, Grover PJ, Kunin CM. Rapid detection of pyuria by leukocyte esterase activity. *JAMA* 1981;245:1653–1655.

32. Sadof MD, Woods ER, Emans SJ, et al. Dipstick leukocyte esterase activity in first-catch urine specimens. *JAMA* 1987;258:1932–1934.

33. Morrison MC, Lum G. Dipstick testing of urine—can it replace urine microscopy? *Am J Clin Pathol* 1986;85:590–594.

34. Hensey OJ, Cooke RWI. Estimation of urine specific gravity and osmolarity using a simple reagent strip. *BMJ* 1983;286:53.

35. Haber MH. *Urinary sediment: a textbook atlas*. Chicago: American Society of Clinical Pathologists, 1981.

36. Koepke JA. Hematology and urinalysis testing in the physician's office. *Pathologist* 1986;40:17–20.

37. Schumann GB, Greenberg NF. Usefulness of macroscopic urinalysis as a screening procedure. A preliminary report. *Am J Clin Pathol* 1979;71:452–456.

38. Craver RD, Abermanis JG. Dipstick only urinalysis screen for the pediatric emergency room. *Pediatr Nephrol* 1997;11:331–333.

39. Addis T. *Glomerular nephritis*. New York: Macmillan, 1948.

40. Ward PCJ. *The urinary sediment videodisc text/atlas*. St. Paul: Image PSL, 1991.

41. Haber MH. *Urine casts, their microscopy and clinical significance*. Chicago: American Society of Clinical Pathologists, 1975.

42. Tamm I, Horsfall FL Jr. A mucoprotein derived from human urine which reacts with influenza, mumps, and Newcastle disease viruses. *J Exp Med* 1952;95:71–97.

43. McKenzie JK, McQueen EG. Immunofluorescent localization of Tamm-Horsfall mucoprotein in human kidney. *J Clin Pathol* 1969;22:335–339.

44. Haber MH, Lindner LE. The surface ultrastructure of urinary casts. *Am J Clin Pathol* 1977;68:547–552.

45. Brody L, Webster MD, Kark RM. Identification of elements of urinary sediment with phase-contrast microscopy. *JAMA* 1979;206:1977–1981.

46. Behrman RA. Urinary findings before and after a marathon race. *N Engl J Med* 1939;225:801–802.

47. Lindner LE, Vacca D, Haber MH. Identification and composition of types of granular urinary casts. *Am J Clin Pathol* 1983; 80:353–358.

48. Jenkins RD, Fenn JP, Matsen JM. Review of urine microscopy for bacteriuria. *JAMA* 1986;255:3397–3403.

49. Roberts AM. Some effects of exercise on the urinary sediment. *J Clin Invest* 1935;14:31–33.

50. Dressendorfer RH, Wade CE, Amsterdam EA. Development of pseudoanemia in marathon runners during a 20-day road race. *JAMA* 1981;246:1215–1218.

51. Fairley FK, Birch DF. Hematuria: a simple method for identifying glomerular bleeding. *Kidney Int* 1982;21:105–108.

52. Corwin HL, Korbet SM, Schwartz MM. Clinical correlates of eosinophiluria. *Arch Intern Med* 1985;145:1097–1099.

53. Corwin HL, Bray RA, Haber MH. The detection and interpretation of urinary eosinophils. *Arch Pathol Lab Med* 1989;113:1256–1258.

54. Wilson DM, Salazer TL, Farkouh ME. Eosinophiluria in atheroembolic renal disease. *Am J Med* 1991;91:186–189.

55. Dunnil MS. A review of the pathology and pathogenesis of acute renal failure due to acute tubular necrosis. *J Clin Pathol* 1974;27:2–13.

56. Schumann GB, Palmieri LJ, Jones DB. Differentiation of renal tubular epithelium in renal transplantation cytology. *Am J Clin Pathol* 1977;67:580-584.

57. Mandal AK. Analysis of urinary sediment by transmission electron microscopy. An innovative approach to diagnosis and prognosis in renal disease. In: Haber MH, Corwin HL, eds. *Clinics in laboratory medicine*. Philadelphia: Saunders, 1988:463–481.

58. Bolande RP. Inclusion bearing cells in the urine in certain viral infections. *Pediatrics* 1959;24:7–12

59. Hudson JB, Dennis AJ, Gerhardt RE. Urinary lipid and the Maltese cross. *N Engl J Med* 1978;299:586.

60. Papanicolaou GN. Cytology of urine sediment in neoplasms of the urinary tract. *J Urol* 1947;57:375–9.

61. Holmquist N. Detection of cancer with urinary sediment. *J Urol* 1980;123:188–9.

62. Bradley M. Urine crystals: identification and significance. *Lab Med* 1982;13:348–353.

63. Trainor LD, Steinberg JP, Austin GW, Solomon HM. Indinavir crystalluria. Identification of patients at increased risk of developing nephrotoxicity. *Arch Pathol Lab Med* 1998;122:256–259.

64. Tashima KT, Horowitz JD, Rosen S. Indinavir nephropathy. *N Engl J Med* 1997;336:138–140.

65. Heptinstall RH. *Pathology of the kidney*, 4th ed. Boston: Little, Brown, 1991.

66. Schreiner GE. The identification and clinical significance of casts. *Arch Intern Med* 1957;99:356–369.

67. Lindner LE, Jones RN, Haber MH. A specific urinary cast in pyelonephritis. *Am J Clin Pathol* 1980;73:809–811.

Section 5
Cytogenetics

Cytogenetics - Introduction

Daniel L. Van Dyke

Section Chief

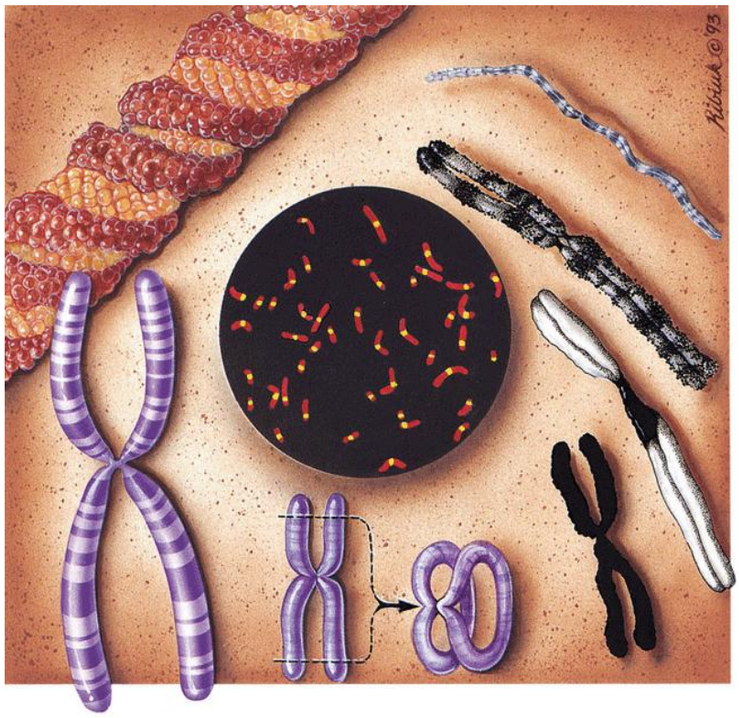


Figure.

29 Basic Cytogenetics

30 Clinical Cytogenetics

31 Prenatal Cytogenetic Diagnosis

32 Cytogenetic Studies in Neoplastic Hematologic Disorders

33 The Chromosome Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics

34 Solid Tumor Cytogenetics

Clinical cytogenetics was born in 1959 with the description of trisomy 21 in Down's syndrome, monosomy X in Turner's syndrome, and XXY in Klinefelter's syndrome. The field came of age in 1970 with the introduction of chromosome-banding techniques. Cytogenetics is maturing further today because of the rapid development of fluorescence-labeled DNA probes that can mark whole chromosomes, centromere regions, or specific genes of interest. Thus, chromosome analysis has become a major independent discipline in clinical pathology, and an ever-increasing number of medical centers and reference laboratories have a cytogenetics laboratory. Most clinical cytogenetics laboratories have experienced dramatic growth throughout the 1980s, and that growth continued into the 1990s. The importance of cytogenetics will continue to expand with the wider application of new technologies such as maternal serum screening in pregnancy, cytogenetic analysis of solid cancers, *in situ* hybridization of DNA probes, and chromosome painting.

Throughout this section, the authors offer critical technical information related to the preparation of specimens for karyotype analysis, such as the fine points of cell culture for analysis of hematologic malignancy, and the problem of confined chorionic mosaicism in the interpretation of prenatal diagnostic results. Newer concepts and insights into understanding the molecular genetic effects of cytogenetic abnormalities are detailed, such as genomic imprinting in Prader-Willi and Angelman syndromes, amplification of short “nonsense repeats” in fragile X syndrome, the molecular genetic defects associated with the inherited chromosome instability syndromes, and alterations involving oncogenes and tumor-suppressor genes in cancer and leukemia.

29

Basic Cytogenetics

Stuart Schwartz

Catherine G. Palmer

HISTORY OF HUMAN CYTOGENETICS

Part of "29 - Basic Cytogenetics"

The initial discoveries leading to the discipline of human cytogenetics occurred in the early 1950s, but their advent was based on a period of technical development starting at the turn of the century. The understanding of human chromosome behavior is based on chromosome methodology and an understanding of aberrations gained from plant and insect cytogenetics, supplemented by methods developed in studying animal tumors and tissue cultures (1).

The early studies of human chromosomes from the 1890s to the 1920s used testicular tissue to observe the meiotic divisions occurring in gametogenesis. The tissue, obtained from recently executed criminals, was fixed immediately postmortem and usually was sectioned. The chromosome counts varied widely, but the work of Winiwarter in 1912 established chromosome counts of 47 in spermatogonia and 48 in oogonia and an XX/XO sex-determining mechanism. T. S. Painter confirmed Winiwarter's work in 1923 and established the human chromosome number to be 48 but disagreed with the sex-determining mechanism, considering it to be XX/XY.

Further progress required tissue culture technology and methodology for handling mitotic cells, not the least of which was the discovery by T. C. Hsu in 1952 of the usefulness of hypotonic pretreatment of tissue culture cells before fixation to disperse the metaphase chromosomes.

Among the methods derived from plant cytogenetics was the use of agents that interfered with the organization of tubulin into the mitotic apparatus. Blakeslee and Avery used the alkaloid colchicine for this purpose and prepared plant chromosomes by a technique involving squashing the cells under a coverglass with gentle pressure. Another plant cytogeneticist, Albert Levan, in John Biesle's tissue culture laboratory, was the first to combine the colchicine and hypotonic pretreatments to obtain chromosome preparations of metaphases from tissue cultures.

In 1956, Levan and Tjio (2) used the same methodology on tissue cultures of human embryonic cells and found the chromosome number of these mitotic cells to be 46 rather than the previously accepted 48. The work was shortly thereafter confirmed by Ford and Hamerton, who observed 23 sets of paired chromosomes or bivalents in squash preparations of human male meiosis.

After the correct normal chromosome number was established, the identification of chromosome abnormalities soon followed. Jerome Lejeune in 1959 described trisomy of a small chromosome in Down-syndrome, and in rapid succession, sex chromosome abnormalities occurring in already described syndromes were identified: 45,X in Turner-syndrome by Ford and co-workers in 1959, 47,XXY in Klinefelter-syndrome, and 47,XXX by Jacobs and associates in the same year. The two other viable trisomic syndromes described, trisomy 13 by Patau and trisomy 18 by Edwards, were described the following year.

Some of these investigators used tissue cultures of skin biopsies and others used bone marrow preparations, but the squash methodology was in use until Rothfels and Simonivitch discovered that rapid air drying on slides of drops of fixed cells resulted in spreading of the metaphases. The culmination of this series of technical advances occurred when Peter Nowell in 1960 discovered that cultures of leukocytes that had been separated from red blood cells using phytohemagglutinin, an extract of the broad bean, were stimulated to divide in culture. This technique, still used today, provided a readily available source of dividing cells and led to the identification of many chromosome abnormalities.

The rapid progress of cytogenetics during these few years required that a common system of nomenclature for the chromosomes be developed, and a series of conferences involving the major investigators in cytogenetics was convened. The first conference, in Denver in 1960, resulted in a system with which to group the chromosomes based on relative length, arm ratio, and centromeric index and proposed a standard system of nomenclature of the groups using a numerical designation, i.e., chromosomes 1–3,4,5,6–12 (Fig. 29.1 and Fig. 29.2). The London Conference in 1963 further described the system and added letters to the groups, A to G plus the sex chromosomes. The meeting in Chicago in 1966 adopted a shorthand system for describing human chromosome abnormalities.

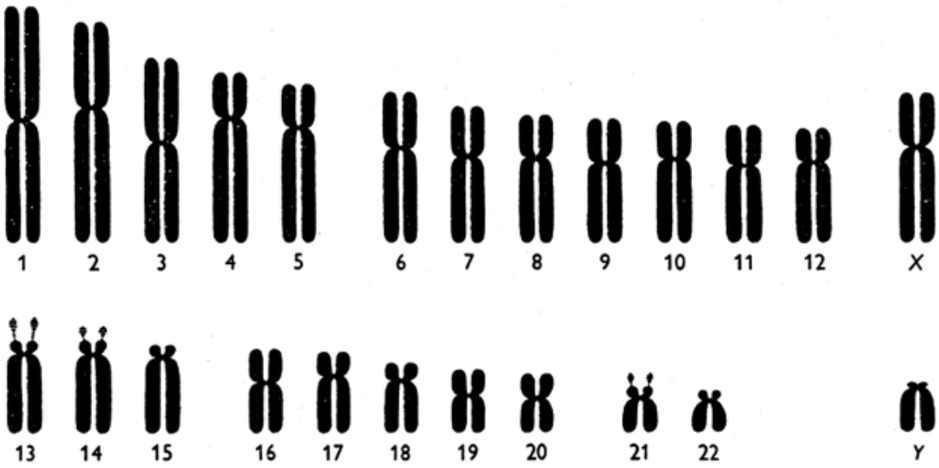


FIGURE 29.1. Idiogram drawn from standard measurements of human chromosomes from the 1960 Denver conference. (From Editorial comments, *Ann Hum Genet* 1960;24:319, with permission.)

Chromosome No.	Relative length				Centromere index			
	A	B	C	D	A	B	C	D
1	9.08	9.08 ± 0.611	9.11 ± 0.53	8.44 ± 0.433	48.0	49.4 ± 3.04	48.6 ± 2.6	48.36 ± 1.166
2	8.45	8.17 ± 0.250	8.61 ± 0.41	8.02 ± 0.397	38.1	39.4 ± 2.05	38.9 ± 2.6	39.23 ± 1.824
3	7.06	6.96 ± 0.352	6.97 ± 0.36	6.83 ± 0.315	45.9	47.6 ± 2.10	47.3 ± 2.1	46.95 ± 1.557
4	6.55	6.62 ± 0.403	6.49 ± 0.32	6.30 ± 0.284	27.6	29.2 ± 2.97	27.8 ± 3.3	29.07 ± 1.867
5	6.13	6.34 ± 0.366	6.21 ± 0.50	6.08 ± 0.305	27.4	29.2 ± 3.03	26.8 ± 2.6	29.25 ± 1.739
6	5.84	6.19 ± 0.516	6.07 ± 0.44	5.90 ± 0.264	37.7	39.1 ± 2.63	37.9 ± 2.5	39.05 ± 1.665
7	5.28	5.60 ± 0.435	5.43 ± 0.47	5.36 ± 0.271	37.3	35.3 ± 2.90	37.0 ± 4.2	39.05 ± 1.771
X	5.80	5.45 ± 0.377	5.16 ± 0.24	5.12 ± 0.261	36.9	41.4 ± 6.16	37.5 ± 2.7	40.12 ± 2.117
8	4.96	5.13 ± 0.307	4.94 ± 0.28	4.93 ± 0.261	35.9	32.7 ± 2.80	32.8 ± 2.8	34.08 ± 1.975
9	4.83	4.81 ± 0.194	4.78 ± 0.39	4.80 ± 0.244	33.3	37.0 ± 3.04	32.7 ± 4.1	35.43 ± 2.559
10	4.68	4.66 ± 0.512	4.80 ± 0.58	4.59 ± 0.221	31.2	35.4 ± 3.81	32.3 ± 2.9	33.95 ± 2.243
11	4.63	4.70 ± 0.289	4.82 ± 0.30	4.61 ± 0.227	35.6	40.7 ± 3.07	40.5 ± 3.3	40.14 ± 2.328
12	4.46	4.66 ± 0.410	4.50 ± 0.26	4.66 ± 0.212	30.9	30.5 ± 3.64	27.4 ± 4.0	30.16 ± 2.339
13	3.64	3.22 ± 0.310	3.87 ± 0.26	3.74 ± 0.236	14.8	-	16.6 ± 3.6	17.08 ± 3.227
14	3.55	3.09 ± 0.212	3.74 ± 0.23	3.56 ± 0.229	15.5	-	18.4 ± 3.9	18.74 ± 3.596
15	3.36	2.83 ± 0.262	3.30 ± 0.25	3.46 ± 0.214	14.9	-	17.6 ± 4.6	20.30 ± 3.702
16	3.23	3.46 ± 0.353	3.14 ± 0.55	3.36 ± 0.183	40.6	42.2 ± 3.57	42.5 ± 5.6	41.33 ± 2.74
17	3.15	3.06 ± 0.377	2.97 ± 0.30	3.25 ± 0.189	31.4	36.6 ± 5.86	31.9 ± 3.3	33.86 ± 2.771
18	2.76	2.98 ± 0.316	2.78 ± 0.18	2.93 ± 0.164	26.1	31.5 ± 4.15	26.6 ± 4.2	30.93 ± 3.044
19	2.52	2.55 ± 0.269	2.46 ± 0.31	2.67 ± 0.174	42.9	48.1 ± 2.48	44.9 ± 4.0	46.54 ± 2.299
20	2.33	2.61 ± 0.144	2.25 ± 0.24	2.56 ± 0.165	44.6	46.5 ± 3.59	45.6 ± 2.5	45.45 ± 2.526
21	1.83	1.34 ± 0.189	1.70 ± 0.32	1.90 ± 0.170	25.7	--	28.6 ± 5.0	30.89 ± 5.002
22	1.68	1.53 ± 0.178	1.80 ± 0.26	2.04 ± 0.182	25.0	--	28.2 ± 6.5	30.48 ± 4.932
Y	1.96	1.82 ± 0.353	2.21 ± 0.30	2.15 ± 0.137	16.3	-	23.1 ± 5.1	27.17 ± 3.182

FIGURE 29.2. Relative length and arm ratios of human chromosomes. A, B, C, and D represent different data sets. Sets B, C, and D were identified by Q banding and stained with orcein or Giemsa 9. Set A is the Denver-London data not preidentified by Q banding. The relative length of each chromosome is the length of the chromosome compared with the length of the haploid (1N) set of chromosomes. There are two ways to designate the centromere position: the arm ratio, which is the length of the longer arm relative to the shorter one, and the centromere index, the ratio of the length of the shorter arm to the total length of the chromosome. (From Harden DG, Klinger HP, An international system for human cytogenetic nomenclature. In: Harden DG, Klinger HP, eds. Birth Defects: Original Article Series, vol. 21. New York: March of Dimes Birth Defects Foundation, 1985. Basel, Switzerland: Karger, 1985:114, with permission.)

Although many chromosome abnormalities were described by 1966, the abnormalities were relatively easily visualized—gains or losses of whole chromosomes or easily observed deletions or rearrangements. Some chromosomes could not be readily differentiated from each other; a few were separable by autoradiographic methods based on differences in the completion of replication of similarly appearing chromosomes. Caspersson in 1970 used a fluorescing dye, quinacrine mustard, to stain human chromosomes and was able to demonstrate a different pattern of fluorescence for each of the 22 pairs of autosomes and

the X and Y chromosomes. Each chromosome could now be defined by a specific pattern of light and dark bands, and the banding pattern was shown to be reproduced by other staining techniques using Giemsa stain or other fluorescent dyes. This technique permitted identification of further rearrangements, deletions, duplications, and inversions not previously recognized. The Paris Conference of 1971 addressed the identification of individual chromosomes, regions, and bands and detailed a method of describing structural changes using the banding nomenclature. All the nomenclature was consolidated in the International System for Human Cytogenetic Nomenclature (ISCN) in 1978 (see Chapter 30, Table 26.1) (3).

The major change in the subsequent years was the use of prophase and prometaphase chromosomes to obtain a greater number of recognizable bands, which permitted recognition of even more subtle abnormalities. This was achieved by Yunis, who synchronized the dividing cells to obtain the early stages of division in sufficient quantity to make them useful.

The description of the prophase chromosomes—the size, width, and number of bands at different levels of band resolution—was incorporated into the ISCN in 1981 in *An International System for Human Cytogenetic Nomenclature—High Resolution Banding 1981*; some further revisions were made in 1985 (Fig. 29.3).

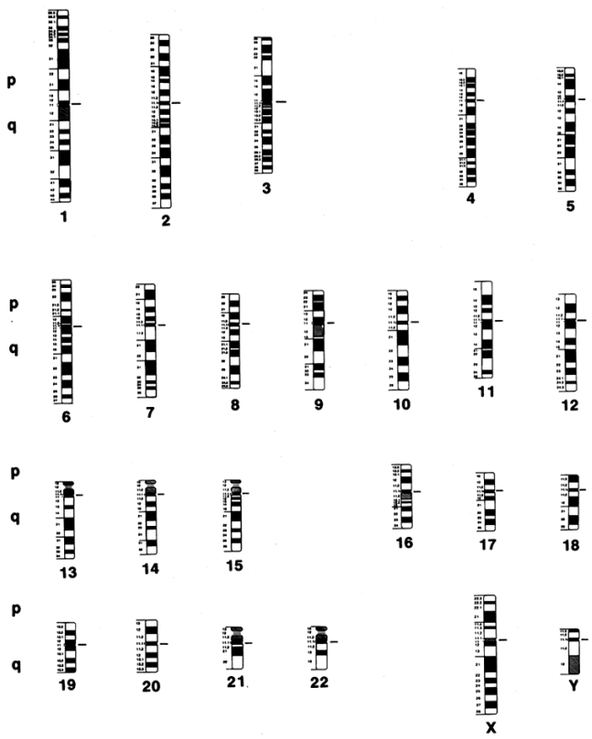


FIGURE 29.3. Diagram showing G-banding pattern of normal human chromosomes at 400-band level. (Adapted from Harden DG, Klinger HP. An international system for human cytogenetic nomenclature. Harden DG, Klinger, HP (eds), In: Harden DG, Klinger HP, eds. *Birth Defects: Original Article Series*, Vol. 21. New York: March of Dimes Birth Defects Foundation, 1985.)

In 1991, recognizing the complexity of describing acquired abnormalities with the current nomenclature, the Study Committee on Human Cytogenetic Nomenclature established new guidelines specific for describing abnormalities encountered in cancer cytogenetics (3a).

The most recent advance in cytogenetics has been the advent of molecular cytogenetics, in which techniques of molecular biology are applied to cytogenetic preparations. These procedures include *in situ* hybridization for localizing DNA probes to specific chromosomes and bands. Radioactive probes labeled with tritiated thymidine were initially used, followed by procedures utilizing immunologic reactions coupled with staining of the product with a fluorescent dye (immunofluorescence). These techniques were useful for both gene localization and the identification of fragments or rearrangements in dividing and nondividing cells.

The greatest advances, however, have involved the utilization of fluorescence *in situ* hybridization (FISH) using a variety of probes labeled with biotin and visualized by fluorochrome staining or probes directly labeled with fluorochromes in conjunction with fluorescence microscopy. This method permits staining of individual chromosomes or regions of chromosomes and allows identification of specific chromosome rearrangements in either metaphase or interphase cells.

As in previous years, an updated version of the ISCN was published in 1995 (4) that addressed the new guidelines that were developed to describe the utilization of FISH. These guidelines proposed an entirely new nomenclature to incorporate this new technology.

Cytogenetics has now become an integral part of the gene mapping attempts and has a wide application to medical questions regarding the relationships between such chromosome abnormalities as infertility, miscarriage, birth defects, mental retardation, and cancer.

CHROMOSOME STRUCTURE

Part of "29 - Basic Cytogenetics"

Studies of chromosomes of a variety of organisms have led to the conclusion that the chromosome is composed of a single continuous DNA fiber in each chromatid (i.e., one of the two daughter strands of a replicated chromosome). The basic subunit is a 10-nm fiber that has been described after digestion by micrococcal nuclease as the “nucleosome”—the familiar “beads-on-a-string” structure, the core particle of which contains 145 bp of DNA helix coiled around an octomer of histones H2a, H2b, H3, and H4 (Fig. 29.4). Adjacent nucleosomes are separated by

DNA linkers of varying length associated with H1 histone, which stabilizes the structure (Fig. 29.5). This fibril is in turn compacted into a 30-nm thick fibril that undergoes further condensation to form the metaphase chromosome. Observations of higher chromosome structure derived from whole-mount or histone-depletion electron microscopy give different perspectives of chromosome architecture (Fig. 29.6) and have resulted in several models of chromosome structure (5).

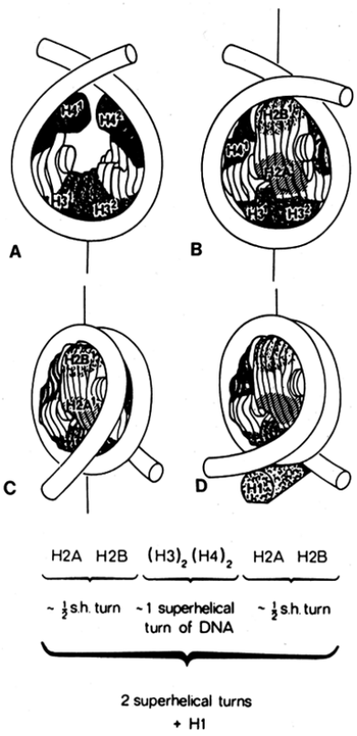


FIGURE 29.4. Assembly of nucleosome core particle from its constituents. (From Richmond TJ, Finch JT, Klug A. Studies of nucleosome structure. *Cold Spring Harb Symp* 1982; 47:493–501, with permission.)

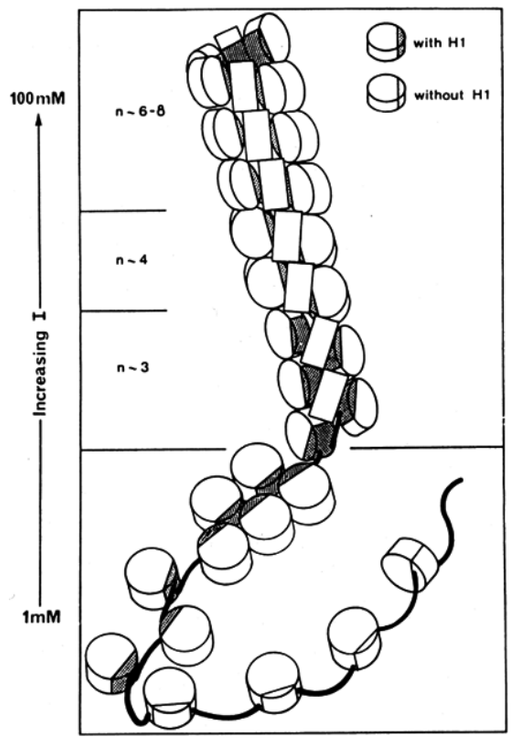


FIGURE 29.5. Idealized drawing of the solenoid structure of a chromosome showing the open zigzag of nucleosomes that form the solenoid. (From Thoma F, Koller TH, Klieg A. *J Cell Biol* 1979;83:403–427, with permission.)

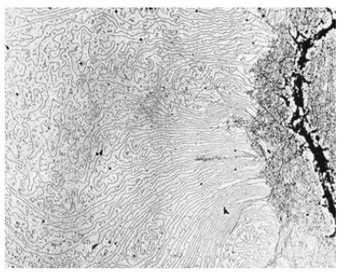


FIGURE 29.6. Electron micrograph of histone-depleted chromosomes shows the organization of the DNA in loops. The chromosome scaffold is to the left with loops extending outward from the scaffold. (From Paulson JR. Scaffolding and radial loops, the structural organization of metaphase chromosomes. In: Adolph KW, ed. *Chromosomes and chromatin, volume III*. Boca Raton, FL: CRC Press, 1988:3–36, with permission.)

When chromosomes are digested to remove most of the histones, leaving only protein and DNA and observed by electron microscopy after spreading, a loose fibrillar protein structure extending through the chromatids and surrounded by a halo of loops of DNA is revealed. These observations have suggested a chromosome structure in which radial loops of DNA are supported by a cytoskeleton or scaffold of fibrillar protein (Fig. 29.7).

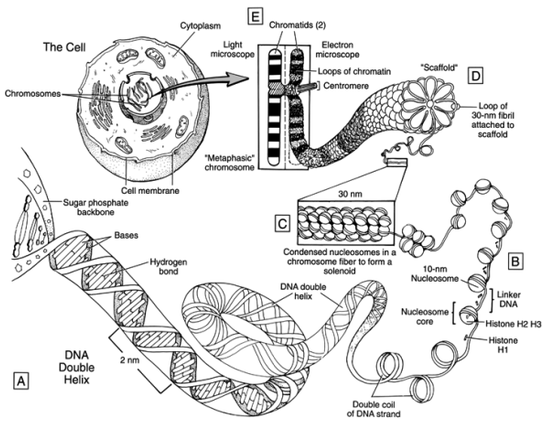


FIGURE 29.7. Model of chromosome structure shows the hierarchy of coiling and looping in the DNA helix **(a)**, nucleosomal 10-nm fiber **(B)**, 30-nm solenoidal fiber **(C)**, the looping of the 30-nm fibers and their attachment to the chromosome scaffold **(d)**, and the chromosome as seen on electron microscopy **(E)**.

Another model derived from whole-mount electron microscopy and light microscopy observations of coiled chromatids suggests that the chromosome is based on a successive series of helical coils of the 30-nm fiber to form a super solenoid, which eventually results in a coiled metaphase chromatid of approximately 1,000 nm in width. All these structures require some degree of compaction of the 10-nm DNA fiber into the condensed chromatin that is visualized at mitotic or meiotic division. The current concept of the chromosome is that the solenoid and looping conformations are both present in the chromosome. The solenoidal structure is the 30-nm fibril that is attached to a scaffoldlike matrix in loops. The whole structure undergoes a higher order of condensation by further coiling as the chromosome shortens in progress toward metaphase (6,7).

Specialized Regions of Chromosomes

The Telomere

The ends of the chromosomes, the telomeres (Fig. 29.8), have several important functions in the cells. The concept of the telomere is based on the observation that when chromosomes are broken by radiography or other means, the broken ends will fuse to each other, but not to unbroken ends. Telomeres protect the ends of the chromosomes from degradation and from becoming involved in end-to-end fusions and allowing complete replication of the end of the chromosome. Telomeres at the end of chromosomes enable cells to distinguish chromosome ends from breaks in DNA. This is essential as unrepaired chromosomes with breaks will cause the cell to become arrested at a checkpoint. Telomeres at the ends of chromosomes allow the cells to pass through the checkpoint. Telomeres also guide localization of the chromosomes within the nucleus as they have the tendency to be positioned near each other (telomere association) and with the nuclear membrane in meiosis and, on occasion, in mitotic prophase.

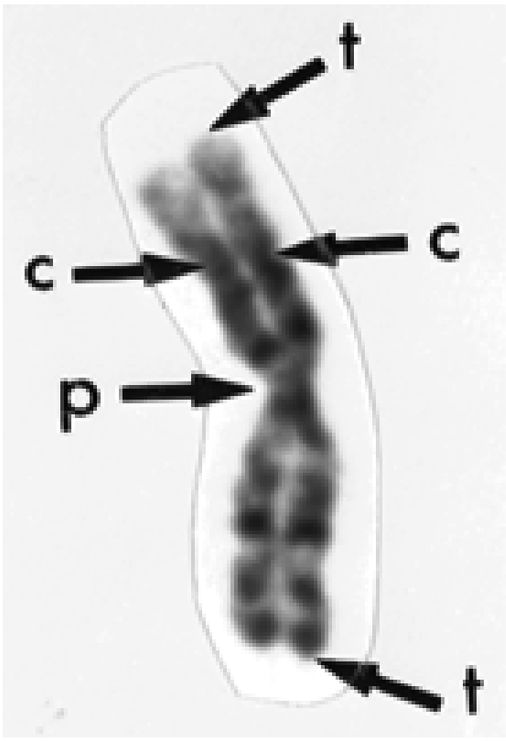


FIGURE 29.8. Each chromosome at metaphase is composed of two chromatids (*c*), each of which has a single continuous strand of DNA extending from end to end. Each chromatid has a centromere, which is located in a constriction, the primary constriction (*p*), and at each end each chromatid is capped by a telomere (*t*).

There is a considerable amount of telomeric DNA [approximately 15 kilobase (kb)] within the cells that is regulated in both species- and cell-specific ways. Additionally, although the telomeres are composed of simple sequences (see below), small changes

can be catastrophic (e.g., division failure). Therefore, it has been hypothesized that telomeres have additional functions including regulation of gene expression, and control over cell division, cell senescence, and neoplastic proliferation. The lack of telomerase will cause the chromosomes to shorten, thereby exposing active genes and eventually leading to cell death. Telomerase activity will immortalize cells, but its mere presence does not confer immortality, indicating that its relationship with cancer is more involved.

The structure of telomeres must account for the characteristics described above as well as for the means of replication of the terminus of a linear DNA molecule. Studies of telomeres have used organisms with many short chromosomes, such as *Tetrahymena* and yeast, to permit purification and molecular analysis of the chromosome ends. The *Tetrahymena* telomere has thus been found to contain 1,500 to 6,000 tandem repeats of the nucleotide sequence TTAGGG. This structure has been found in organisms as widely separated as birds and reptiles, as well as

yeast and other microorganisms. The repeating sequence TTAGGG also occurs at the termini of human chromosomes as well (Fig. 29.9) (8). These tandemly repeated DNA sequences are conserved among vertebrates; however, the number of repeats will vary from 500 to 300. In humans, the telomeres in germ line cells are significantly longer than in the somatic tissue cells. The homology of base sequences in the telomeres can account for telomere association of homologous and nonhomologous chromosomes. In cases in which telomere association persists to metaphase, it is possible that an endonuclease involved in replication has failed to function.

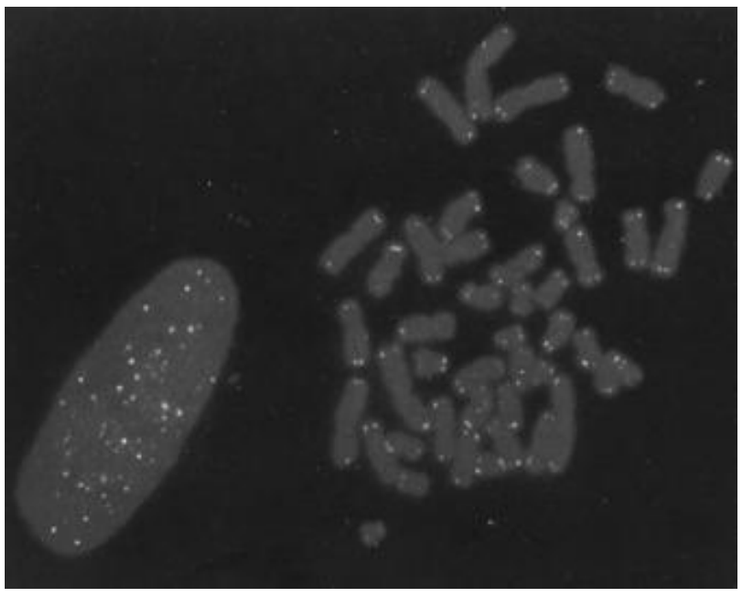


FIGURE 29.9. Telomeres of chromosomes shown after *in situ* hybridization with biotin-labeled repetitive telomeric sequences to metaphase chromosomes. (From Moyis RK, Buckingham JM, Cran LS et al. A highly conserved repetitive DNA sequence (TTAGGG), present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* 1988;85:6622–6626, with permission.)

There is telomere-associated DNA (TAS) located immediately internal to the terminal telomeric DNA. Both the telomeres and TAS are dynamic structures that are subjected to both deletion and amplification, apparently involving homologous recombination. Internal to the TAS are unique subtelomeric sequences, which are unique among all the chromosome arms. These are the sequences utilized for the FISH subtelomeric probes (as described below).

There are special problems in replicating the ends of the DNA strand, namely, the requirement of a DNA polymerase to have both a template to copy and a primer to extend because polynucleotide synthesis can proceed in only one direction (from 5′ to 3′). The 3′ ends of the strands cannot be replicated because there is no more DNA left on which to synthesize the RNA primer. How then does the telomere replicate?

An unusual telomerase that functions as a special reverse transcriptase has been found in*Tetrahymena*. The RNA of the telomerase provides its own template for the synthesis of the telomeric repeats that are added on during replication of the ends of the chromosome. A similar telomerase has been identified in human and mouse cells, suggesting that the *Tetrahymena* system may be present in all cells. It has been shown that human telomerase contains a RNA subunit of approximately 450 nucleotides with no sequence homology to any of the lower eukaryotic telomerase RNA components. Proteins act with the RNA template to maintain telomerase. In humans, two important proteins have been identified: TRF1 and TRF2 (9,10). Both of these proteins bind to the TTAGGG repeat and are localized to the telomere of human metaphase chromosomes and are present throughout the cell cycle. TRF1 functions as a suppressor of telomere elongation as overexpression of TRF1 leads to telomere shortening. However, a mutation will lead to a dominant-negative effect and telomere elongation. TRF2 prevents end-to-end fusion and contributes to maintenance of correct structure of telomeres. Overexpression of this protein has no effect; however, expression of a deleted TRF2 protein leads to rapid growth, senescence, and fusion of telomeres (11,12 and 13).

The Centromere

The centromere, the region of the chromosome also referred to as the primary constriction, divides the chromosome into arms (Fig. 29.8). The centromere may be in the middle of the chromosome, resulting in a metacentric chromosome, may be asymmetrically placed on the chromosome to give a submetacentric and/or acrocentric chromosome, or may be at the end of a chromosome in a telocentric chromosome. The centromere is composed primarily of repetitive DNA and associated with the kinetochore (see below) (14).

The centromere of a chromosome has several important functions: (a) it is essential for chromosome attachment to the spindle apparatus, (b) it is essential for the proper segregation of the chromosomes in both mitosis and meiosis, and (c) it serves as a point of attachment for both sister chromatids. As indicated above, the centromere contains repetitive DNA. One important class of tandemly repeated centromeric DNA, which constitutes one component of the functional centromere, is the alpha-satellite DNA family. The amount of this DNA is extremely variable between centromeres, consisting of anywhere from 200 kb to greater than 5,000 kb. The alpha- satellite DNA is located at the primary constriction, and these repeats comprise approximately 3% to 5% of the genome. The basic unit of the alpha-satellite DNA is a 171-bp monomer, which is repeated anywhere from 100 to 1,000 times. Although the monomers between different chromosomes are similar, the sequence may vary by as much as 10% to 40%. This allows the development of specific DNA probes for FISH to detect individual centromeres. It has been hypothesized that alpha-satellite DNA is necessary for centromeric function. A 17-bp of alpha-satellite DNA is recognized by centromere protein B (CENP-B). These CENP-B boxes are regularly distributed at high frequency in the alpha-satellite DNA and analogous to other boxes in other species. Additionally, when alpha-satellite DNA has been introduced to cell lines, this DNA confers centromeric activity. In opposition to these findings, there have now been several reports of chromosomes with no detectable DNA or CENP-B but which normally segregate in cells, suggesting that alpha-satellite is necessary for a functioning centromere (15,16).

The kinetochore is a specialized proteinaceous structure located on either side of and closely associated with the centromere. This centromere/kinetochore complex consists of the kinetochore, heterochromatin, and associated proteins. It is the location

for the mechanochemical motor components of chromosome movement in mitosis. This complex has three structural domains: (a) inner pairing domain, containing proteins and DNA; (b) central domain, consisting of heterochromatin; and (c) outermost domain (kinetochore), containing the kinetochore structure, associated proteins, and chromatin. This outermost domain is the primary point of spindle attachment.

The kinetochore functions as the region of attachment of the subunits of the spindle apparatus, the microtubules, to the DNA of the chromosome during division. As indicated previously, the kinetochore of a mammalian chromosome is a multilayered disk that stains with heavy metals; when visualized by electron microscopy, it appears to have a trilaminar structure composed of an outer plate of condensed fibrillar or granular material 40 nm thick that contains DNA (Fig. 29.10). The kinetochore microtubules appear to terminate in this outer plate. Beneath this is a middle and more electron-translucent 30-nm thick layer that appears to be loosely aggregated and fibrillar in nature. The inner layer, also composed of DNA, is similar in thickness to the outer layer and approximates the underlying chromatin. In addition to the DNA and RNA contained in the kinetochore, there are specific proteins associated with the outer and inner kinetochore plates (17).

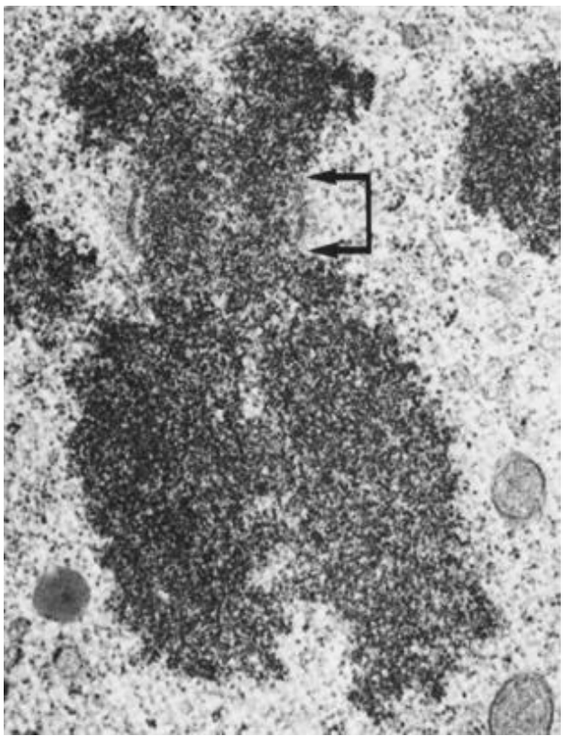


FIGURE 29.10. An electron micrograph of a thin section of a mouse chromosome shows the primary constriction (*arrows*) and the kinetochore, which appears along the lateral surface of the constriction. (From Rattner JB. In: Adolph KW, ed. *Chromosomes and chromatin, vol II*. Boca Raton, FL: CRC Press, 1988:29–49, with permission.)

Patients with autoimmune diseases, such as scleroderma and CREST syndrome, have antibodies that react with these CENPs. These CENPs consist of several different species, all located at the centromere/kinetochore complex. CENP-C, the CliPs (chromatio linking proteins) and Inceps (inner centromere proteins) are localized to the inner pairing domain, CENP-A and CENP-B to the central domain, and CENP-E, and CENP-F to the outermost domain. CENP-B was the first centromeric protein to be isolated and cloned. As described above, its main function is an interaction and binding with alpha-satellite DNA and is conserved among vertebrates. There is a variation in intensity of CENP-B, based on the amount of satellite DNA present. CENP-C has a direct role in kinetochore structure and activity. It is directed only at the active centromere, whereas CENP-B can be also found at inactive centromeres. CENP-C is necessary for both kinetochore assembly and establishment of stable chromosome attachment to the spindle. CENP-E is a unique motor protein. It is associated with microtubule and chromosome movement. Similar to CENP-C, it is localized to only the active centromere. CENP-A is a centromere-specific histone H3 homolog, which is located at the inner plate of the kinetochore. Similar to CENP-C and CENP-E, it is associated with only an active centromere. CENP-A is important in the construction of the nucleosome core particles (18,19).

Heterochromatin

As described previously, a major component of the centromeric region is constitutive heterochromatin (repetitive DNA). First described by Heitz in 1928, heterochromatin is chromatin that remains condensed telophase and remains visible during interphase. This out-of-phase or allocyclic behavior of heterochromatin is also reflected in its DNA synthesis, which occurs late in the replication of the chromosome. Two types of heterochromatin may be differentiated.

The first type, constitutive heterochromatin, contains DNA that is not transcribable and is located mainly at centromeric and telomeric sites, although in some organisms, its location is not limited to these regions. In human chromosomes, constitutive or c-heterochromatin is centromeric in all the chromosomes, with additional paracentromeric areas in chromosomes 1, 9, and 16. It makes up much of the long arm of the Y chromosome and may be readily demonstrated by C banding (described later in this chapter). The second type of heterochromatin, facultative heterochromatin, is composed of transcribable DNA that has been inactivated and that has assumed the other characteristics of heterochromatin, including allocyclic condensation and late replication (see X-chromosome Inactivation section).

Constitutive heterochromatin is composed largely of a type of DNA, satellite DNA, which differs in base composition and physical characteristics from the remainder of the DNA. When DNA is fragmented and ultracentrifuged, the lightest and heaviest fragments settle at a different rate than the remainder of the DNA, appearing as separate satellite bands. Satellite DNA is composed of multiple tandem repeats of simple-sequence DNA, 2 to 12 bp in length, resulting from amplification of these simple sequences. Divergence owing to mutation produces related but distinct satellites within an organism. In addition to the satellite DNA, other repeated sequences are found in the human genome. These include short interspersed repeated sequences, which are approximately 300 bp in length, Alu elements (which are repeated DNA sequences cleaved by the restriction endonuclease AluI), and long interspersed repeated sequences. Eight different

satellites identified by gradient centrifugation have been defined in human DNA and have been localized in the centromeric and telomeric regions of the chromosomes by *in situ* hybridization. Analysis of human genomic DNA using restriction endonucleases (enzymes that cleave DNA molecules at specific recognition sites) has revealed additional satellite DNA composed of tandem repeats that also hybridize to centromeric regions. One of these, alpha-satellite DNA (see above), is composed of a number of subfamilies that differ with respect to the sequence and location of sites of cleaving by restriction endonucleases.

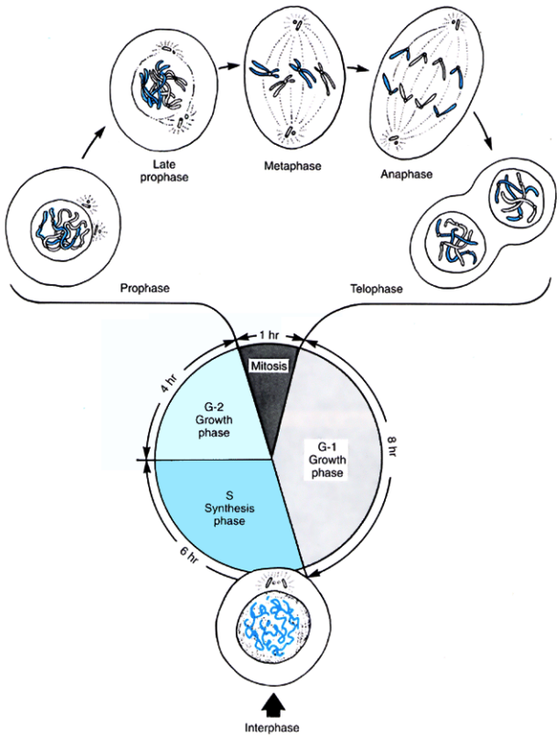


FIGURE 29.11. The mitotic cycle. The mitotic division comprises only a small segment (0.75 to 1.0 hour) of the cycle, which is characterized by an S or synthesis phase and two growth phases, all occurring during interphase.

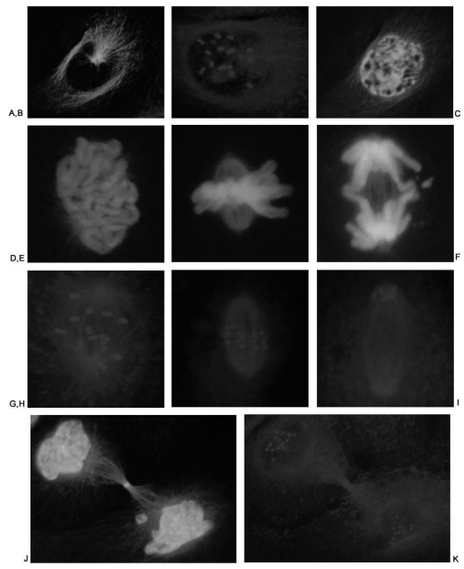


FIGURE 29.12. Components of the mitotic apparatus after staining with specific antibodies and fluorescent stains. Cells are from a mammalian cell line (PtK). **A:** The microtubules are stained green with antitubulin antibodies and fluorescein-labeled secondary antibodies. **B:** The kinetochores are stained red with anticentromere antibodies from a patient with scleroderma plus rhodamine-labeled secondary antibodies. **C:** The chromatin of an early prophase cell. The chromosomes are stained blue with DAPI stain. Centrosome activity is revealed in the two asters of the cytoplasmic microtubules (green), but the nuclear envelope blocks interaction with kinetochores. **D–F:** Prometaphase through anaphase. The chromosomes are stained blue with DAPI and the spindle green with antitubulin antibodies and fluorescein-labeled secondary antibodies. **G–I:** Prometaphase through anaphase stained to show kinetochores (with anticentromere antibodies and rhodamine-labeled secondary antibodies). **J, K:** A pair of cells in late telophase-interphase, when nuclei are essentially reestablished, but the two cells are connected by a cytoplasmic bridge. **J:** Same cells stained to show chromatin and spindle as above. **K:** Same cell stained to show kinetochores. (From McIntosh J, Koonce MP. Mitosis. *Science* 1989;246:622–628, with permission.)

Alpha-satellite sequences are located at the centromeres of human chromosomes. The alpha-satellite regions of most human chromosomes differ from each other (20).

Mitosis

DNA packaging changes during the cell cycle. From great extension during transcription, there is an increasing formation of loops as metaphase approaches, concurrent changes in the methylation of chromatin, and the presence of a variety of associated histone and nonhistone proteins as the chromosome is again compacted.

The culmination of contraction of the interphase chromosome is the mitotic division. Mitosis, which is necessary to distribute the chromosomes equally at cell division, is only a transient phase occupying approximately 1 hour of a cell cycle, which comprises some 18 to 24 hours (Fig. 29.11). At the initiation of the cell cycle, cells may be activated from a resting state, G-0, and enter the cell cycle. During the G-1 stage of the cell cycle, events occur that are necessary for the subsequent S (synthesis) phase in which the chromosomes are replicated. DNA replication is followed by G-2, a period in which further changes occur in preparation for formation of the spindle, in nuclear membrane breakdown, and in further shortening of the chromosomes composed now of two chromatids. At the beginning of metaphase, the chromosomes align themselves on the metaphase plate so that each short thick chromatid is oriented toward one pole of the mitotic spindle. As anaphase occurs, the two chromatids of each chromosome separate beginning at the kinetochores and move to opposite poles of the spindle. By telophase, the chromosomes are grouped at the poles. A nuclear membrane is reformed and the chromosomes uncoil, forming the interphase nuclei of the daughter cells (Fig. 29.12) (21).

Meiosis

Whereas mitotic divisions result in two identical daughter cells, the essence of meiosis is pairing, recombination, and a reduction in chromosome number to the haploid 1N number. DNA replication during premeiotic S is three to six times longer than the S phase of mitotic cells. The DNA is almost completely replicated, with a small portion remaining unreplicated until replication is finally completed at zygotene. Meiosis results in a reduction to the haploid chromosome number because two cell divisions occur with only one replication of DNA. The synaptonemal complex (SC), a unique structure not found in mitotic divisions, holds the paired homologous chromosomes together during meiosis I and facilitates the exchange of chromatin that results in crossing over.

Meiosis is characterized by a series of stages that are microscopically distinct, with significant changes in chromosome behavior seen as well at the electron microscope level (Fig. 29.13). The chromosomes are at their greatest length at leptotene. They are still unpaired, but because DNA replication has occurred, two chromatids may be observed. At this stage, the telomeres are attached to the nuclear membrane and synapsis of the chromosomes begins from these ends. The homologous chromosomes are brought together by the protein chromosome cores to which the loops of chromatin are attached. Synapsis is complete at zygotene, the core proteins from the homologous chromosomes forming the SC. This structure, specific to meiosis, is essential for maintaining the alignment of the homologous chromosomes, which are now closely paired (Fig. 29.14).

At the pachytene stage of meiosis, the paired homologous chromosomes (or bivalents), each composed of four chromatids (tetrads), are held in register by the SC as crossing over occurs. The SC, a ladderlike structure composed of two lateral elements on either side of the chromosomes and a central fibrillar region, is complete. Recombination nodules, units that mediate the molecular process of exchange, appear on the SC and chiasmata are visible. The latter are evidence that crossing over has occurred.

The recombination nodules correspond in number and position with chiasmata. At diplotene, the SC begins to disappear, but the chromosomes are held together by chiasmata. If no chiasmata are present, the bivalent falls apart and the homologous chromosomes will segregate independently. Thus, the chiasmata have an additional important function in holding the paired homologous chromosomes together.

At diakinesis, the chromosomes shorten, and as they shorten, the chiasmata appear to move to the ends of the chromosomes (terminalize), and there is a reduction in the total number of chiasmata. The bivalents move to the periphery of the nucleus. The nuclear membrane disappears and chromosomes align on the spindle. At metaphase as the paired homologs align, the kinetochores are widely separated facing opposite poles of the spindle. Orientation of centromeres of nonhomologous chromosomes in the metaphase spindle is random, resulting in independent assortment of the several pairs of chromosomes.

When the homologs separate, the telophase groups of chromosomes at the poles contain half the number of chromosomes of the original cell. However, each chromosome still has two chromatids, and each is referred to as a dyad.

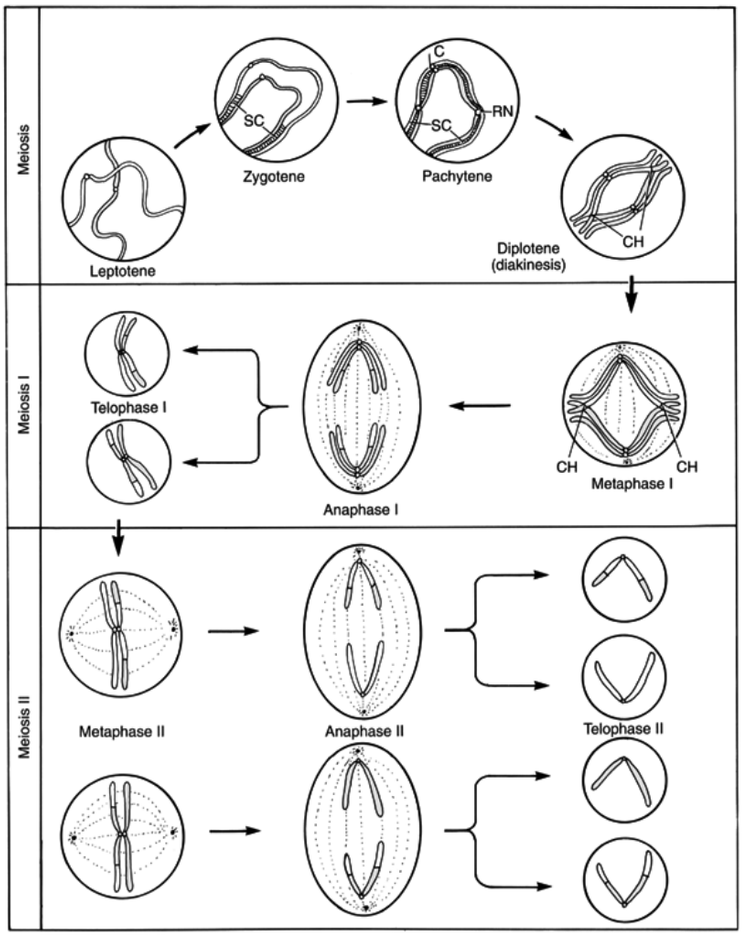


FIGURE 29.13. The meiotic division. Note the extended chromatin of leptotene, the beginning of the formation of the synaptonemal complex (SC) at zygotene, the completed synaptonemal complex showing recombination nodules (RN) at pachytene, the four strands showing crossing over at diplotene to diakinesis with the centromeres (C) widely separated and the chromosomes held together by chiasmata (CH), and the alignment of homologs at metaphase I and separation at anaphase I (only one daughter cell shown). Prophase II is short and is not shown. At metaphase II the sister chromatids separate, and at telophase II the haploid number results.

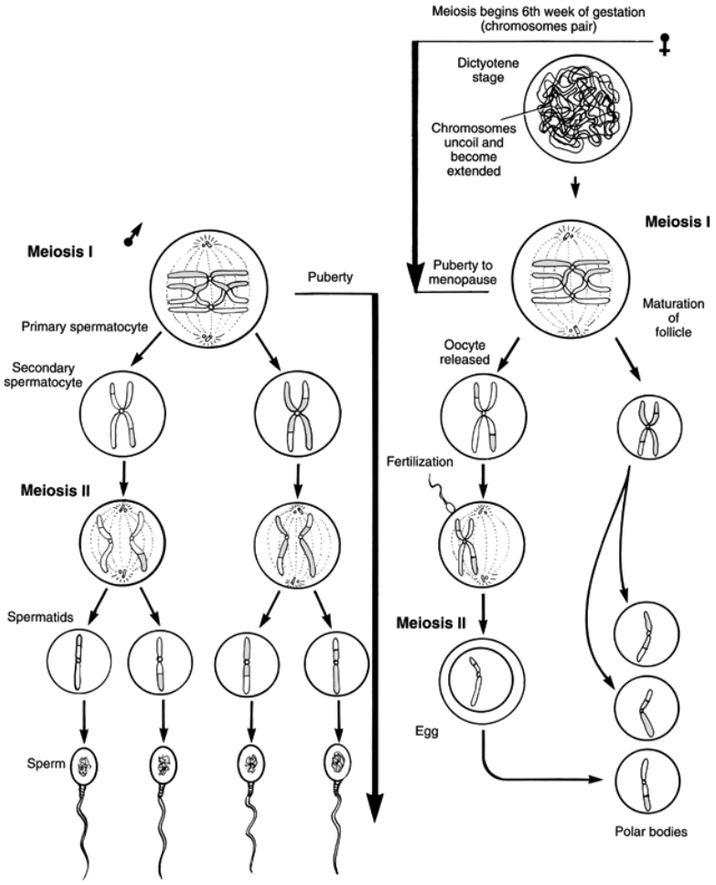


FIGURE 29.14. Comparison of meiosis in male and female gametogenesis.

The subsequent interphase is short and without a period of DNA replication, and the dyads separate at the second meiotic division. Each of the products of meiosis now has half the number of chromosomes and half the amount of DNA of the original germinal cells.

In males, meiosis occurs in germ cells, which begin to divide meiotically at puberty. Meiotic divisions are of much longer duration than mitosis, requiring 15 to 32 hours to complete the division. The spermatogonia initially turn into primary spermatocytes. These spermatocytes enter meiosis I and emerge as haploid

secondary spermatocytes. Secondary spermatocytes enter meiosis II, emerging as spermatids. Approximately 100 to 200 million spermatids are in each ejaculate. The transition from spermatogonia to mature sperm is a relatively rapid process, requiring 60 to 65 days.

In females, the first meiotic division is initiated prenatally at approximately 3 to 6 weeks of gestation. The initial process of transition from primordial germ cell to oogonia involves 20 to 30 mitotic divisions. The progression from oogonia to mature ova involves a complex series of intermediate events and is completed as each egg matures and is released from the follicle. This may occur at any time from puberty to menopause. The second

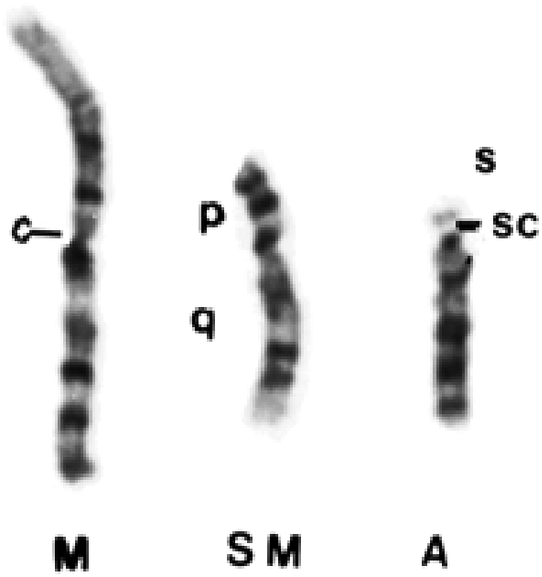


FIGURE 29.15. Morphologic characteristics used to group human chromosomes. M, metacentric; SM, submetacentric; A, acrocentric; c, centromere; p, short arm; q, long arm; s, satellites; sc, secondary constriction.

meiotic division occurs in the fallopian tube after fertilization. There is a very lengthy interval between the onset of meiosis and its eventual completion. Each of the meiotic divisions produces a polar body so that at maturation, the egg and two polar bodies result (Fig. 29.14).

There are several distinct differences between mitosis and meiosis. Mitosis occurs in all tissues, meiosis only in the testis and ovary. Diploid somatic cells are produced through mitosis, whereas meiosis produces haploid sperm and egg cells. There is normally one round of DNA replication per cell division in mitosis and one round of replication in meiosis; however in the latter, there are two cell divisions during this interval. Prophase is very short in mitosis (approximately 30 minutes in human cells), but is long and complex in meiosis I and can take years to complete. Recombination is rare and abnormal in mitosis but is normally seen at least once for each pair of homologs in meiosis. All the daughter cells in mitosis are genetically identical, but they are different in meiosis because of recombination and independent assortment of homologs.

CHROMOSOME IDENTIFICATION

Part of "29 - Basic Cytogenetics"

Human chromosome morphology is usually based on the appearance of mitotic metaphase chromosomes. Metaphase chromosomes are composed of two identical chromatids resulting from the prior replication of DNA during the S phase. Chromosomes may differ in length and in position of the centromeres. The chromosomes appear constricted at the centromere or primary constriction, and this constriction usually divides each chromosome into two arms. The arms are designated p (the short arm) and q (the long arm) (Fig. 29.15). These arms may be equal, in which case the chromosome is referred to as metacentric; if the arms are unequal, the chromosome is called submetacentric. The centromere may be near one end of the chromosome (acrocentric chromosome) or terminal (telocentric chromosome).

The morphology of the human chromosome was initially defined by the relative length of the chromosome compared with the total length of a normal X-containing haploid set and by two parameters relative to the position of the centromere: the arm ratio and the centromeric index (Fig. 29.2).

Another defining feature of the chromosome is the occurrence of secondary constrictions, a thinning or indentation of the chromosomes, which may be visible in only a portion of the cells. These constrictions occur in areas of heterochromatin and at the tip of the acrocentrics, where they separate a small segment, the satellite, from the chromosome. The secondary constrictions vary in size; in human acrocentric chromosomes, the region of the constriction between the main part of the chromosome and the satellite contains the nucleolar-organizing region.

Chromosome Staining Patterns

Chromosome bands in human chromosomes were first demonstrated by Caspersson in 1970. A band is defined as an area of the chromosome that is darker or lighter than its adjacent segments. The chromosome consists of a continuous series of dark and light bands. Several staining procedures have been developed, and bands that appear dark with one procedure may be light with another. Chromosome banding patterns are a means of identifying individual chromosomes (22,23 and 24).

Q and G Bands

Caspersson used quinacrine mustard, a fluorescent dye, to demonstrate areas of bright and dull fluorescence on the arms of the chromosomes. The bands produced are referred to as Q-bands. Q-banding permits identification of each chromosome in the human karyotype. In particular, the Y chromosome is readily differentiated from the remaining acrocentric chromosomes by the brilliant fluorescence of the distal half of Yq, which is also visible in the interphase nucleus as a brilliant Y body. With Q-banding, centromeric regions and satellites demonstrate variations in staining intensity and size that are inherited (Fig. 29.16). The staining reaction of Q-banding is based on interaction with specific bases, adenine and thymidine. Other adenine-thymidine (AT)-specific fluorochromes such as Hoechst 33258 and DAPI produce a similar banding pattern.

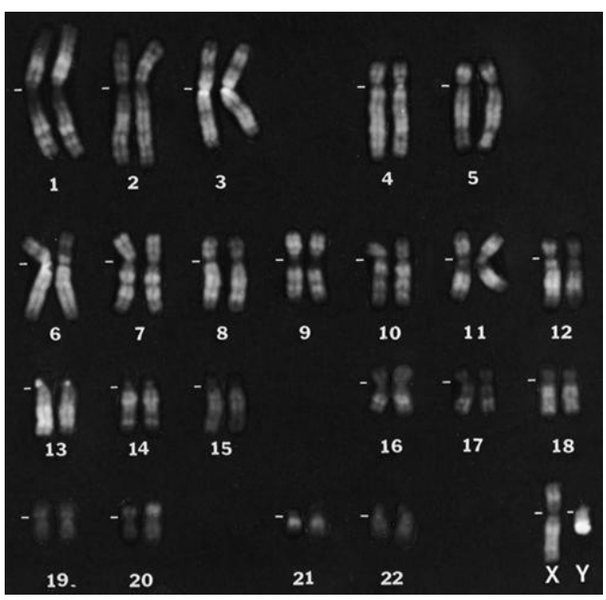


FIGURE 29.16. Q-band karyotype. (Reproduced from Lin CC, Vehida IA. Fluorescent banding of chromosomes. In: Kruse PF, Patterson MK, eds. Tissue culture methods and applications. New York: Academic Press, 1973:778–781, with permission.)

G-bands are produced by pretreatment with mild salt or proteolytic enzymes, followed by staining with one of the Romanowsky dyes (Giemsa, Wright's, or Leishman's stain) (Fig. 29.17). The patterns of G- and Q-bands are similar and correspond to the pattern of chromomeres seen in the chromosomes of meiotic pachytene. The G/Q-bands condense early in mitosis and meiosis, and replication studies with tritiated thymidine (^{3}HT) or bromodeoxyuridine (BrdU) show that these regions replicate late in the cell cycle.

Although darkly staining G-bands generally correspond to bright Q-bands, there are some differences in staining, particularly in some paracentromeric regions that stain negatively by Q-banding but stain darkly with G-banding. Similarly, the segment of the Y chromosome that stains brilliantly with Q-banding stains with G-banding uniformly with the remainder of the chromosome. Thus, although there are similarities in banding patterns, the two banding methods have different uses in cytogenetic

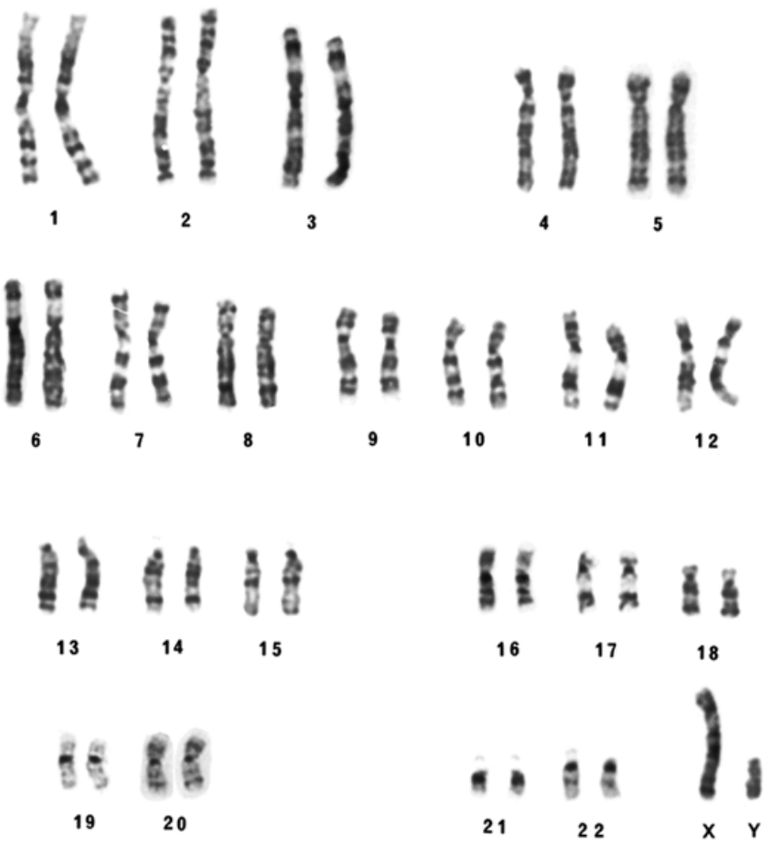


FIGURE 29.17. G-band karyotype of a normal male.

studies, G-banding being used more routinely for cytogenetic studies and Q-banding reserved for special purposes requiring visualization of polymorphisms on the Y chromosome.

Method for G-Banding with Trypsin

To obtain sharply defined bands with trypsin banding, slides must be aged. Freshly prepared slides are placed in a drying oven at 65°C overnight or longer, depending on humidity, and then trypsinized. The materials needed are 0.1 g Difco powdered trypsin (1:250) in 100 mL isotonic buffered saline (Isoton), isotonic buffered saline, phosphate buffer (pH 6.8), and Giemsa working solution (1 mL Giemsa and 49 mL phosphate buffer at pH 6.8).

Slides are treated in trypsin solution for 3 to 7 seconds, rinsed in Isoton and fetal calf serum and then in Isoton alone. The slides are then stained for approximately 2 to 4 minutes in Giemsa solution, rinsed in phosphate buffer (pH 6.8), then in distilled water, and allowed to dry.

Method for QM Staining

The materials needed are 5 mg quinacrine mustard dihydrochloride in 80 mL KH_2 PO_4 buffer (store in light-tight container) and buffer (18.16 g KH_2 PO_4 in 2 L of distilled H_2O).

Slides do not need to be aged. Stain slides 1 to 2 minutes and rinse in distilled water. Place slides in KH_2 PO_4 buffer for 1 to 2 minutes. Place two or three drops of KH_2 PO_4 buffer on slide, add coverslip, blot excess buffer, and seal edges of coverslip. Store in the dark. Observe with fluorescent microscope using a wavelength of 450 to 500 nm.

R-Bands

R-banding is produced by treatments that selectively denature AT-rich DNA; for example, heat denaturation of chromosomes at 85°C followed by staining with Giemsa or with acridine orange. Fluorochromes with an affinity to guanine-cytosine (GC) bp such as chromomycin A3 or olivomycin will also produce

a fluorescence pattern similar to that of R-bands. R-banding may also be produced by BrdU incorporation, a replication banding technique (RBG) that takes advantage of the late replication of R-bands (discussed Replication Banding section).

R-banding is the reverse of G/Q-banding. Regions that appear light with G-banding are darkly stained with R-banding (Fig. 29.18). After R-banding, the ends of the chromosomes appear darkly stained. Thus, it is useful to utilize R-banding when an abnormality involving the chromosome ends, such as a terminal deletion, is suspected. The extent of deletions may be more precisely determined from the darkly staining R-bands than with G-banding, in which the same band is lightly staining.

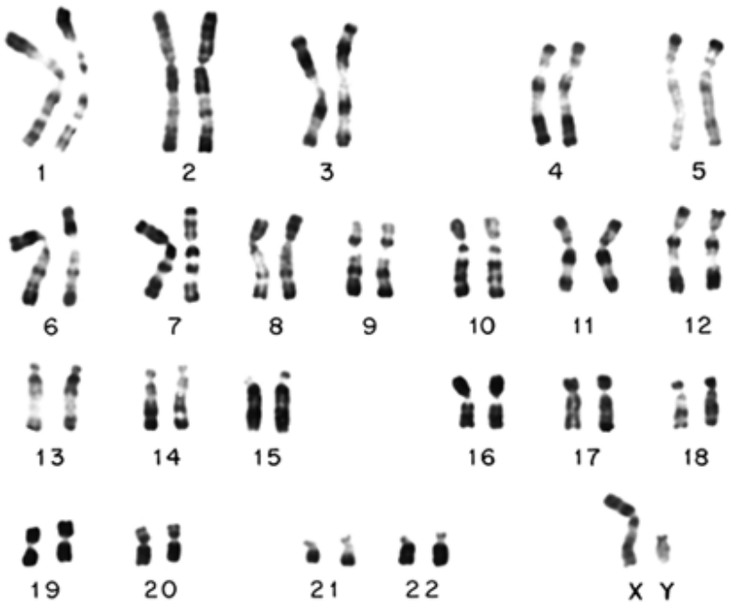


FIGURE 29.18. R-band karyotype of a normal male. (From Drouin R, Richer CL. High-resolution R-banding at the 1250-band level. II. Schematic representation and nomenclature of human RBG-banded chromosomes. *Genome* 1989;32:428, 429, with permission.)

R-bands are relatively GC rich, replicate in the first half of the S phase, and condense late in prophase. *In situ* hybridization of DNA complementary to mRNA sequences indicates that the coding DNA is localized in R-bands. Among genes and genetic diseases localized to specific bands on the chromosome, the greatest proportion (74%) is located in G-negative bands.

R-Banding Method

Aged slides are incubated in Sorensen's buffer pH 6.5 at 85°C for approximately 8 minutes, rinsed in the same buffer at room temperature, stained in Giemsa for 5 to 10 minutes, rinsed in distilled water, and allowed to dry.

C-Bands

C-banding is derived from the methodology used by Pardue and Gall for *in situ* hybridization of mouse satellite DNA to mouse centromeric heterochromatin. Arrighi and Hsu found that this method produced darkly staining centromeric heterochromatin

without the hybridization step by staining with Giemsa after denaturation with sodium hydroxide (barium hydroxide may be substituted) and subsequent incubation in a salt solution. The bands at the centromeres are readily observed (Fig. 29.19). These heritable regions vary in size, may differ in homologs, and may vary in position in the long or short arm pericentromeric region. C-banding may also be obtained after treating fixed cells on slides with the restriction endonucleases AluI, DdeI, HaeIII, MboI, RoaI, or HinfI.

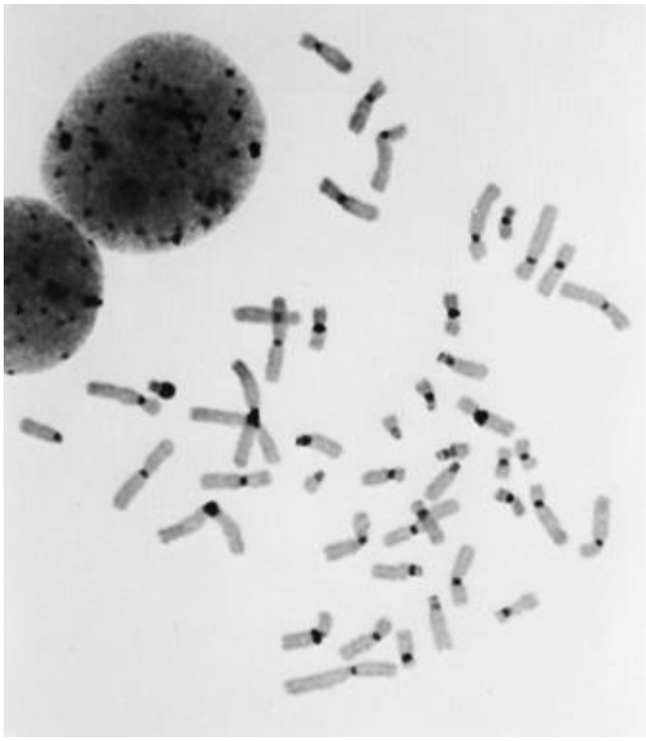


FIGURE 29.19. C-banded metaphase: centromeric and paracentromeric regions of chromosomes 1, 9, 16, and Yq stain. (Courtesy of J. Zunich.)

C-banding is used to identify the heterochromatic segments (for example, in suspected dicentric chromosomes) and to identify chromosome variants when there is a question of whether an unusual finding is a variant or an abnormality. Familial variants may also be used to trace the origin of specific abnormalities (e.g., Down syndrome).

C-Banding Method

Slides that have been aged at least 1 week are incubated for 1 hour in 0.2N HCl, rinsed in distilled water, and incubated for approximately 3 minutes in a saturated solution of $Ba(OH)_2$ diluted with distilled water [15 mL $Ba(OH)_2$ and 35 mL H_2O]. Slides are rinsed in distilled water until the $Ba(OH)_2$ film disappears and are then incubated in 2×SSC at 60°C for 1 hour. They are rinsed and stained in Giemsa for 10 minutes, rinsed again, and dried.

Chromosome Heteromorphisms

Constitutive heterochromatin in human chromosomes may vary in size or position in the chromosome. This variation in amount of highly repetitive DNA results in chromosome variants or heteromorphism. Homologous chromosomes may vary in the amount or position of C-bands and in the intensity of fluorescence or size of Q-bands. Individual variation in C heterochromatin is inherited and the heteromorphisms segregate in a mendelian fashion.

C-band heteromorphisms include variation in the size and position of the centromeric heterochromatin of all chromosomes

and of the paracentromeric heterochromatin of chromosomes 1, 9, and 16 and the distal long arm of Y. Variation in the size of short arms or satellites may also be documented by C-banding. The variation in size is continuous and may be classified by length into very small, small, intermediate, large, and very large. In some cases, large C-bands may appear to be composed of two segments. The position of C-bands may be in the long or short arm or part may be in the long arm and part in the short arm. Although Q- and C-band variants have been studied in normal and abnormal populations, no clinical significance has been found.

THE SIGNIFICANCE OF BANDING AND CHROMOSOME STRUCTURE

Part of "29 - Basic Cytogenetics"

Several lines of evidence have suggested that G-bands are more AT rich than R-bands. Evidence from base-specific fluorochromes such as Hoechst 33258 supports this observation. Autoradiographic studies using tritiated thymidine or tritiated guanidine and buoyant density data suggest that G- and R-band DNA differ in base ratio. The differences in R- and G-bands may be related to the distribution of families of interspersed repeated sequences. The short interspersed repeated sequences, of which the Alu DNA sequence family is predominant, occur mainly in R-bands. Alu is 56% guanine plus cytosine. In G-bands, the long interspersed repeated sequences, or L_1 family, occur more frequently. L_1 is composed of 58% adenine plus thymidine. Thus, the difference in distribution of these repeated sequences may account for the banding with fluorescent dyes and the reaction with base-specific antibodies. The basis for banding with proteolytic enzymes and Giemsa staining may be related to differences in the distribution of interacting proteins.

C-bands are areas of the chromosome that are composed of condensed and highly repetitive, nontranscribed satellite DNA, which is highly methylated (see Heterochromatin section). C-banding procedures result from the differential extraction of DNA in which more DNA remains in the C-band region. This may be related to a difference in protein-DNA interaction interfering with DNA extraction. Alpha-satellite DNA has been found by restriction analysis and *in situ* hybridization to be located in the C heterochromatin of all human chromosomes. Subsets of alpha-satellite sequences have been localized to specific centromeric regions (although some are present in more than one chromosome).

The alphoid sequences are hypermethylated and may be related to the inactivity of C heterochromatin. Thus, the studies of the basis for staining procedures used for the identification of human chromosomes have led to a further understanding of the structure of chromosomes. Some of the functional and biochemical characteristics of chromosomes are summarized in Table 29.1 (25,26,27,28 and 29).

TABLE 29.1. FUNCTIONAL AND BIOCHEMICAL ATTRIBUTES OF CHROMOSOME BANDS

	Bands		
Attribute	**G/Q**	**R**	**C**
Base composition of DNA	AT rich	GC rich	Mainly AT rich
Timing of DNA replication	Last half of S phase	First half of S phase	Late S phase
Time of condensation	Early mitosis	Late mitosis	Remains condensed in interphase
Concentration of expressed genes	Low	High	None
Enriched repeated DNA sequences	Lines (L_1)	Sines (Alu)	Satellite

AT, adenine thymidine; GC, guanine Cytosine.

DA/DAPI Staining Method

4′6-Diamidino-2-phenylindole (DAPI) is a fluorochrome that, when combined with a nonfluorescent counterstain, distamycin A (DA), shows bright fluorescence at secondary constrictions of chromosomes 1, 9, 16, the proximal 15 (15p11), and distal Yq (Fig. 29.20). It is particularly useful in identifying marker chromosomes that have originated from 15p, which fluoresce brightly after DA/DAPI staining. Both DAPI and DA bind preferentially to AT-rich DNA.

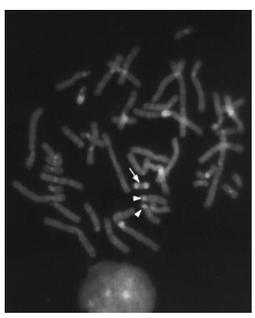


FIGURE 29.20. Cell with an i(15p) marker after DA/DAPI staining. The paracentric heterochromatin of chromosomes 1, 9, 16, Yq, and 15p (*arrowheads*) stain brightly. An extra marker chromosome (*arrow*) derived from 15p is also shown to fluoresce after DA/DAPI staining.

Slides are flooded with DA (DA) 0.2 mg/mL of McIlvaine's citric acid, Na_2 HPO_4 buffer at pH 7.0), covered with a coverslip, and incubated at room temperature for 15 minutes. The coverslip is removed and the slide rinsed with buffer at pH 7.0.

The DAPI solution (0.2 to 0.4 μg/mL) is added to the slide, and the slide is coverslipped and incubated in the dark at room temperature for 15 minutes, rinsed, mounted in buffer, and sealed. The slides are observed with a fluorescence microscope using a 360-nm excitation filter and a 460-nm emission filter.

Nucleolar Organizing Region Bands

The specific chromosome region in the satellite stalk that organizes the nucleolus, the nucleolar organizing region (NOR) may be observed in metaphase chromosomes by a silver impregnation method (Ag-NOR banding) (Fig. 29.21). This technique stains nonhistone residual proteins that appear to be located specifically at functional NORs and localizes the areas coding for the 18s and 28s ribosomal RNA. Each of the acrocentric chromosomes may bear an active NOR. Usually not all 10 of the acrocentric NORs are active in any one cell. NORs may vary in size and in some cases, double NORs may be present in a chromosome. These variations are consistent and heritable from parent to child.

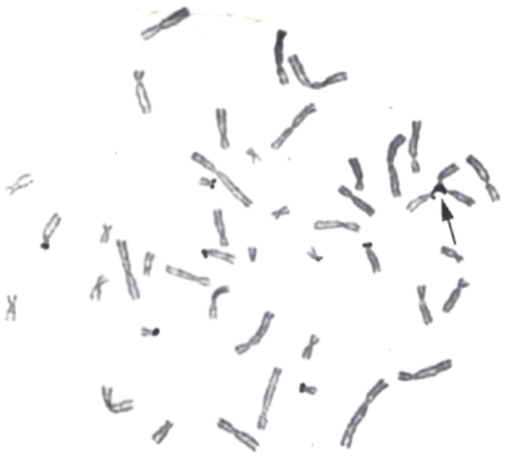


FIGURE 29.21. The active nucleolar organizing regions (NOR) [Ag-NOR staining (silver impregnation method)] are seen as darkly staining areas at the ends of each of the 10 acrocentric chromosomes. *Arrow* indicates three acrocentric chromosomes in association and the NOR regions of the three fused.

Replication Banding

Early cytogenetic studies using ^{3}H-thymidine and autoradiography demonstrated that different chromosomes and regions of chromosomes completed their replication at different times during the S phase. More recently, it has been possible to demonstrate these replication differences and the differences in replication time of G (late) and R (early) bands using BrdU, an analog of thymidine. When BrdU is added to the culture, it is incorporated into DNA during the cell cycle. If the cells are treated with Hoechst 33258 and subjected to photolysis and differential staining with acridine orange (RFA technique) or with Giemsa (FPG technique), R-bands are produced. BrdU also interferes with chromosome condensation. This banding technique is often used to produce R-bands in previously synchronized cultures. Because R-bands have completed their replication early in S phase, the regions incorporating BrdU during the last half of S will be G-bands. Thus, the R-bands will be dark and the G-bands light after FPG staining. Alternatively, the same staining procedure may be used to show G-bands if the cells are grown initially in BrdU and thymidine is added late in replication (Fig. 29.22).

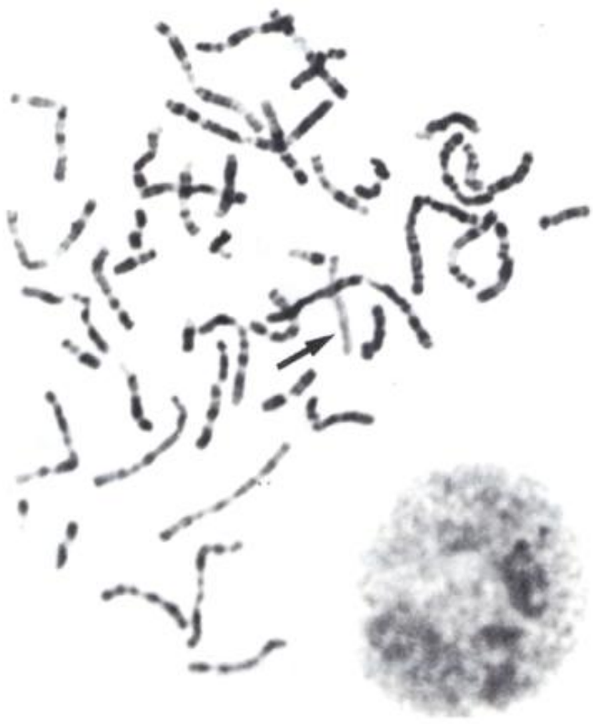


FIGURE 29.22. Replication banding (RBG). Banding after treatment of the cell for the last 5 hours with bromodeoxyuridine. This technique is frequently used to demonstrate the late-replicating X (*arrow*). The chromosomes show an R-banding pattern.

The BrdU/FPG method is most useful in demonstrating the differences in the two X chromosomes because the late-replicating X chromosome stains light and the early-replicating one dark with this method. However, much of this work has been supplanted by molecular studies of X-inactivation (see below).

FPG Method

Lymphocytes are grown for 48 to 54 hours in RPMI 1640 and are synchronized with excess thymidine (concentration 1.24×10^{-3} M) for 17 hours. The block is released after rinsing by resuspending in TC199 plus BrdU (10^{-3} mol/L) for 4.5 hours, and cells are harvested without colcemide exposure.

Slides are stained with Hoechst 33258 (2.5 g/mL) for 15 minutes. The slides are exposed to ultraviolet light for 120 to 150 minutes and stained with Giemsa (1%) for 5 minutes.

High-Resolution Banding

Although most laboratories use metaphase chromosomes for routine chromosome analysis, the need for procedures that demonstrate subtle gains or losses of single bands has led to the development of methodology for high-resolution imaging of chromosomes. Identification of small deletions involved in Prader-Willi (PWS), Miller-Dieker, and other deletion syndromes requires the use of methods that will provide chromosomes with identifiable bands on the order of 800 to 1,000 bands. High-resolution chromosomes may be obtained by using compounds such as ethidium bromide or actinomycin D, which interfere with the condensation of the chromosomes and produce longer chromosomes with subdivided bands. Alternatively, Yunis' method of synchronization of cells (which uses antimetabolites to arrest cells in S phase) may be used, followed by release with thymidine for specific periods to accumulate cells in particular stages of prophase or prometaphase. Although standard metaphase chromosomes demonstrate 400 to 500 bands, high-resolution methods may resolve 800 to 1,200 bands in prometaphase or prophase chromosomes. (The latter technique may also be used in conjunction with replication banding with BrdU to give high-resolution R banding.) The chromosome procedures may also be altered by increasing hypotonic treatment and reducing colcemide time to permit maximum length and greatest spreading of the more extended chromosomes.

High-resolution chromosomes are not only extended compared with metaphase chromosomes but also have a greater number of bands. Both light and dark bands seen at the 400-band level may be resolved at the earlier stages of division into multiple light and dark bands (Fig. 29.23).

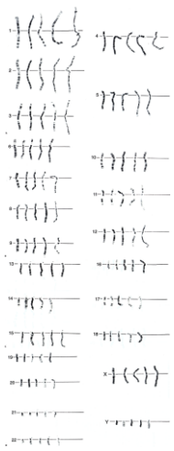


FIGURE 29.23. A-D: High-resolution banding of human chromosomes. The first chromosome of each series is shown at approximately the 475-band level, with increasing length and band resolution as the chromosomes become more extended in prometaphase.

Method of Synchronization for High-Resolution Chromosomes

After incubation of leukocyte cultures grown in RPMI 1640 for 72 hours, methotrexate (final concentration 10^{-7} mol/L) is added to the cultures. Seventeen hours later, the incubated cells are released from the block by rinsing in RPMI 1640 and are grown for an additional 5 to 5.5 hours, treated with colcemide (0.05 µg/mL at room temperature) for 10 minutes, and harvested. Alternatively, the block may be released with 0.1 ml thymidine (1×10^{-3} mol/L) added to each 10-mL culture.

Ethidium Bromide Method for Prometaphase Chromosomes

Leukocyte cultures grown for 72 hours are harvested by addition of 0.05 mL of 6 mg/mL of colcemide and 0.1 mL of a 1 mg/mL solution of ethidium bromide for 2 hours. The cells are fixed and processed in a standard manner.

MOLECULAR CYTOGENETICS

Part of "29 - Basic Cytogenetics"

In the past 10 years, the burgeoning field of molecular cytogenetic has had a monumental effect on the study of chromosomes. Before the introduction of this methodology, chromosomes could only be analyzed using conventional methods already detailed (i.e., G-banding). This limited both our innate understanding of the structure of chromosomes and the resolution allowing detection of cytologic changes. These molecular cytogenetic methods include both FISH as well as the utilization of microsatellite markers to define chromosome structure (29).

Overall, the objective of *in situ* hybridization is the hybridization of a labeled, denatured, single-stranded DNA probe to a target DNA segment that has been denatured in place on a microscope slide. Initial studies on metaphase chromosomes (30) were done using radioactive DNA probes. This had limited application (and almost no clinical application) because of the difficulty involved with the technology. Many of the studies involved repetitive DNA sequences, although there were limited studies for gene localization. However, the technology became much more accessible and widespread after the initial work of Pinkel et al. (31) with nonradioactive *in situ* hybridization. Most probes (and almost all commercially obtained probes) now are directly labeled with a fluorochrome and can easily be detected after hybridization. Probes generated for research studies may be labeled with biotin or digoxigenin and detected indirectly with fluorochrome-labeled antibodies. The utilization of fluorescent probes is combined with fluorescent microscopy, in which one or more probes, each labeled with a different color, can be detected. New methodology (see below) can now be utilized to have the chromosomes fluoresce with 24 different colors, one for each chromosome pair. Less frequently, hybridization and detection of probes are performed using enzymatic methods.

The technology has advanced so that there are many different types of DNA probes that can be used for FISH analysis. These include (a) repetitive DNA segments, the most common which is alpha-satellite DNA found at the centromere regions of all chromosomes; (b) whole chromosome libraries, which uses DNA specific to one chromosome and the entire DNA from that chromosome is used as a probe for hybridization; and (c) and unique (single-copy) sequences of DNA that can be used as a probe for hybridization. Many applications of these probes have been developed, making the possibilities of this field endless. For example, the use of chromosome paints and libraries has been extended to include spectral karyotyping and multicolor FISH, which uses 24 colors to study metaphase chromosome. This technology has become increasingly useful in analyzing preparations from cancer specimens and markers and *de novo* unbalanced rearrangements in constitutional abnormalities. Single copy probes can vary from commercial probes (e.g., those commonly used to detect microdeletion syndromes) to home-brew probes, which can range in size from <1 kb to >1 MB [e.g., yeast artificial chromosomes (YACs)]. The utilization of these home-brew probes has become greater over the past several years especially with the advent of a project to create a map in which bacteria artificial chromosomes (BACs) are localized to every 1 MB in the genome. One unique

type of probe that has been developed is the subtelomeric probe, which is a chromosome-specific probe localized to subtelomeric regions in the chromosomes. This type of probe has been shown to be very useful in the detection of subtle rearrangements involving the ends of chromosomes. The utilization and specific applications of the above mentioned probes are discussed in more detail in the next several chapters (32,33).

Listed below are basic protocols for labeling probes as well as for hybridizing them to the target of interest. Only general protocols are listed:

Bionicking cosmids, BACs, plasmids, phage, and YACs:

In a 500-µL Eppendorf tube, add:

1 µg cosmid, BAC, plasmid or phage DNA

× µL distilled H_2O

10 µL dNTP deoxy-Nucleotide Triphosphate

10 µL enzyme

100 µL total

(If YACS are used, 20 µL YAC DNA is utilized in a total of 250 µL.)

After vortexing and spinning each tube, place each tube into in a 15°C water bath for 1 to 2 hours.

Add 10 µL stop buffer, vortex and spin and store at -20°C until use.

Hybridizing cosmid, BAC, plasmid, and phage probes:

Concentrate the probe overnight at -20°C:

50 ng labeled probe

15 µL Cot-1 DNA

1/10 volume 3M sodium acetate, pH 5.2

2.5 × volume cold 100% ethanol

The concentrated probe is spun at 15,000*g*, aspirated, and dried on a 37°C hot plate.

The pellet is resuspended in 10 to15 µL of 50% hybridization buffer, vortex, spun and placed in a 37°C water bath for 30 minutes. It is then allowed to reanneal in a 37°C water bath for 3 hours.

The denatured probe is then hybridized to the denatured slide at 70°C for 2 minutes and then placed at 37°C overnight. The price of treatments will vary depending on the target DNA to be hybridized.

The slides are then washed in 50% fomamide at 43°C for 15 minutes, then in 2XSSC at 37°C for 8 minutes.

The biotin is detected using a fluorescein-labeled avidin.

(This protocol will change depending on the type of probe, the target DNA, and the source of the probe. Most commercial probes are direct labeled and this last step is not necessary.)

All the discussed probes can be hybridized to a variety of target DNA. The most common and familiar target is the metaphase chromosome. Interpretation of the FISH studies on these specimens is very straightforward and metaphase chromosomes are the most commonly utilized specimen. FISH can also be used on interphase cells. The interphase cells can be either from direct preparations (e.g., amniotic fluid cells used to detect aneuploidy) or from cultured cells. Additionally, preparations from tissue specimens can be utilized, where interphase cells can be studied either as dispersed cells or in paraffin sections, where the original tissue architecture is intact. This latter methodology is extremely useful for Her2-neu analysis in breast cancer.

Another molecular methodology that can be utilized for the analysis of chromosomes is the use of the polymerase chain reaction (PCR) of highly polymorphic di-, tri-, and tetranucleotide microsatellite markers. As these can be used to detect the presence or absence of specific DNA sequences, they are an effective way to study chromosome structure when parental and proband DNA is available. In combination with FISH, microsatellite markers can delineate both deletions and duplications. In cases of nondisjunction, they are the most efficient means for determining the origin of nondisjunction.

The advances in molecular cytogenetics in the past several years have greatly recast the way clinical and research cytogenetic studies proceed. As is demonstrated in the following chapters, its use has influenced the evaluation of somatic abnormalities seen in neoplasia as well as the study of constitutional abnormalities (e.g., detected postnatally in peripheral blood specimens or prenatally in amniotic fluid and chorionic villus samples).

THE SEX CHROMOSOMES

Part of "29 - Basic Cytogenetics"

In humans, females have two X chromosomes and males an X and a Y chromosome. The fact that this difference in the number of X chromosomes did not result in differences in expression of X-linked genes between males and females led H. J. Muller in 1931 to propose the concept of dosage compensation. In humans and most mammals, dosage compensation is achieved by inactivation of excess copies of the X chromosome, and equal amounts of gene product result from transcription of the single active X in both males and females (34,35).

The X Chromatin Body

In 1948, Murray Barr and his graduate student E. C. Bertram identified a nucleolar satellite that was present in histologic preparations of neurons of female cats but not of male cats. After further verification on human preparations, they called the sex-specific chromatin mass “sex chromatin.” (Sex chromatins are now also called Barr bodies.) Subsequently, Barr observed that individuals with ovarian dysgenesis (Turner syndrome) had male-type nuclei—lacking a sex-chromatin body—and that biopsies from males with Klinefelter syndrome were similar to those of normal females in that a sex-chromatin body was present. When the chromosome defects in these syndromes were observed, the correspondence of the number of Barr bodies to the number of inactive X chromosomes became evident.

Further contribution to an understanding of X chromatin came with the demonstration of asynchronous replication of the sex chromosomes, one of the two X chromosomes replicating later in the S phase, and of the mosaicism in the expression of X-linked genes.

In humans, females heterozygous for glucose-6-phosphate dehydrogenase (G6PD) deficiency demonstrated two types of erythrocytes, and in the mouse, coat color variegation resulting

from X-linked coat color mutants or autosomal genes translocated to the X chromosomes occurred in female mice but not in male mice.

These observations led Lyon in 1961 to propose the single active X hypothesis that one of the two X chromosomes in each cell of a female is transcriptionally inactive, that inactivation occurs early in embryogenesis at which time random inactivation of maternally and paternally derived X chromosomes occurs, and that inactivation in the lineage of an X is stable and is transmitted to all descendants (clonal).

This hypothesis received confirmation from a variety of experiments demonstrating mosaicism in the expression of X-linked genes in mice, humans, and other mammals. Single cells cloned from tissue cultures derived from females heterozygous for G6PD demonstrated either A or B electrophoretic variants of G6PD, but never both. Similar behavior of other X-linked enzymes and the presence of two cell populations in heterozygous females who carry the gene for disorders such as ocular albinism and hypohidrotic ectodermal dysplasia provided additional confirmation and acceptance of the Lyon hypothesis. X-inactivation occurs in all mammals, including marsupials but not in other groups (36,37).

The Timing of Inactivation

The onset of inactivation may be demonstrated by the appearance of the sex-chromatin body, the change to asynchronous and late replication of one of the X chromosomes, hypoacetylation of histone proteins, or the appearance of differences in X-linked enzyme activity. These events indicate that the typical euchromatin on the X chromosome converts to heterochromatin. These changes may be shown directly at different stages of embryogenesis or indirectly by the use of chimeric mice. By transplanting single cells from an embryo heterozygous for pigment mutants at different stages of embryogenesis into blastocysts carrying a different mutant, it has been possible to demonstrate (a) lack of inactivation by the appearance of three different colors in patches or (b) if inactivation has occurred, by the occurrence of only two of the pigment mutants.

Using the tools described, it has been shown that inactivation does not occur in all tissues simultaneously but occurs at different times in cell lineages as differentiation occurs. Thus, in the blastocyst that occurs at 3 days in the mouse, X-inactivation is limited to the trophectoderm. Inactivation occurs between 6 and 13 days in the inner cell mass and at the differentiation of the primitive ectoderm. Inactivation in the germ-cell progenitors appears to occur by day 12, but, because both Xs are active in the mature oocyte, reactivation of the X chromosomes must occur before meiotic prophase.

Exceptions to the Lyon Hypothesis

Nonrandom Inactivation

Although the tenets of the Lyon hypothesis as previously explained still hold today, there are some notable exceptions to these rules. Although inactivation is random with respect to the parental origin of the X, there are cases that demonstrate nonrandom X-inactivation. Exceptions to random inactivation of maternally and paternally derived X chromosomes occur in apparent nonrandom inactivation of X chromosome abnormalities, in nonrandom inactivation of paternal X chromosomes in trophoblastic tissues, and in nonrandom inactivation of X chromosomes bearing particular X-linked genes.

Apparent preferential inactivation of X-chromosome abnormalities such as isochromosome X, ring X, and deleted X occur regularly in humans. In each case, random inactivation of the normal X would result in cells nullisomic for a portion of the X chromosome and inviability or lessened viability. In balanced X-autosomal translocations, random inactivation appears to be followed by selection, presumably as a result of the spreading of inactivation into autosomal segments. Cells with such autosomal segments inactivated will have reduced viability, and, as a result, cells with the translocation active and the normal X inactivated usually predominate.

Conversely, in unbalanced X-autosome translocations, the abnormal X appears to be preferentially inactivated as a result of selection of cells with a balanced karyotype. In these instances, spreading of inactivation to the autosomal segment may be incomplete, and phenotypic expression of the extra segment may occur.

Additionally there are some cases in which individuals have two normal X chromosomes but still show preferential inactivation of one particular X chromosome. Some of these cases have been shown to be owing to a mutation in the *Xist* (X inactive specific transcripts) gene, whereas the cause for other cases is still unknown. Many of these cases can be traced within families.

Preferential expression of the maternally derived X chromosome has been shown to occur in tissues of marsupials and in the extraembryonic membranes of rodents and humans. Unlike X rearrangements, the preferential inactivation of paternal alleles does not result from selection but appears to be initial nonrandom inactivation owing to differences in parental origin—“imprinting” (38,39).

Active Genes

Implied in the Lyon hypothesis is that all the genes on the inactive X chromosome are inactive; however, this is not always the case. Early studies indicated that the inactivated X chromosome is not completely inactive. The pseudoautosomal region, the distal portion of the Xp (a segment of the X that undergoes recombination with the Y chromosome) escapes inactivation in humans and mice. In this region is MIC2, a cell surface antigen that is located at the proximal end of the pseudoautosomal region in humans. Several DNA probes have been localized to the human pseudoautosomal region with DXYS14 at the telomeric end. Steroid sulfatase (STS), a gene resulting in X-linked ichthyosis, and the Xga blood group map close to the pseudoautosomal region and escape inactivation in humans. More recently, it was shown that this phenomenon is much more widespread (40).

In a study of X-linked genes, Brown et al. (41) used a rodent/human somatic-cell hybrid to study the expression of 33 X-linked genes. They were able to isolate both active and inactive X chromosomes individually in the rodent/human somatic-cell hybrids. Their studies indicated that six of the 33 genes showed that they escaped X-inactivation, and each demonstrated expression

from both the inactive as well as the active X chromosome. Although many of the genes in this initial study were chosen because of some prior knowledge concerning their inactivation pattern, four of the 17 genes selected, without prior knowledge, were found to escape inactivation. This suggests that a higher proportion of genes may escape inactivation than was originally anticipated. More recent evidence showed that greater than 16 genes are known to escape activity (42). It has been estimated that as many as one third of all the X-linked genes may escape inactivity.

Permanence

One last exception to the Lyon hypothesis concerns the fact that inactivation is permanent. Although one X chromosome is always inactive in somatic cells, this inactivation is reversible in the development of the germ cells. In female germ cells, both X chromosomes are active. In contrast, the single X chromosome in male germ cells is inactive.

DETECTION OF X-INACTIVATION

As explained above, initial studies utilized BrdU incorporation for the cytologic identification of the inactive X chromosome. More recent work provided molecular methodology to quantitate more precisely X-chromosome inactivation. The easiest and most common method of X-chromosome inactivation involves the use of differential methylation of the androgen receptor gene (43). A polymorphic CAG repeat in the first exon is amplified by PCR and selective digestion of the DNA with a methyl-sensitive enzyme can distinguish the active and inactive X chromosome. Other methods that parallel this procedure include the use of a transcribed polymorphism of the *Xist* gene (44), a transcriptional polymorphism of gene *p55* (45) and differential methylation at the FMR-1 locus (46).

Inactivation Center

Studies of X-autosome translocations in mice and humans suggest that inactivation proceeds from a region controlling inactivation, the inactivation center, and spreads to the rest of the chromosome. Thus, in reciprocal translocations between an X and an autosome, only one of the two segments will show inactivation. The X controlling element, the X_{ce} locus, is considered to be the major controlling center for X-inactivation in mice. Evidence from allelic variants of X_{ce} suggests that this locus controls the probability of inactivation of the X chromosome and can also control the preferential inactivation of paternal X chromosomes in extraembryonal membranes.

In humans, Therman proposed the presence of an inactivating center between Xq11.2 and Xq21.1 on the basis of failure of inactivation in the absence of this segment in various deletions and rearrangements and on the evidence of the occurrence of bipartite sex chromatin when this region of the X is duplicated (Fig. 29.24).

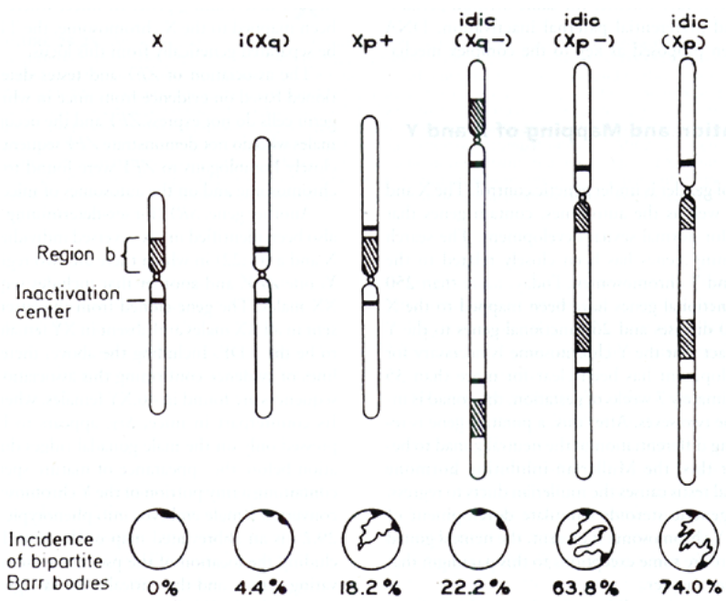


FIGURE 29.24. The location of the X-inactivation center of Xq and the presumably active (b) region on Xp in six structurally different X chromosomes, plus the incidence and hypothetical structure of bipartite Barr bodies formed by them. (From Therman E, Sarto GE. Inactivation center on the human X chromosome. In: Sandberg AA, ed. *Cytogenetics of the mammalian X chromosome, part A*. New York: Alan R. Liss, 1983:315–325, with permission.)

The inactivating center, identified by Hunt Willard, appears to be at Xq13 in humans. The *Xist* gene is transcribed from the X_{ce} on the inactive X chromosome. The *Xist* gene transcribes a large mRNA, approximately 15 to 17 kb, which is retained in the nucleus. This mRNA seems to coat the inactive X. Before inactivation, there appears to be only a dot of RNA on each chromosome. However, after inactivation, the level of *Xist* RNA is increased and the inactive chromosome appears to be coated. This increased level of *Xist* RNA appears not to be owing to increased transcription but rather to a stabilization of transcription.

As described previously, once inactivation has been initiated, it spreads outward on the inactive X. Early cytologic evidence from X-autosome translocations in humans and mice suggests that autosomal regions close to the point of translocation are more likely to be inactivated. Studies of similar translocation in human chromosomes suggest that inactivation may skip some regions and inactivate others farther from the translocation point. More recent quantitative work, (47) confirmed these cytologic observations. This group studied 20 transcribed sequences in 4q24-qter, using reverse transcription PCR analysis. The der(X)t(X;4)(q22;q24) had been isolated in a somatic cell hybrid, and these studies indicated three known genes and 11 expressed sequence tags (ESTs) were not expressed (inactive), three genes and three ESTs escaped inactivation. This study molecularly confirmed the cytologic conclusion that the spreading of inactivation occurred, but is incomplete and noncontiguous (48).

To attribute X- inactivation to any mechanism, it is necessary to explain the initiation, promulgation, and maintenance of X-inactivation and to consider whether the same method pertains to both random and preferential paternal inactivation. DNA methylation has been proposed as key to the complex mechanism involved.

Sex Determination and Mapping of X and Y Chromosomes

The determination of gender is under genetic control. The X and Y chromosomes, as well as the autosomes, contain genes that must be functional for normal sexual development. The search for the sex-determining genes has been closely related to the mapping of the X and Y chromosomes. Today, more than 250 diseases and 300 functional genes have been mapped to the X chromosome and 20 diseases and 26 functional genes to the Y chromosome. The fact that the Y chromosome is necessary for normal sexual development has been clear for more than 35 years. Before approximately 7 weeks of gestation, the gonad is indistinguishable in the two sexes. After this, a putative gene is responsible for inducing differentiation of the neutral gonad to become a testis. After this, the Mullerian-inhibiting hormone secreted from the fetal testis causes the mullerian ducts to regress, and released androgenetic steroids stimulate development of male genitalia. If no Y chromosome is present, the neutral gonad forms an ovary. There are some exceptions to this paradigm that are covered in another chapter.

There have been several candidates for the sex-determining gene or genes, described previously. Initially, the H-Y antigen, a male-specific antigen, was for many years considered to be the testis-determining factor (TDF) subsequent to this, *ZFY*, a zinc finger gene on the Y chromosome, was enthusiastically received as the possible TDF. Both are now considered to be related to sperm maturation rather than to the initial differentiation of the testes.

The H-Y antigen was initially demonstrated when skin transplanted from male isogenic mice was rejected by females of the same strain. The antigen has also been demonstrated by serologic tests using antisera produced after inoculation of females with male cells or by cytotoxicity assays with T cells. When interspecific comparisons demonstrated conservation of the soluble antigen in evolution, it was hypothesized that the H-Y antigen and TDF were the same. More recent evidence indicates that the serologic and transplantation antigens are coded by two different genes and further evidence indicates that neither is the TDF. A gene with properties similar to the serologically detected antigen has been localized to chromosome 6, and its presence in females as well as X-chromosome abnormalities removes it from contention as TDF. The T-cell toxicity antigen has been localized to the long arm of the Y, whereas evidence from XX males indicates that TDF is on the short arm of the Y chromosome.

The identification of XX males in whom the distal portion of the Yp chromosome has been transferred to the X and t(Y;15) females with a corresponding deletion of the Y chromosome led to the cloning of a 140-kb segment of the Y chromosome that coded for a regulatory protein, a zinc finger protein (the human gene is designated *ZFY*). Although it was suggested that the H-Y antigen may be the biochemical expression of the TDF (because both have at least one locus in or close to the pseudoautosomal region), evidence accumulated to indicate that *ZFY* is not the sex-determining gene. In mice, where the H-Y antigen has been mapped to the X chromosome, the TDF (Tfy in mice) can be separated genetically from this locus.

The association of *ZFY* and testes determination was questioned based on evidence from mice in which fetal testes lacking germ cells do not express *ZFY* and the occurrence in humans XX males who do not demonstrate *ZFY* sequences. In addition, genes closely homologous to *ZFY* were found to be present on the X chromosome and on the autosomes of mice and marsupials.

Another gene, *SRY* (for *s*ex-determining *r*egion on the *Y*), had also been identified in sex-reversed individuals, i.e., a female with X and a t(Y;22) in which there are two regions deleted from the Y, one *ZFY*, and another that includes sequences present in all XX males. The gene cloned from the latter region, *SRY*, is present in all XX males and absent in XY females, and now is known to be the TDF. Including the above, there are several different lines of evidence confirming this association. Mutations in *SRY* sequences are found in 46,XY females, where *SRY* is not deleted. Its counterpart in mice, *Sry*, appears to be testis specific, expressed only on the male gonadal ridges during testis differentiation before the appearance of mature sperm. Transgenic mice containing a tiny portion of the Y chromosome, containing *SRY*, converted female embryos into phenotypic normal mice. Figure 29.25 is an abbreviated map of the X and Y chromosomes including the location of the pseudoautosomal region, the inactivating center, and the critical area—a segment that must be intact in both X chromosomes to permit normal ovarian development in females. The *SRY* gene was one of the first major genes to be cloned by *positional cloning* as a result of the Human Genome Project (49,50).

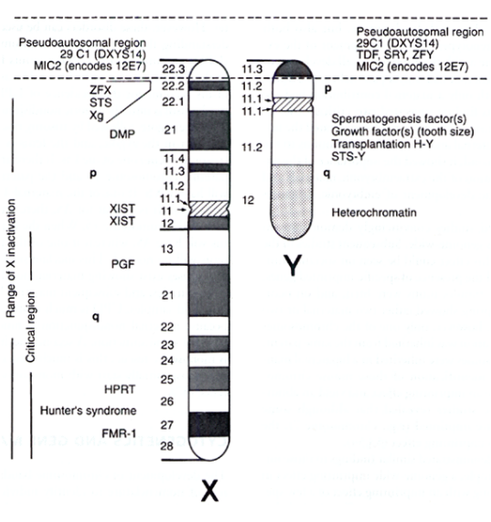


FIGURE 29.25. An abbreviated map of the X and Y chromosomes shows loci and areas related to sexual differentiation. ZFX and ZFY, loci for zinc finger protein on X and Y chromosomes; DXYS14, pseudoautosomal locus defined by probe 29C1, the most distal of a group of probes localized to this region; MIC2, pDP1002 encoding cell surface antigen 12E7; pseudoautosomal region, region of pairing of X and Y chromosomes (this region is not inactivated); STS and Xg, steroid sulfatase and Xga blood group (both are located close to the pseudoautosomal region and both escape inactivation); SRY, sex-determining region of the Y chromosome; HY, locus-encoding transplantation histocompatibility HY; MIC2, locus-encoding histocompatibility cell surface antigen 12E7; gonadal development, regions of Xp and Xq that, when deleted, result in gonadal dysgenesis; critical region, region whose integrity is important in gonadal development of balanced X/autosomal translocation carriers; inactivation center, region controlling X inactivation.

The usefulness of chromosome abnormalities of X and Y and of cytogenetic analyses of unusual XX males and XY females in resolving the question of the location of the sex-determining gene cannot be discounted. The usefulness of chromosome abnormalities to mapping of the human genome is discussed further in the following section.

Imprinting

Although an equal number of autosomal chromosomes are inherited from both the mother and the father of an individual, the genetic contribution from both parents may be different. Genomic imprinting, which refers to the differential modification of both the paternal and maternal contributions to the zygote, results in the differential expression of paternal genes. For example, in some genes the paternal and maternal alleles function differently. The best example of this is in the proximal region of the long arm of chromosome 15, where the *SNRPN* gene is functioning in the paternal genome only, whereas the *UBE3A* gene is functional in the maternal genes only. Loss of the paternal genes in 15q results in PWS, and loss of the maternal genes results in Angelman syndrome (AS). These syndromes are discussed in detail in a later chapter.

Most of the imprinting studies were first described in mouse models and can best be described using these models. Zygotes containing either two maternal haploid complements, two paternal haploid complements, or one complement from each parent have been studied. To create these complements, either the male or female pronucleus complement has been microsurgically removed from the fertilized oocyte and the desired complement constructed. Logically, it would follow that if the paternal complement had no effect on development, then all three scenarios would yield equivalently normal zygotes. However, these studies indicated that only the complement with both a maternal and paternal chromosome complement developed normally. In zygotes with maternal or paternal complements only, not only did

the fetus demonstrate abnormal development, but also both classes had different phenotypes. Good development of the extraembryonic membrane and lack of the inner cell development were noted in the zygotes that had only a paternal contribution, whereas the zygotes with only a maternal contribution had the reverse phenotype. This finding suggested two critical observations regarding imprinting. First, it clearly demonstrated the importance of both the paternal and maternal contributions to the developing embryo. Second, it showed the importance of the paternal loci in the formation of the extraembryonic materials and the maternal loci in the development of embryonic structures (39,51,52).

The aforementioned studies convincingly demonstrate an imprinting effect that is genome-wide. Subsequent studies in the mouse indicated that this effect could be seen on specific chromosomes and indicated the presence of specific imprinted genes. Mice heterozygous for translocations were bred, and on occasion, the resultant offspring showed either two maternal or two paternal chromosomes; however, only one of the chromosomes involved in the translocation was inherited from the same parent. The rest of the chromosomes were inherited in a biparental manner. This allowed the identification of those mouse chromosomes that demonstrate an imprinting effect and yield an abnormal phenotype. These studies revealed that although some mouse chromosomes are imprinted (e.g., chromosome 7), the vast majority shows no imprinting effect (53,54).

Studies in humans demonstrated similar findings to those initiated in mice. For example, a genome-wide imprinting effect in humans can be seen along with an imprinting effect of a few specific chromosomes. Genome-wide imprinting effects include hydatidiform moles and ovarian teratomas. Moles are an abnormal product of conception in which the chorionic villi are hydropic and hypoplastic and have a predilection to become malignant. There are two types of moles, complete and partial. Complete moles have no fetal development or normal villi, whereas in partial moles, there is fetal development, but the villi are abnormal. The chromosome complement in complete moles is 46,XX with the entire complement of androgenic origin. Partial moles are triploid (69 chromosomes) with two of the three complements paternal in origin. In contrast to these, ovarian teratomas are benign cysts with two maternal and no paternal chromosome complement. They are thought to represent the parthenogenetic outcome of an unfertilized oocyte. The ovarian teratomas are not malignant but are very well differentiated (55).

In addition to the genome-wide imprinting abnormalities, there are also chromosome-specific imprinting effects in humans. In some cases, disorders have been associated with imprinting of specific chromosomes, such as for chromosome 15 (PWS and AS), chromosome 11 (Beckwith-Wiedemann syndrome), and chromosome 6 (neonatal diabetes). For other chromosomes, nonspecific findings have been associated with imprinting (e.g., chromosomes 7 and 14). Many of the imprinted genes in humans appear to be clustered, such as in proximal 15q (15q11-q13), and they can also be found in the orthologous region in mice (chromosome 7) (56).

The phenotypic description of the two most widely studied imprinting syndromes, PWS and AS, is discussed in a later chapter. However, these disorders can be used as a paradigm for understanding the mechanism of autosomal uniparental disomy (UPD) and imprinting. UPD accounts for approximately 25% of the PWS cases, yet only 3% of AS cases. Two major mechanisms account for the incidence of UPD. The most common mechanism is initially meiotic nondisjunction leading to trisomy 15 in the zygote, followed by trisomy rescue. Three chromosome 15s start in the zygote, and the fetus is rescued by eliminating (trisomy rescue) one of the 15s. If there are two maternal and one paternal chromosome 15 and the paternal 15 is lost, the fetus will have PWS. If one of the maternal 15s is lost, the fetus will be normal. It is similar for AS; there are two paternal and one maternal chromosome 15. When the maternal 15 is lost, the fetus will have AS, whereas if one of the paternal 15s is lost, the fetus will be normal. This mechanism has been confirmed because cases exist showing maternal nondisjunction in chorionic villus samples, and subsequent maternal UPD detected in amniotic fluid samples. UPD is much more commonly seen in PWS because maternal nondisjunction is much more frequent than paternal nondisjunction. A second mechanism leading to UPD is monosomy rescue; this is much less common because the zygote must initially start with monosomy 15, an unlikely occurrence (57).

CYTOGENETICS AND GENE MAPPING

Part of "29 - Basic Cytogenetics"

The development of chromosome banding methods and a system of nomenclature to identify individual chromosome arms and bands was an important step needed to allow the assignment of human genes and DNA probes to individual chromosomes, regions, and bands (gene mapping). As banding procedures and identification of chromosome variants and chromosome abnormalities became available, techniques of gene mapping were simultaneously developed.

Cytogenetics has made important contributions to gene mapping by a number of different approaches, including the use of classic mapping techniques with chromosome variants and abnormalities used as markers, deletion mapping with and without concurrent dosage studies, somatic cell hybridization, and *in situ* hybridization, especially FISH.

Although assignment of genes to the X chromosome based on sex linkage was possible before cytogenetic studies, the assignment of genes to specific autosomes was possible only after chromosome variants were observed. The first gene assignment, the Duffy blood group, used the segregation of a familial chromosome 1 variant, an extended paracentric heterochromatic band of chromosome referred to then as “uncoiler,” and the concurrent segregation of the Duffy blood group.

The presence of chromosome variants in a family made it possible to establish how often the variant was inherited with specific markers. The degree of linkage of genes on a chromosome is related to their closeness in position on a particular chromosome. Those farther apart are more likely to undergo crossing over and separation at meiosis. Recombination, when genes are close together, may be close to zero; if far apart or on different chromosomes, it will approach 50%. This is the recombination

frequency. Gene mapping is based on the degree of recombination, the frequency serving as a measure of genetic distance, and the relative distance apart of the variant and gene on a genetic map.

Thus, investigators used C-band variants to localize and order a number of genes on chromosome 1. Similarly, segregation in families of translocations or other cytogenetically marked chromosomes have been used to localize genes to chromosomes. For example, it was possible to use a fragile site on chromosome 16 (at 16q21-22), segregating as an autosomal dominant trait, to demonstrate linkage with alpha-haptoglobin by following segregation and recombination of the two in a large family and thus to localize the alpha-haptoglobin gene close to this fragile site.

A second approach to gene mapping, somatic cell hybridization, involves the fusion of human cells, fibroblasts, or lymphoblastoid cell lines with rodent cells. If both parental lines are drug sensitive, the use of selective media will allow growth of the cell hybrids but not the parental cell lines. Fusion is facilitated by use of Sendai virus or polyethylene glycol. Although the hybrid cells initially contain chromosomes of both parent lines, human chromosomes are gradually lost from unstable hybrid cells. Eventually only one or a few human chromosomes are retained in a stable cell line. Cytogenetic recognition of the individual human chromosomes or, alternatively, recognition of a protein associated with a specific chromosome permits identification of a number of unique cell lines. A collection of such cell lines, each with a different chromosome or chromosomes (a cell panel) allows the presence of a particular chromosome to be correlated with a unique gene product. Thus, a specific gene is localized to a chromosome.

More recently, linkage studies using cell hybridization have employed lymphoblastoid cell lines derived from individuals with different chromosome abnormalities to develop panels with multiple overlapping rearrangements of the same chromosome. These facilitate localization of a particular gene to a specific region or band.

The first use of deletions to map chromosomes was in *Drosophila*. A series of chromosome deletions seen in the salivary gland chromosomes were compared with the expression of specific genes, and the position of the gene on the *Drosophila* chromosomes was established. In human chromosomes, deletion mapping has taken several forms. One relates reduced enzyme levels, loss of an allele, and occurrence of a deletion to localize a gene. For example, Ferguson-Smith used a series of abnormalities of chromosome 2 to localize red cell phosphatase (ACP_1) to the region 2p23. Trisomies and duplications have also been used to localize genes by gene dosage studies or by the presence of three different alleles. For example, individuals trisomic for 6p21-6pter with three different histocompatibility haplotypes facilitated the localization of the histocompatibility locus to 6p21 to 6pter.

A panel of abnormalities of chromosome 9, including duplications and deletions, has been used in dosage studies of human red cell adenylate kinase (AK-1). Demonstration of an increase of AK-1 activity by 43% in one duplication permitted the assignment of AK_1 to 9q33-qter. In this case, the chromosome involved had been previously identified by cell hybridization, and dosage studies with the panel of cell lines permitted further localization to a region. A classic demonstration of mapping a disease gene involves the localization of Duchenne's muscular dystrophy (DMD). An Xp deletion in a male with DMD localized this disorder to Xp21, which was later confirmed by a series of females having DMD, with X/autosomal translocations containing an Xp21 breakpoint.

The most direct cytogenetic method of gene localization has been *in situ* hybridization, which permits direct visualization of the location of a probe on a chromosome. This method originated in attempts to localize satellite DNA by hybridizing DNA labeled with tritiated nucleotides to chromosomes that had been previously denatured. After a period of incubation, slides that had been dipped in photographic emulsion subsequent to hybridization were observed. The appearance of silver grains on the slides showed the location of the decaying isotope and hence of the hybridizing DNA.

Although this method was initially used for localizing satellite DNAs and other repetitive DNAs and RNA to chromosomes, Harper and Saunders subsequently adapted the methodology to permit its use with single-copy probes. Their adaptations included the use of dextran sulfate in the hybridization reaction and probes contained in a vector that had been nick-translated to label the DNA with tritiated nucleotides. These adaptations resulted in networking of vectors and accumulation of the label at specific sites on the chromosome. The labeled chromosomes may be photographed for grain location and then banded and rephotographed or the grain developed and the chromosomes banded, usually with a fluorescent dye, and observed simultaneously. A further development of *in situ* hybridization was the use of chromosomally abnormal cells, such as those bearing a translocation, for regional mapping. In such cases, the mapping of the probe to the translocated segment can identify the specific region or band to which the grain is confined (58).

Most recently, hybridization with radioactive probes has been replaced by the use of FISH. With this method, the nick-translated probe is commonly a biotin-labeled probe; random primer, also conjugated with avidin, can be used to label and can be subsequently visualized by use of antiavidin antibodies and immunofluorescence with fluorescein on counterstained chromosomes (this technology was explained in more detail earlier). This method has the advantage that, by use of two separate fluorescent dyes, it is possible to visualize the chromosomes at two different wavelengths, one producing banding and the other demonstrating the localization of the probe. The process is faster and the hybridized probe is more specifically localized (Fig. 29.26). One of the problems with the use of radioactive probes was the occurrence of grain on other chromosomes of the cell as a result of scatter of radioactive emissions, which necessitated grain counting and statistical evaluation of the results. The fluorescence approach is more specific and is useful for both repetitive and single-copy probes. What has become especially efficacious about this technique is the ability to use almost any size probe (1 kb to >1 MB), for mapping. These probes can be labeled and easily mapped to the appropriate chromosome. If a probe is too small to detect a signal, it can be labeled and hybridized

to either cosmid, BAC, or PAC libraries to obtain a larger probe that can be used with FISH (59,60 and 61).

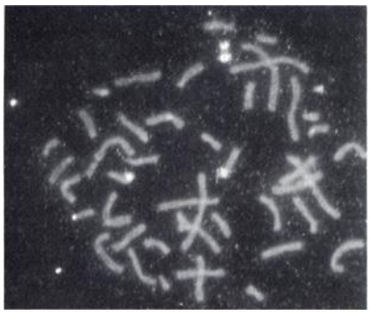


FIGURE 29.26. Localization of alpha-satellite probe from chromosome 15 to both chromosome 15s and to a marker chromosome derived from chromosome 15, i(15p).

The overall usefulness of the fluorescence techniques has not only been demonstrated in metaphase chromosome spreads, but also by its ability to delimit the position of linked genes on a chromosome by using interphase cytogenetics and labeling with several fluorescent dyes emitting in different wavelengths. This provides much greater resolution when trying to map a gene. Fiber-FISH provides even greater mapping resolution. This is a technique in which chromosomes are mechanically stretched and then the probe applied. This provides much greater spatial resolution allowing the precise ordering of closely mapped genes and DNA segments (62).

Gene mapping studies have included combinations of the above techniques with molecular methods. Although chromosome variants were initially useful, the variation in DNA by the use of restriction endonuclease digestion has been so much greater that chromosome variants are now rarely used for conventional linkage studies. In turn, microsatellite markers (STSs, as described previously) are now routinely utilized for these studies, and soon single nucleotide polymorphisms will be routinely used. The cosegregation of these molecular polymorphisms and recombination frequency of these markers in families can be used for genetic linkage studies. It is this approach that has led to the identification of probes diagnostic for genetic disorders and ultimately to their chromosome localization, either by linkage to previously located genes or by *in situ* hybridization.

Since 1973, International Human Gene Mapping Workshops have met regularly to summarize the assignment of genes to chromosomes. From 25 genes mapped in humans in 1973, the conference of 1991, the Eleventh International Workshop, reported 2,325 genes identified and localized (63,64). The data generated at these conferences have been summarized in published reports and include information on the genetic constitution of each of the chromosomes, assignment of mendelian disorders and clinical disorders, and catalogs of cloned genes and DNA fragments. An example of the extent of mapping is shown in Fig. 29.27, where selected anchor loci are indicated for chromosome 1 and a human disease phenotype for that chromosome is also shown. The progress of mapping the human genome is rapidly increasing, with international cooperation in this effort and with molecular biology, computer science, and technology making contributions to progress in this field. It should be noted that the mapping effort has become so extensive that each individual chromosome has individual meeting and workshop reports.

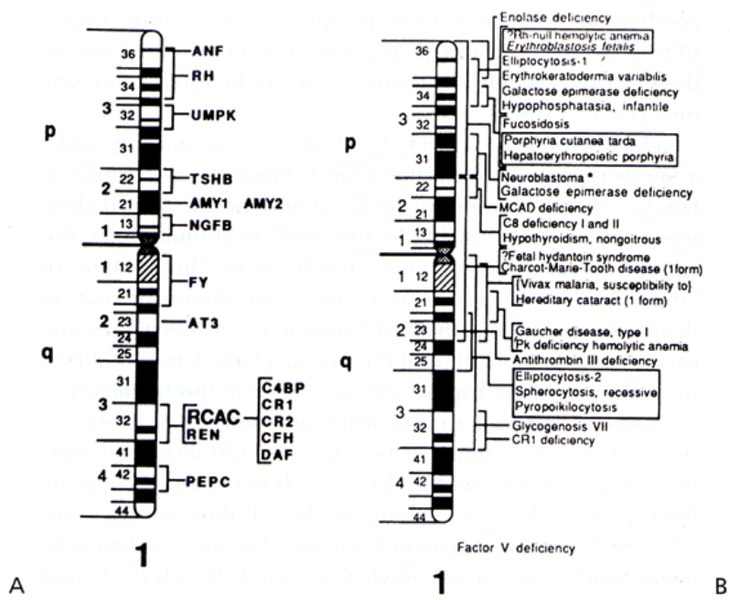


FIGURE 29.27. Gene localization on chromosome 1. **A:** Selected anchor loci. ANF, natriuretic factor; RH, rhesus blood group; UMPK, uridine monophosphate kinase; TSHB, thyroid-stimulating hormone, ß subunit; AMY1, amylase (salivary); AMY2, amylase (pancreatic); NGFB, nerve growth factor ß; FY, Duffy blood group; AT3, antithrombin III; RCAC, regulator of complement activation center; C4BP, complement component 4 binding protein; CR1, complement component 3b; CR2, complement component 3d; CFH, complement factor H; DAF, decay accelerating factor of complement; REN, renin; PEPC, peptidase C. **B:** Disorders for which the mutation has been mapped to chromosome 1. (From McKusick VA. *Mendelian inheritance in man*, 8th ed. Baltimore: The Johns Hopkins University Press, 1988 pg xliii, with permission.)

REFERENCES

1. Hsu TC. *Human and mammalian cytogenetics: an historical perspective*. New York: Springer-Verlag, 1979.

2. Tjio JH, Levan A. The chromosome number of man. *Hereditas* 1956;42:1–6.

3. Harnden DG, Klinger HP. An international system for human cytogenetic nomenclature. In: Harnden DG, Klinger HP, eds. *Birth defects: original article series*, vol. 21. New York: March of Dimes Birth Defects Foundation, 1985.

3a. Mitleman F, ed. *ISCN: Guidelines for cancer cytogenetics, supplement to an international system for human cytogenetic nomenclature*. Basel: S Karger, 1991.

4. Mitelman F, ed. *ISCN: An international system for human cytogenetic nomenclature*. Basel: S Karger, 1995.

5. Benbow RM. Chromosome structures. *Sci Progr* 1992;76:425–450.

6. Marsden MPF, Laemmli U. Metaphase chromosome structure: evidence for a radial loop model. *Cell* 1979;849–858.

7. McGhee D. The structure of interphase chromatin. In: Risley MS, ed. *Chromosome structure and function*. New York: Van Nostrand Reinhold, 1986:1–37.

8. Moyzis RK, Buckingham JM, Cram L.S., et al. A highly conserved repetitive DNA sequence, (TTAGGG), present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* 1988;85:6622–6626.

9. Smith S, de Lange T. TRF1, a mammalian telomeric protein. *Trends Genet* 1997;13:21–26.

10. Ancelin K, Brun C, Gilson E. Role of the telomeric DNA-binding protein TRF2 in the stability of human chromosome ends. *Bioessays* 1998;20:879–883.

11. Muniyappa K, Kironmai KM. Telomere structure, replication and length maintenance. *Crit Rev Biochem Mol Biol* 1998;33:297–336.

12. Pardue M-L, DeBaryshe PG. Telomeres and telomerase: more than the end of the line. *Chromosoma* 1999;108:73–82.

13. Gottschling DE, Stoddard B. Telomeres: structure of a chromosome's aglet. *Curr Biol* 1999;9:R164–R167.

14. Mitchell AR. The mammalian centromere: its molecular architecture. *Mutat Res* 1996;372:153–162.

15. Jabs EW, Persico MG. Characterization of human centromeric regions of specific chromosomes by means of alphoid DNA sequences. *Am J Hum Genet* 1987;41:374–390.

16. Lee C, Wevrick R, Fisher RB, et al. Human centromeric DNAs. *Hum Genet* 1997;100:291–304.

17. Pluta AF, Mackay AM, Ainsztein AM, et al. The centromere: hub of chromosomal activities. *Science* 1995;270:1591–1594.

18. Sullivan BA, Schwartz S, Willard HF. Centromeres of human chromosomes. *Environ Mol Mutagen* 1996;28:182–191.

19. Warburton PE, Cooke CA, Bourassa S, et al. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr Biol* 1997;7:901–904.

20. Korenberg JR, Rykowski MC. Human genome organization: alu, lines, and the molecular structure of metaphase chromosome bands. *Cell* 1988;53:391–400.

21. McIntosh JR, Koonce MP. Mitosis. *Science* 1989;246:622–628.

22. Sumner AT. Chromosome banding and identification. Fluorescence. *Methods Mol Biol* 1994;29:83–96.

23. Caspersson TO. The William Allen memorial award address: the background for the development of the chromosome banding techniques. *Am J Hum Genet* 1989;44:441–451.

24. Verma RS, Babu A. *Human chromosomes: manual of basic techniques*. New York: McGraw-Hill, 1995.

25. Holmquist G, Gray M, Porter T, et al. Characterization of Giemsa dark- and light-band DNA. *Cell* 1982;31:121–129.

26. Bickmore WA, Sumner AT. Mammalian chromosome banding an expression of genome organization. *Trends Genet* 1989;5:144–148.

27. Holmquist GP. Chromosome bands, their chromatin flavors, and their functional features. *Am J Hum Genet* 1992;51:17–37.

28. Craig JM, Bickmore WA. Chromosome bands—flavours to savour. *Bioessays* 1993;15:349–354.

29. Chang SS, Mark HF. Emerging molecular cytogenetic technologies. *Cytobios* 1997;90:7–22.

30. Pardue ML, Gall JG. Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *PNAS* 1969;64:600–604.

31. Pinkel D, Straume T, Gray JW. Cytogenetic analysis using quantitative, high sensitivity, fluorescence hybridization. *PNAS* 1986;83:2934–2938.

32. Lichter P. Multicolor FISHing: what's the catch? *Trends Genet* 1997;13:475–479.

33. Raap AK. Advances in fluorescence in situ hybridization. *Mutat Res* 1998;400:287–298.

34. Heard E, Clerc P, Avner P. X-chromosome inactivation in mammals. *Annu Rev Genet* 1997;31:571–610.

35. Lyon MF. X-chromosome inactivation. *Curr Biol* 1999;9:R235–R237.

36. Gartler SM. Mammalian X chromosome inactivation. *Annu Rev Genet* 1983;17:155–190.

37. Lyon MF. The William Allan memorial award address: X-chromosome inactivation and the location and expression of X-linked genes. *Am J Hum Genet* 1988;42:8–16.

38. Graves JA. Mammals that break the rules: genetics of marsupials and monotremes. *Annu Rev Genet* 1996;30:233–260.

39. Bartolomei MS, Tilghman SM. Genomic imprinting in mammals. *Annu Rev Genet* 1997:31:493–525.

40. Disteche CM. Escape from X inactivation in human and mouse. *Trends Genet* 1995;11:17–22.

41. Brown CJ, Carrel L, Willard HF. Expression of genes from the human active and inactive X chromosomes. *Am J Hum Genet* 1997;60:1333–1343.

42. Carrel L, Cottle AA, Goglin KC et al. A first-generation X-activation profile of the human X chromosome. *PNAS* 1999;96:14440–14444.

43. Schmucker B, Meindl A, Achatz H et al. A PCR based X-chromosome inactivation assay for carrier detection in X-linked immunodeficiencies using differential methylation of the androgen receptor gene. *Immunodeficiency* 1995;5:192.

44. Rupert JL, Brown CJ, Willard HF. Direct detection of nonrandom X chromosome inactivation by use of a transcribed polymorphism in the Xist gene. *Eur J Hum Genet* 1995;3:333–343.

45. Luhovy M, Liu Y, Belickova M, et al. A novel clonality assay based on transcriptional polymorphism of X chromosome gene p55. *Biol Blood Marrow Transplant* 1995;1:81–87.

46. Carrel L, Willard HF. An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. *Am J Med* Genet 1996;64: 27–30.

47. White WM, Willard HF, Van Dyke DL et al. The spreading of X inactivation into autosomal material of an X; autosomal translocation evidence for a difference between autosomal and X-chromosomal DNA. *Am J Hum Genet* 1998;63:20–28.

48. Kay GF. Xist and X chromosome inactivation. *Mol Cell Endocrinol* 1998;140:71–76.

49. Page DC, Fisher MC, McGillivray B, et al. Additional deletion in sex determining region of human Y chromosome resolves paradox of X, t(Y;22) female. *Nature* 1990;346:279–281.

50. Goodfellow PN, Lovell-Badge R. SRY and sex determination in mammals. *Annu Rev Genet* 1993;27:71–92.

51. Hurst LD, McVean GT. Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends Genet* 1997;13: 436–443.

52. Brannan CI, Bartolomei MS. Mechanisms of genomic imprinting. *Curr Opin Genet Dev* 1999;9:164–170.

53. Morison IM, Reeve AE. A catalogue of imprinted genes and parent-of-origin effects in humans and animals. *Hum Mol Genet* 1998;7:1599–1609.

54. Beechey CV. Appendix: imprinted genes and regions in mouse and human. *Results Probl Cell Diff* 1999;25:303–323.

55. Moore T, Reik W. Genetic conflict in early development: parental imprinting in normal and abnormal growth. *Rev Reprod* 1996;1:73–77.

56. Ledbetter DH, Engel E. Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. *Hum Mol Genet* 1995;4:1757–1764.

57. Nicholls RD, Saitoh S, Horsthemke B. Imprinting in Prader-Willi and Angelman syndromes. *Trends Genet* 1998;14:194–200.

58. Magenis RE. Application of structural rearrangements and DNA probes to gene mapping. In: Daniel A, ed. *The cytogenetics of mammalian autosomal rearrangements*. New York: Alan R. Liss, 1988: 855–893.

59. van Ommen GJ, Breuning MH, Raap AK. FISH in genome research and molecular diagnostics. *Curr Opin Genet Dev* 1995;5:304–308.

60. Heng HH, Spyropoulos B, Moens PB. FISH technology in chromosome and genome research. *Bioessays* 1997;19:75–84.

61. Palotie A, Heiskanen M, Laan M, et al. High-resolution fluorescence in situ hybridization: a new approach in genome mapping. *Ann Med* 1996;28:101–106.

62. Heiskanen M, Peltonen L, Palotif A. Visual mapping by high resolution FISH. *Trends Genet* 1996;12:379–382.

63. Human Gene Mapping 11. London Conference (1991). Eleventh International workshop on Human Gene Mapping. London, UK, August 18–22, 1991. *Cytogenet Cell Genet* 1991;58:1–984.

64. Human Gene Mapping 11. London Conference (1991). Eleventh International Workshop on Human Gene Mapping. London, UK, August 18–22, 1991. *Cytogenet Cell Genet* 1991;58;986–2156.

SUGGESTED READINGS

Adolph KW. *Chromosomes and chromatin. Vols. I, II, and III*. Boca Raton, FL: CRC Press, 1988.

Bergsma D, ed. *Chicago Conference: Standardization in human cytogenetics. Birth defects* (Original articles series, vol. II, No. 2). New York: Alan R. Liss, 1966.

De La Chapelle, Sorsa M. Microspreading and the synaptonemal complex in cytogenetic studies. In: *Chromosomes today, volume 6*. Amsterdam: Elsevier Science Publishers, 1997:71–82.

Ellis N, Goodfellow PN. The mammalian pseudoautosomal region. *Trends Genet* 1989;5:389–429.

Morton CC, Christian SL, Donlon TA, et al. Report of the fourth international workshop on human chromosome 15 mapping 1997. *Cytogenet Cell Genet* 1999;84:12–21.

Richmond TJ, Finch JT, Klug A. Studies of nucleosome structure. *Cold Spring Harb Symp Quant Biol* 1983; 47:493–501.

Risley MS. The organization of meiotic chromosomes and synaptonemal complexes. In: Risley MS, ed. *Chromosome structure and function*. New York: Van Nostrand Reinhold, 1986:126–151.

Sandberg AA, ed. *Cytogenetics of the mammalian X chromosome. Part A. Basic mechanisms of X chromosome BEH. (Progress and topics in cytogenetics series, volume. 3A)*. New York: Alan R. Liss, 1983.

Sandberg AA, ed. *The Y chromosome. Part A. Basic characteristics of the Y chromosome. (Progress and topics in cytogenetics series, volume 6A)*. New York: Alan R. Liss, 1985.

Sumner AT. The nature and mechanisms of chromosome banding. *Cancer Genet Cytogenet* 1982;6:59–87.

Thoma F, Koller T, Klug A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* 1979;83:403–27.

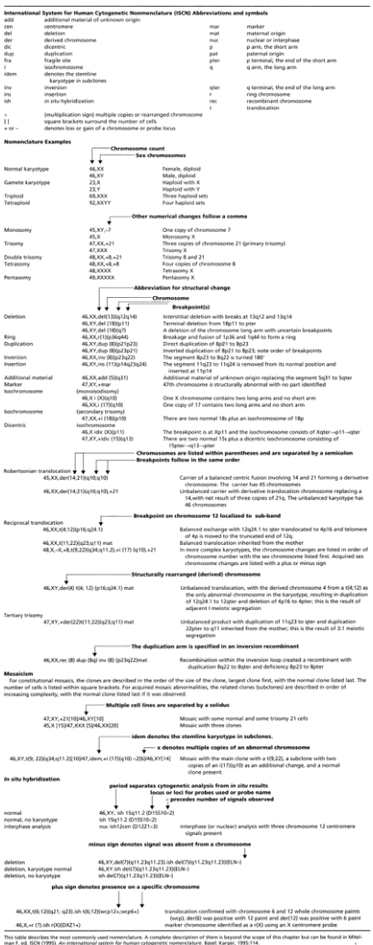
30
Clinical Cytogenetics

Daniel L. Van Dyke

Anne Wiktor

In Chapter 29, normal chromosome structure and the normal karyotype were described. This chapter addresses cytogenetic pathology associated with birth defects, mental retardation, infertility, and pregnancy loss. Approximately 0.6% of newborns have a chromosome abnormality, constituting a major cause of lethal or debilitating birth defects and mental retardation. Among couples who experience two or more miscarriages, 5% have a balanced chromosome rearrangement. Most of these are familial, and some also confer an increased risk of having a liveborn child with an abnormal karyotype. Approximately 50% of first-trimester miscarriages have a chromosome abnormality, making this by far the most common cause of pregnancy loss, and perhaps 100% of leukemia and cancer acquire an abnormal karyotype. We begin with a discussion of the variety of numerical and structural chromosome abnormalities and continue with examples and descriptions of cytogenetic changes associated with birth defects, mental retardation, and male and female infertility. The main indications for chromosomal study are summarized, followed by techniques for chromosome preparation and guidelines for analysis. The recommendations of the International System for Human Cytogenetic Nomenclature (1) are used to describe karyotypes throughout this section (Table 30.1).

TABLE 30.1. STANDARD CYTOGENETIC NOMENCLATURE[a]



- CHANGES IN CHROMOSOME NUMBER AND STRUCTURE
- AUTOSOMAL CHROMOSOME SYNDROMES
- SEX CHROMOSOME SYNDROMES
- INDICATIONS FOR CHROMOSOME STUDIES (SUBPOPULATIONS)
- METHODS

CHANGES IN CHROMOSOME NUMBER AND STRUCTURE

Part of "30 - Clinical Cytogenetics"

Changes in Chromosome Sets (Ploidy)

The normal karyotype contains two copies of every autosome plus two sex chromosomes. This is diploid, or 2n. A sperm or ovum is haploid, or 1n, and normally contains one copy of each autosome plus an X or a Y chromosome.

Triploidy (three haploid sets, or 3n) results when two sperm fertilize one ovum (Fig. 30.1A). Approximately 7% of first-trimester miscarriages are triploid. Rare triploid stillbirths and liveborn infants have occurred. The karyotypes 69,XXY and 69,XXX are most common, whereas 69,XYY is far less common and 69,YYY is never seen. Triploidy can also occur from fertilization of a diploid ovum, but most of these conceptions are unable to survive long enough to be observed among miscarriage specimens. This implies that two maternal contributions are more detrimental than are two paternal contributions, probably because of imprinting difference between the maternal and paternal haploid sets, as described in chapter 29. Triploidy appears to occur sporadically among miscarriages, and the risk of recurrence in subsequent pregnancies is not increased. In the *in vitro* fertilization laboratory, a higher concentration of sperm in the fertilization culture can increase the rate of formation of triploid embryos.

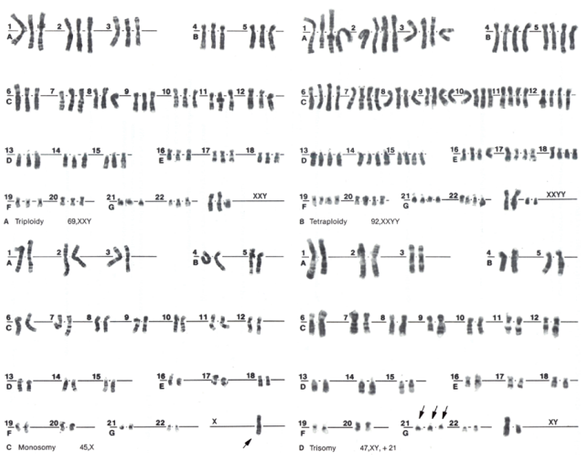


FIGURE 30.1. Numerical changes. An abnormal number of haploid sets (n) is found in triploidy (3n) **(a)** and tetraploidy (4n) **(B)**. Monosomy X **(C)** (*arrow*) is the most common monosomy (loss of a single chromosome, 2n-1), and trisomy 21 **(D)** (*arrows*) is the most common trisomy among liveborn infants (gain of a single chromosome, 2n+1).

Tetraploidy (4n) usually results from failure of cell division after a round of mitosis, and the karyotype is typically 92,XXXX or 92,XXYY (Fig. 30.1B). There is evidence from tumor studies that tetraploidy can also arise from fusion of two diploid cells. Approximately 2.5% of first-trimester miscarriages are tetraploid. A few patients with birth defects have been reported as having a mixture of diploid and tetraploid cells. However, because tetraploidy can also arise in cell culture, this is a difficult diagnosis to confirm, and the significance of such cells must be questioned. Tetraploidy appears to arise sporadically, and there is no increased risk of recurrence in subsequent pregnancies.

Changes in Chromosome Number in a Diploid Background

In the haploid, diploid, triploid, and tetraploid karyotypes, there are one, two, three, and four copies of each chromosome, respectively. The normal karyotype is euploid (Latin, *true number*). When an individual chromosome is lost or gained, the karyotype is aneuploid (Latin, *not true number*). The loss or gain of a single chromosome from a diploid background is monosomy or trisomy, respectively.

Down's syndrome (trisomy 21), identified by Lejeune in 1959, was the first clinical condition in which a gain of a single chromosome was reported (Fig. 30.1D). Trisomy 18 and 13 (Edward's and Patau's syndromes, respectively, after the men who described their chromosomal etiologies in 1960) are also classic trisomy syndromes associated with major birth defects, mental retardation, and a high risk of perinatal death. Trisomy X is mild by comparison. Most of the other autosomal trisomies are associated with fetal demise and are not commonly found among liveborn infants (Table 30.2). Trisomy 8 is a common abnormality in leukemia and trisomy 7 is common in many solid cancers.

TABLE 30.2. AUTOSOMAL TRISOMY AND ISOCHROMOSOME CONDITIONS THAT HAVE BEEN OBSERVED IN MISCARRIAGE SPECIMENS AND AMONG LIVEBORNS

Chromosome	% of Trisomic Miscarriages[a]	Liveborns[b]			
		Trisomy[b]	Trisomy Mosaic	Isochromosome	2° Trisomy (47, +isochromosome)
1	0				
2	5.6		V rare		
3	0.8		V rare		
4	2.5		V rare		
5	0.1		V rare		
6	0.3				
7	4.5	None?	V rare		
8	3.7	None?	1/10,000[c]		+i(8p) V rare
9	2.8	Rare	Rare	i(9q) V rare	+i(9p) Rare[c]
10	1.9		V rare		
11	0.2		V rare		
12	0.8		V rare		+i(12p) Uncommon[c]
13	5.7	1/5,000[c]	Uncommon[c]	i(13q) Rare	
14	4.2	V rare	V rare		
15	7.3	V rare	V rare		+dic(30)(q1) Uncommon[c]
16	32.4		V rare		
17	0.7		V rare		
18	5.1	1/6,000[c]	Uncommon[c]	i(18q) Rare	+i(18p) Uncommon[c]
19	0.3		None?		
20	2.7	None?	Rare[c]		
21	8.3	1/800[c]	Uncommon[c]	i(21q) Rare[c]	+i(21q) V rare
22	10.1	V rare	V rare	i(22q) V rare	+dic(22)(q11)

[a] Autosomal trisomy is identified in 25% to 30% of miscarriage specimens. The percentage of each chromosome among autosomal trisomies is shown. Data from ref. 2.

[b] More than 50 reported cases but less than 1 per 10,000 livebirths is considered uncommon, 10 to 50 reported cases is considered rare, and fewer than 10 reported cases is considered very rare. The designation “none?” means that this has been reported but the finding is controversial. A blank space indicates that the condition has not been observed.

[c] Syndrome is described in the text.

During meiosis I (see Chapter 29), the chromosome homologs segregate to opposite sides

of the meiotic cell. During meiosis II or mitosis, sister chromatids segregate to opposite sides of the cell. This is normal disjunction. If sister chromatids fail to segregate normally (primary nondisjunction), the daughter cells become aneuploid (Fig. 30.2). Most pentaX (49,XXXXX) or 49,XXXXY cases probably arise as a result of maternal nondisjunctional events during both meiosis I and II, with a normal paternal contribution of an X- or Y-bearing sperm.

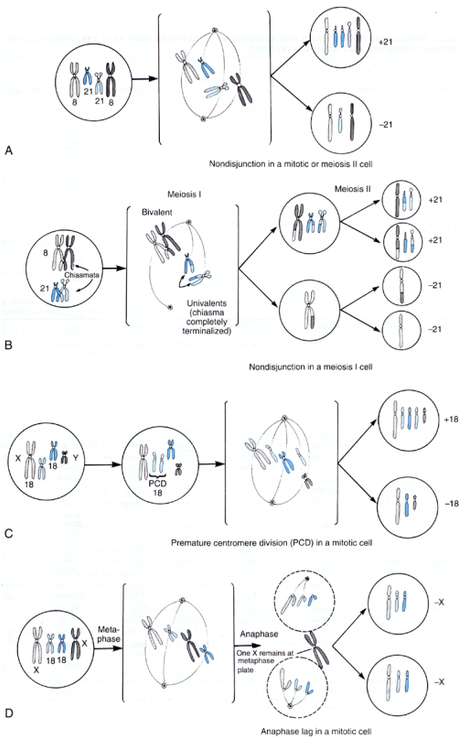


FIGURE 30.2. Origins of aneuploidy. Nondisjunction **(a)**, an important source of trisomy and monosomy, occurs during mitosis or meiosis II when sister chromatids (or when homologs in meiosis I, as shown in **B**) move together to the same daughter nucleus. A second source of aneuploidy **(C)** is premature division of the centromere, which permits the site chromatids to drift apart and segregate independently during anaphase to daughter nuclei. Monosomy can also arise through anaphase lag **(d)**, in which one chromatid is excluded from either daughter nucleus.

Because trisomy 21 (Down's syndrome) is the most common single identified cause of mental retardation, the biological basis of nondisjunction is of great interest. The parental origin and meiotic stage of the nondisjunctional error can often be identified from studies of chromosome 21 short arm variants and restriction fragment length polymorphism using DNA probes from 21q in Down's syndrome patients and their parents. From such studies, it has been established that more than 90% of the nondisjunctional events occur during maternal meiosis and the rest during paternal meiosis or an early embryonic mitosis. Most of these errors occur during meiosis I. The average maternal age for trisomy 21 children is approximately 30 years compared with the population average maternal age of approximately 25 years. Maternal age is also advanced for 47,XXX and 47,XXY and the other autosomal trisomies (Table 30.3). The biological basis for the maternal age effect is not completely understood, but the frequency of recombination (crossing over) appears to be lower in most cases, with zero or only one recombination event, and the recombinations appear to be mostly near the end of the chromosome. This results in a much weaker physical connection between the two chromosome 21 early in meiosis I and allows them to drift apart (Fig. 30.2B). In other cases of maternal age-associated nondisjunction, an opposite effect may be occurring: a greater number of recombination events than normal or a recombination event close to the centromere may lead to a stronger physical connection between the chromosomes (an entanglement perhaps) such that they cannot separate normally at the first meiotic division, and both chromosomes travel together to the same daughter cell (4).

TABLE 30.3. CHROMOSOME ABNORMALITY RATES PER THOUSAND IN LIVEBIRTHS BY 5-YEAR MATERNAL AGE INTERVALS*a*

Maternal Age	Down's Syndrome[b]	Edwards' Syndrome	Patau's Syndrome	47,XXY	47,XXX
15–19	0.6	0.06	0.03	0.4	0.4
20–24	0.7	0.07	0.04	0.4	0.4
25–29	0.9	0.09	0.05	0.4	0.4
30–34	1.4	0.14	0.07	0.6	0.4
35–39	4.2	0.42	0.21	1.1	0.7
40–44	14.2	1.42	˜0.60	2.6	1.9
45–49	˜50	?1.42	?0.6	6.6	3.5

[a] The statistics are from ref. 3.

[b] Average of studies of rates in New York and Sweden.

Many other risk factors for nondisjunction have been postulated, but none has been clearly demonstrated as important, including paternal age, decreasing sexual activity with advancing

age, ionizing radiation, maternal thyroid dysfunction, inbred population groups, chromosome 21 short arm variants, and rearrangements involving other chromosomes.

An individual who is trisomic, such as a female with trisomy X or 21, has an increased risk of having offspring with the same trisomy. The formation of a 24,X,+21 or 24,XX ovum in a trisomy 21 or trisomy X female, respectively, is called secondary nondisjunction. One would expect 50% of the conceptions of a 47,XXX mother to be either 47,XXX or 47,XXY and 50% of the conceptions of a 47,XX,+21 mother to have trisomy 21, as products of secondary nondisjunction.

The only monosomy observed in liveborn infants or miscarriages is 45,X. This is important in Ullrich-Turner syndrome (Fig. 30.1C) and is observed in approximately 9% of miscarriages. Monosomy 21 and 22 are rarely identified in miscarriages. Conceptions with autosomal monosomy appear unable to survive long enough to be recognized and are probably lost before the time of implantation. Conversely, monosomy 7 is a common change in acute myeloid leukemia and myelodysplasia, and monosomy 22 is an important abnormality in meningioma (see also Chapter 32 and Chapter 33). Although maternal age is a major risk factor in trisomy, it is not advanced in monosomy X.

Mosaicism

Nondisjunction or anaphase lag in mitotic cell division creates mosaicism, a mixture of cells with different karyotypes. Mosaics for trisomies 13, 18, and 21 and monosomy X are all well known and in general have a similar or less severe phenotype than the nonmosaics. Several other mosaic trisomies have been observed in liveborn infants (Table 30.2). Patients with trisomy 8 are probably all mosaic with both normal and trisomic cell lines. It has been suggested that cultured fibroblast cells from trisomy 8 mosaics typically have a higher proportion of +8 cells than do lymphocytes. Different cell types probably tolerate some chromosome abnormalities differently. Several tissue-limited or tissue-dependent mosaic chromosome abnormalities have been identified in patients who have mental retardation, asymmetric skeletal or other malformations, and a variegated, swirled, or linear streak pattern of skin pigmentation, similar to hypomelanosis of Ito (Table 30.4). Many of these subjects have a normal blood karyotype with mosaicism found only in cultured fibroblast cells. Although the true frequency of clinically important mosaicism is uncertain, trisomy 8 and diploid/triploid mosaicism are among the most important distinct mosaicism syndromes.

TABLE 30.4. TISSUE-LIMITED MOSAICISM*a*

Characteristics common to tissue-limited mosaicism
Mental retardation and minor malformations and
Joint contractures (arthrogryposis)
Diaphragmatic hernia (in Killian syndrome)
Pigmentary dysplasia (e.g., hypomelanosis of Ito)
Chromosome abnormalities frequently showing tissue-limited mosaicism
Diploid/triploid mosaicism
Warkany's syndrome
Pallister-Killian syndrome
Supernumerary marker
Incontinentia pigmenti, mental retardation
Turner's syndrome with mental retardation
46/69 mosaicism
Trisomy 8
Secondary trisomy 12p
Various 46/47,+mar or +ring chromosome
X; autosome translocation with Xp11 breakpoint
46, X/46,X,r(X) with loss of XIST function

[a] If tissue-limited mosaicism is identified, uniparental disomy may also be present and may further complicate the phenotype.

Mosaic autosomal monosomy is very rare (other than in cancer cells), and we are aware of only a very few cases of -21 or -22 mosaicism in liveborn infants, prenatal diagnostic specimens, or miscarriages. We have seen one amniotic fluid specimen with -21/diploid mosaicism. The parents elected to interrupt the pregnancy, so it is uncertain whether the fetus would have been viable. The source of mosaic monosomy/diploidy could be mitotic nondisjunction, premature centromere division, or anaphase lag (Fig. 30.2). Because lymphocyte chromosome preparations in most laboratories have approximately 5% of cells with 45 chromosomes, it is difficult to detect a low-frequency mosaic monosomy against this background level of aneuploidy. Cell cultures from some subjects with autosomal rings or dicentrics (see below) have a low frequency of monosomy owing to loss of the mitotically unstable chromosome, but the *in vivo* viability of such cells is unclear. Mosaic monosomy X, with either 46,XX or 46,XY or altered sex chromosome structure, is common among patients with Ullrich-Turner syndrome, gonadal dysgenesis, and ambiguous genitalia. Approximately one in 2,500 prenatal studies is 45,X/46,XY, usually resulting in a phenotypically normal male infant. The frequency of cells 45,X,-Y in males and 45,X,-X and 47,XXX in females increases with advancing age. In females, this age effect becomes evident by age 30, and in males by age 55.

Changes in Chromosome Structure

Changes in chromosome structure fall into three broad categories: deletion, duplication, and rearrangement. A structural change can involve a single base pair (bp) or millions of bases of DNA. The lower limit of resolution through the microscope, and therefore the limit for classic cytogenetic techniques, is approximately 2,500,000 bp [2,500 kilobase (kb)]. The average chromosome contains approximately 100,000 kb, much of it nontranscribed, and the average gene is approximately 40 kb in length. The average 2,500-kb segment is thought to contain approximately 60 genes.

Deletion

A deleted chromosome has lost part or all of one of its arms (Fig. 30.3A, Fig. 30.3B). A chromosome with a deletion of its centromere region (acentric) is usually lost within a few cell divisions. Most chromosome deletions appear through the microscope to be terminal; that is, the telomere (chromosome end) appears to have been lost. However, it is doubtful whether a chromosome with a terminal deletion is stable because the telomere appears to be an essential structure (discussed in more detail in Chapter 29). For example, deletions induced by ionizing radiation have “sticky ends” and frequently form unstable chromosome rearrangements, whereas deletions associated with birth defects are highly stable. Some R-banding studies have suggested that the terminal R band is not lost from deletions that appear by G banding to be terminal. Nevertheless, the ISCN (1) provides for interpretation of a deletion as terminal if it appears to be so through the microscope.

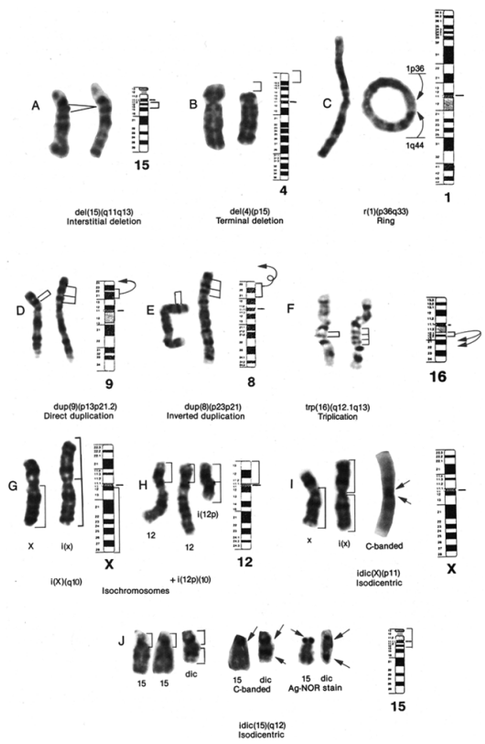


FIGURE 30.3. Structural changes. A chromosome deletion can be interstitial **(a)**, terminal **(B)**, or involving both chromosome ends to form a ring **(C)**. A tandem duplication can be direct **(d)** or inverted **(E)** with respect to the centromere. A tandem triplication is shown in **F**. (Courtesy of The Children's Hospital, Denver, CO.) A monocentric or dicentric isochromosome can replace a normal homolog **(G, I)** or represent an extra element in the karyotype **(H, J)**.

Deletions have been observed for many but not all regions of the karyotype. Several deletion syndromes (a syndrome is a constellation of features typical of patients with the defect) were characterized before the advent of chromosome banding because the structural change was obvious even on uniformly stained chromosome preparations. These include 4p- and 5p- (Wolf-Hirschhorn and cri-du-chat syndromes, respectively), 18p- and 18q- syndromes, and X chromosome deletions associated with gonadal dysgenesis. Some patients with the classic deletion syndromes have been identified only by using high-resolution cytogenetics, because the deletion can be quite small (the 4p- in Fig. 30.3B was not seen in metaphase spreads with shorter chromosomes). For genetic counseling, the parents of patients with deletions need to be karyotyped because 10% to 15% are derived from familial, balanced chromosome rearrangements. If the parents' karyotypes are normal, the risk of recurrence is very low. Deletions are also important in carcinogenesis, permitting the loss of expression of tumor suppressor genes (see Chapter 34).

Ring

Breakage in both arms of a chromosome and repair by fusion at the two breakpoints results in a circular chromosome, or ring (Fig. 30.3C). A ring can arise in any cell, and many patients with rings are mosaic with a normal cell line. Complicated structures that arise during replication can produce daughter cells with multiple rings and rings that have zero, two, or even more centromeres. This instability often generates cell lines with rings of different sizes and with monosomy for the chromosome involved. Rings of every chromosome have been observed, and a few have been sufficiently common to delineate syndromes. However, differences among patients in breakpoints of the rings and the frequent mitotic rearrangement of rings lead to wide variation in the phenotypes of patients who appear to have the same ring.

Although the risk of recurrence in future pregnancies is considered very low, chromosome studies of parents are necessary because some familial rings have been found. If the deletion at both ends of the chromosome is very small, the phenotypic effects can be minimal. Some individuals with rings have had a normal or nearly normal phenotype and have reproduced, but their offspring are at an increased risk for birth defects. For example, we identified a four-generation family with a ring chromosome 21 (5). One mildly mentally retarded child had a ring 21 with more material lost than the ring in her mother's cells. If an individual is mosaic with normal diploid cells, an error early in embryogenesis is likely, and the parents would be expected to have normal chromosomes, although reversion from a ring to a stable rod chromosome has been reported.

Duplication

A cytogenetically visible duplication almost invariably results in a miscarriage or a livebirth with birth defects and mental retardation. The exceptions primarily involve duplications of heterochromatic (C band–positive) segments and acrocentric short arm variants. Gene duplication has been important in evolution, allowing species to experiment with nonessential, duplicate genes. The hemoglobin and immunoglobulin gene families are classic examples of evolution after gene duplication.

A segment can be duplicated on the same chromosome or on a nonhomologous chromosome (intrachromosomal or interchromosomal duplication, respectively). An intrachromosomal duplication can be immediately adjacent to the normal segment

(tandem or contiguous, ABCDEEFG) or, less frequently, separated by nonduplicated material (noncontiguous, ABCD*EF*G). When the duplicated segment has a normal orientation with respect to the centromere, it is a direct duplication (ABCcenD *EFEF*G), otherwise it is an inverted duplication (ABCcenD *EFFE*G) (Fig. 30.3D, Fig. 30.3E). In practice, it is not always possible to determine the orientation of a duplicated segment or whether the duplication is contiguous. The source of a small duplicated segment is often somewhat speculative, and biochemical, molecular, or molecular cytogenetic studies are often useful to confirm a cytogenetic impression.

A direct duplication is the most common intrachromosomal duplication, probably because it can arise from a simple unequal exchange between sister chromatids in a mitotic or meiotic cell, whereas an inverted tandem duplication requires three chromosome breaks (Fig. 30.4A, Fig. 30.4B). A similar unequal exchange mechanism probably accounts for many deletions as well. The mean parental age is significantly advanced in patients with inverted duplications, suggesting a meiotic origin, but it is normal in patients with direct duplications. Most tandem duplications are new mutations with a low risk of recurrence. Chromosome studies of the parents are necessary in all duplication cases to look for a familial chromosome rearrangement and to provide accurate genetic counseling. For an interchromosomal duplication, where the extra material is derived from a different chromosome, the odds that a parent carries a balanced translocation are significantly greater than the 10% to 15% for a deletion. Duplications and deletions associated with familial chromosome rearrangements are discussed further below.

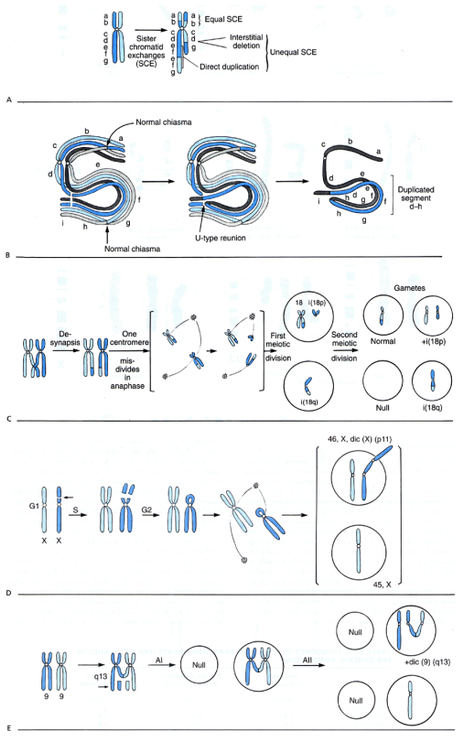


FIGURE 30.4. Origins of duplications and isochromosomes. A direct duplication **(a)** can arise in a mitotic or meiotic cell by unequal sister chromatid exchange or recombination. The reciprocal product is a deletion. An inverted duplication **(B)** probably results from a three-break rearrangement between homologous or sister chromatids during meiosis I. The repair of the three breaks would include a U-type reunion. It is shown here arising in meiosis I, with one break in the proximal long arm and two breaks in the distal long arm. The repair process results in duplication of the intervening segment. Several mechanisms of isochromosome formation have been proposed. In meiosis I, desynapsis of the two chromosome 18s can lead to misdivision of a centromere in anaphase I to form monocentric isochromosomes of each chromosome arm **(C)**, and gametes could have either an i(18q) replacing a normal 18, or a supernumerary i(18p). As shown in this example, both 46,i(18q) and 47,+i(18p) have been observed. Formation of a dicentric isochromosome X can result from repair of a chromosome break with a U-type reunion. In a mitotic cell (as shown in **D**), this would produce an individual with 45,X/46,X,idic(X) mosaicism and in some cases a 46,XX line as well. The acentric short arm remnants would be lost. In contrast to the mitotic origin of the dicentric that replaces a normal homolog (as in **D**), during meiosis, a supernumerary dicentric isochromosome is more likely to arise. In **E**, a chromatid break in both chromosome 9s and U-type reunion are shown. In anaphase I, the homologous chromosomes are attached by the rearrangement and travel together to the same daughter cell. One might expect an anaphase bridge to occur because the two centromeres typically are drawn to opposite poles, but “centromere cooperation” appears to occur in dicentric chromosomes in which the centromeres are close together. The final result is one gamete with a normal chromosome 9 plus the dicentric, one normal gamete, and two null gametes. The acentric fragments are lost.

Isochromosome

In an isochromosome, the arms are mirror images (Fig. 30.3G,Fig. 30.3H,Fig. 30.3I and Fig. 30.3J and Fig. 30.4C,Fig. 30.4D and Fig. 30.4E). The isochromosome can replace a normal chromosome (monoisodisomy) or be an extra chromosome

in an otherwise normal karyotype (secondary trisomy). It can be monocentric with two long or two short arms of the chromosome, or dicentric with the material between the centromeres continuing the mirror image: ABCDcenDCBA or ABCDcenEF::FEcenDCBA, respectively. Isochromosome Xq is associated with gonadal dysgenesis and features of Ullrich-Turner syndrome. Isochromosome Yp and Yq are associated with male infertility and with abnormal sex differentiation, and several autosomal isochromosome syndromes have been identified (Table 30.2). Even though the origin of secondary trisomies is often a meiotic error, mosaicism is frequent because dicentric chromosomes are mitotically unstable, and the karyotype reverts to normal in cells that lose the isochromosome (compensating nondisjunction). Parental karyotypes are usually normal but should be studied because some rearrangements, mainly paracentric inversions (see below), can give rise to an apparent isochromosome. Parental mosaics have also been described. Isochromosomes are also important in the cytogenetic evolution of leukemia and solid cancer.

Rearrangement

The three main categories of balanced chromosome rearrangements are robertsonian translocations, reciprocal translocations, and inversions. One in 175 to 250 phenotypically normal subjects carries a balanced rearrangement (Table 30.5). The identification of balanced rearrangements is important because carriers have increased risks for having miscarriages and malformed liveborn infants, and some male carriers have reduced fertility. A rearrangement can be carried by many family members, conferring risk to many relatives. Furthermore, new mutation carriers of apparently balanced reciprocal translocations and inversions (but not centric fusions) are overrepresented among the mentally retarded. There are several possible explanations for this, including loss of gene function because of breakpoints within genes, undetected deletion at one or both chromosomal breakpoints, or position effects (changes in gene expression because of different location in the genome, such as movement away from or adjacent to controlling elements or heterochromatin).

TABLE 30.5. FREQUENCY OF BALANCED REARRANGEMENTS (PER THOUSAND) IN THE GENERAL POPULATION*a*

	Newborn (Unbanded)[a]	Prenatal + Newborn (Banded, to 1982)	Prenatal (1982–1990)
Sample size	56,952	12,923	11,500
Translocations			
Robertsonian	0.9	0.9	1.4
Reciprocal	0.9	1.5	1.8
Inversions			
Pericentric	0.1	0.7	1.7
Paracentric	0	0.3	0.9
Total	1.9	3.4	5.8

[a] Estimates from newborn surveys using routine stained (unbanded) chromosome preparations were accurate for robertsonian translocations because the chromosome count is 45. Reciprocal translocations and pericentric inversions were identified only if centro mere positions changed. No paracentric inversions were found. Estimates based on prenatal genetic studies using banded chromosomes vary somewhat, the differences at least partly owing to improvements in the quality of banding over time. The statistics are from various sources summarized in refs. 6,7 and 8 and Van Dyke et al., unpublished data.

Robertsonsian Translocation

A centric fusion or robertsonian translocation arises from the fusion of two acrocentric chromosomes (chromosomes 13, 14, 15, 21, and 22) (Fig. 30.5). During the resting phase of the cell cycle multiple nucleoli form, each involving the nucleolar organizing region (NOR) of one acrocentric chromosome. With time, the number of nucleoli becomes fewer and their size larger as they coalesce. In a metaphase chromosome preparation, remnants of nucleolar activity are evidenced by acrocentric associations (Fig 30.5B). Acrocentric association no doubt plays a role in the formation of centric fusions. For reasons that are not understood, the frequency of different robertsonian translocations is nonrandom (Table 30.6).

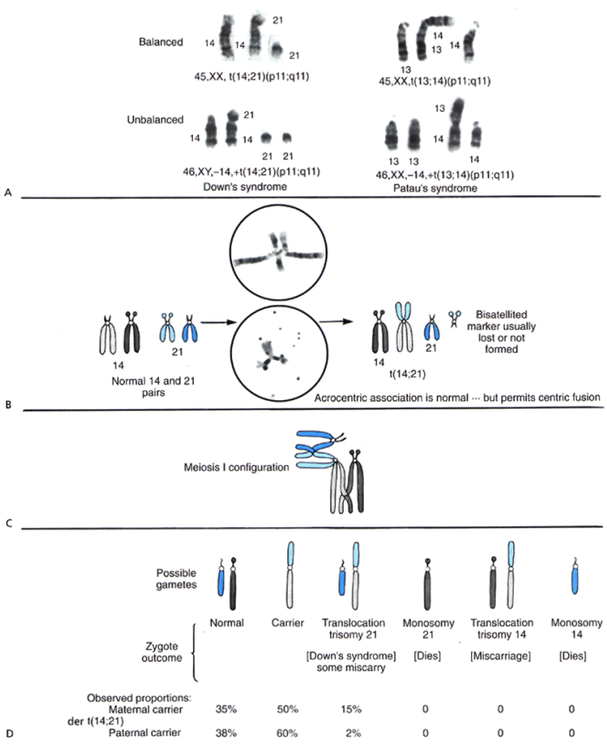


FIGURE 30.5. Robertsonian translocation (centric fusion). The most important robertsonian translocations are der(14;21) and der(13;14). Partial karyotypes of a balanced and an unbalanced set are shown in **A**. Acrocentric association **(B)** is normal chromosome behavior. It is a remnant of nucleolar activity, as shown by the common stream of silver staining (AgNOR staining). This close association probably accounts for the relatively high frequency of centric fusions compared with other balanced rearrangements. The reciprocal product of a centric fusion, a bisatellited marker chromosome, is usually not found. A meiosis I configuration **(C)** formed in a carrier of a der(14;21) is shown, along with the six possible gametic products **(d)**, of which only three are ever observed. Frequency statistics are based on prenatal diagnosis results in carriers (9).

TABLE 30.6. DISTRIBUTION OF EACH TYPE OF ROBERTSONIAN TRANSLOCATION IN THREE POPULATION SAMPLES*a*

Translocation	Unbiased Surveys (%)	Multiple Miscarriages (%)	Unbalanced Translocations (%)
13;13	2.3	3.6	3.3
13;14	71.0	59.6	12.8
13;21	0.9	0.4	2.6
14;21	9.3	8.0	40.7
15;21	0.9	1.2	3.8
21;21	2.3	0.4	28.4
21;22	0.5	0.8	2.8
All others[b]	12.6	26.0	5.5

[a] The statistics summarize the distribution of robertsonian translocations among subjects karyotyped in newborn surveys and routine screening programs such as prenatal diagnostic studies, couples who experienced multiple miscarriages or reduced fertility, and probands with malformations and an unbalanced robertsonian translocation (approximately two thirds of whom had Down's syndrome). The unbiased surveys show that der(13;14) is by far the most common robertsonian translocation. The pattern for the multiple miscarriage and infertility group is similar to that of the unbiased surveys. The der(14;21) and der(21;21) are relatively more important clinically than their frequencies in the general population would suggest. Together they comprise only 12% of robertsonian translocations, but because of their association with Down's syndrome, they account for two thirds of those with unbalanced translocations. Statistics are adapted from ref. 41, pooled with data from our laboratory.

[b] All other robertsonian translocations include der(14;14), der(15;15), der(22;22), der(13;15), der(13;22), der(14;15), der(14;22), and der(15;22).

Individuals who carry a *de novo* balanced robertsonian translocation have no increased risk for having malformations or mental retardation; these are not seen in increased frequency among patients with mental retardation or birth defects. However, rare patients with a *de novo* centric fusion involving chromosome 15 have been described as having Prader-Willi syndrome (PWS) or Angelman syndrome (AS) (described later).

Some patients with trisomy 21 (Down's syndrome) and trisomy 13 (Patau's syndrome) have unbalanced forms of robertsonian translocations (Fig. 30.5). This is the primary reason for studying the karyotype of a Down's syndrome patient because the diagnosis can usually be made clinically. The risk that a carrier of a balanced translocation will have chromosomally unbalanced offspring depends on the translocation involved, the gender of the parent, and the likelihood of survival of the abnormal conceptus (see the sections on trisomies 13 and 21, below). The expected 1:1 ratio of normal to balanced karyotypes among offspring is not found, as more than half of the normal offspring are translocation carriers. This observation is not yet understood.

The frequency of robertsonian translocations in the general population is approximately one in 900, and they are more common among couples with multiple miscarriages and males with infertility. In studies of consecutive newborns or prenatal diagnosis specimens, the sex ratio of centric fusion carriers is approximately equal, whereas among the parents of carriers or unbalanced progeny, many more mothers are carriers. The difference

is probably related to the observed reduced fertility in some male carriers of centric fusions.

Reciprocal Translocation

A reciprocal translocation involves the exchange of chromosome segments, usually involving breakpoints on each of two nonhomologous chromosomes. One in 500 to 650 liveborn infants carries a balanced reciprocal translocation. Unbalanced translocations and complex rearrangements involving more than two breakpoints are less common. One translocation, t(11;22) (q23;q11.2), is seen in approximately 1 in 5,000 normal subjects. Other than recurring translocations in cancer and leukemia, this is the only important exception to the practically random distribution of chromosomal breakpoints among reciprocal translocations. The 22q "hot spot" for chromosome breakage includes repetitive DNA sequences, a low copy repeat cluster, LCR22, that may be a predisposing factor for the t(11;22) and for the 22q11 deletion syndrome (described later).

A major factor in the reproductive risks to a translocation carrier is the viability of the possible unbalanced meiotic products (Fig. 30.6). [Several chapters in Daniel's book (10) are devoted to this issue.] In general, risks of having a child with an unbalanced form of a reciprocal translocation range from nearly zero to 25% to 30%; the risk of miscarriage averages approximately 25%; and some males have reduced fertility. Unbalanced products with small duplications or deletions are more likely to be viable, and a gain of 1% or loss of 0.5% of the haploid complement is usually viable, whereas a gain of 4% or loss of 2.5% is likely to be lethal. The short arms of chromosomes 17 and 20 each constitute approximately 1% of the haploid complement, so either arm is a handy point of reference for predicting the viability of a rearrangement. However, viability is determined more by the genes that reside on a chromosomal segment than by its length, so there is wide variation, with some larger chromosomal segments viable as duplication or deletion and a few smaller segments never observed in a chromosomal imbalance.

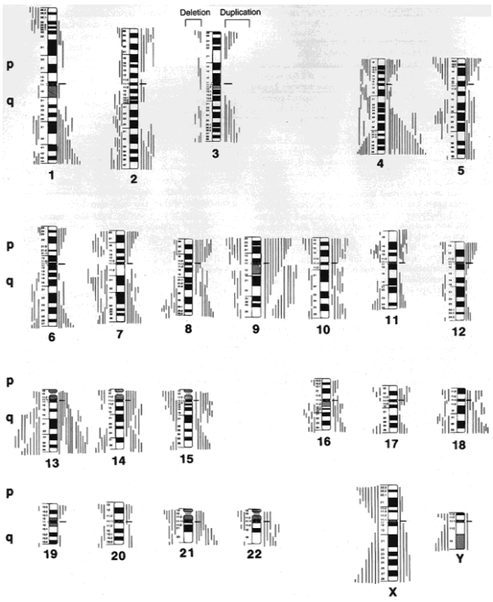


FIGURE 30.6. Chromosome deletions and duplications. The idiogram depicts the deletions (at left of each chromosome) and duplications (at right) that have been observed among liveborn infants. Mosaic cases with a normal cell line are excluded. The length of each line represents the chromosome segment that was lost or duplicated. The relative frequency of each chromosomal imbalance is not represented here. In general, smaller amounts of imbalance are more common, but there are many exceptions. For example, gain and loss of chromosome 13 and chromosome 18 segments are among the most common, whereas even the smallest changes to chromosome 19 are very rare.

Of the unbalanced meiotic products, adjacent I segregants are the most commonly observed (Fig. 30.7). Adjacent II segregants usually have greater imbalances and are rarely seen. If one of the translocation chromosomes is small, then a 3:1 segregation product may result in tertiary trisomy. For the t(11;22) mentioned previously, the risk is approximately 5% for the viable tertiary trisomy, and a maternal age effect may be superimposed on this risk, analogous to the maternal age effect in primary trisomy.

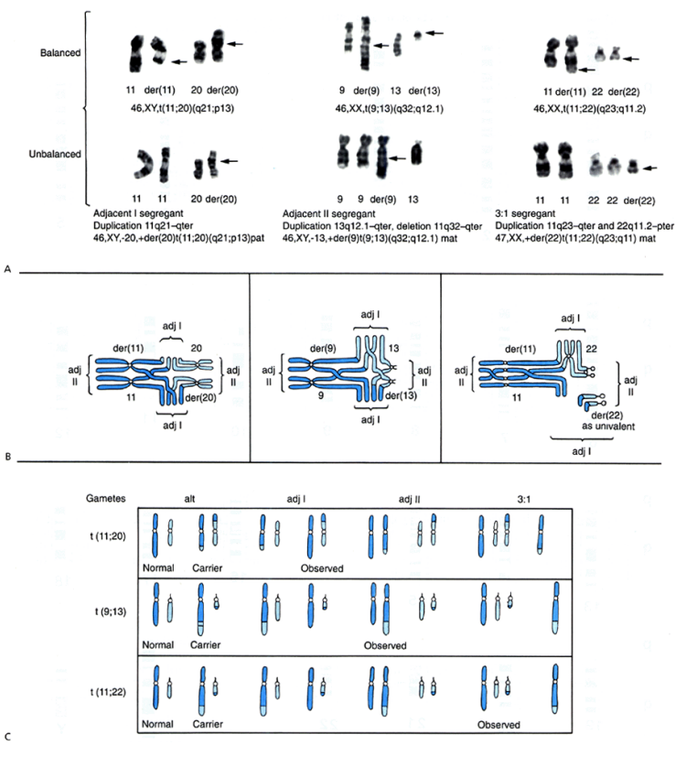


FIGURE 30.7. Segregation in reciprocal translocations. In **A**, three different translocations are shown in their balanced form and in an unbalanced form. In **B**, a meiosis I prophase quadrivalent is depicted for each translocation. The theoretically possible meiosis I segregants include alternate (alternate chromosomes segregate together), adjacent I (nonhomologous centromeres segregate together), adjacent II (homologous centromeres segregate together), and 3:1 segregation (more common when one element is small). Only alternate segregation leads to chromosomally balanced gametes. The meiotic products are shown in **C**, and for the specific translocation used as an example in **A** and **B**, the viable gametes that have been observed are noted.

Insertion and Shift

An insertion is a three-break translocation resulting from removal of a chromosomal segment (two breakpoints) to another site in the karyotype (third breakpoint). A shift is an insertion within the same chromosome (Fig. 30.8). Although insertions are uncommon, they are clinically important because the risk of a chromosomally unbalanced conception in a carrier can be

50%. As with a reciprocal translocation, the risk to have a viable livebirth with a deletion or duplication depends greatly on the size of the inserted segment.

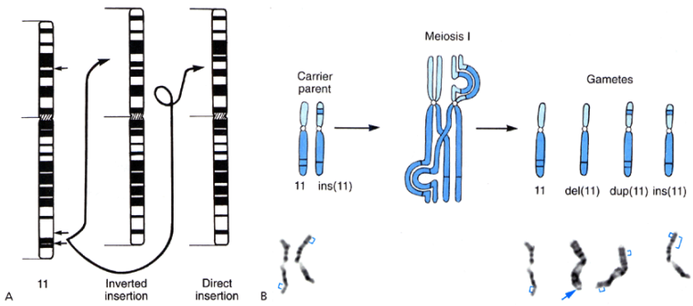


FIGURE 30.8. Recombination in insertion (shift). An intrachromosomal shift requires three breakpoints, as shown in **A** for a chromosome 11 shift. The insertion can be inverted with respect to the original centromeric orientation, or direct. In **B**, the chromosome 11s of the carrier parent are shown diagrammatically and as a partial karyotype. Because the inserted segment is small, it is appropriate to show the meiosis I prophase configuration without synapsis of the insertion segment. Nonrecombinant gametes perpetuate the parental karyotypes, whereas recombination will produce deletion and duplication gametes. Both were observed in siblings with mental retardation and mild malformations. (From Forsythe MG, Walker H, Weiss L, et al. Duplication and deletion 11q23–q24 recombinants in two offspring of an intrachromosomal insertion ("shift") carrier. *Henry Ford Hosp Med J* 1988;36:183–186, with permission.)

Inversion

An inversion is a two-break rearrangement in which a chromosome segment is turned 180° (12) (Fig. 30.9). A paracentric inversion involves only one arm. A pericentric inversion includes the centromere and material from both arms. Balanced pericentric and paracentric inversions are more difficult to identify than translocations, so estimates of their frequency in the general population have increased as the quality of chromosome preparations has improved (Table 30.5). Inversions have been called crossover suppressors because the meiosis crossover products are often nonviable recombinants. Meiotic pairing is expected to create an inversion loop, but an inverted or noninverted segment that is small may remain unpaired during meiosis. Recombination within the inversion loop creates unbalanced products with duplication of material from one end of the chromosome and loss of material from the other end; the inverted segment is neither duplicated nor deleted (15). For a paracentric inversion one recombinant product is dicentric and one acentric.

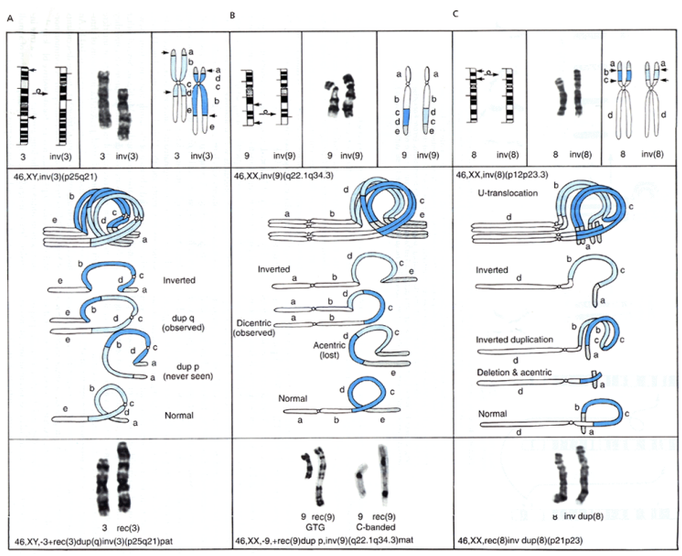


FIGURE 30.9. Recombination in inversion. Three examples are provided to show the results of recombination in pericentric and paracentric inversions. A partial karyotype of the carrier parent and affected offspring is provided at top and bottom, respectively, with an illustration of the meiosis I prophase configuration and recombinant and nonrecombinant products. A pericentric inversion 3 **(a)** and a paracentric inversion 9 **(B)** are shown. A two-break exchange within the paracentric inversion loop **(C)** during meiosis I in a paracentric inversion 8 carrier mother resulted in a child with an inverted duplication of 8p. The inv(9) family was reported by Worsham et al. (13) and the inv(8) family by Feldman et al. (14).

Among pericentric inversions, the risk for producing a malformed livebirth or a recognized miscarriage with a recombinant chromosome varies greatly. In general, a small inversion has a lower likelihood of meiotic exchange between the inverted region and the normal homolog, and therefore a smaller proportion of gametes carry an unbalanced recombinant. Moreover, such recombinants have a greater imbalance and the consequence is lower viability. Conversely, a larger inversion has a greater chance of recombination within the inversion loop, and the recombinants have a smaller imbalance, increasing the chance of producing a miscarriage or malformed livebirth. One can apply some of the same empirical risk data from translocations to pericentric inversions: that viability is dramatically reduced if the imbalance is greater than 0.5% deletion or 1% duplication.

The frequency of balanced pericentric inversions in the general population is approximately one in 600. Pericentric inversions are found more frequently among couples with multiple miscarriages, and probands with unbalanced recombinants are found among patients with birth defects and occasionally in miscarriage specimens. A pericentric inv(3)(p25q21) is responsible for at least 30 patients with a recombinant chromosome having duplication of 3q21–qter and deficiency 3p25–pter (Fig. 30.9A). This inv(3) may have originated in France and been carried by descendants to Canada and the United States, including Detroit, where we identified a branch of the family that came from Newfoundland to Ontario and southeast Michigan. Several other pericentric inversions are considered normal variants because they are not known to create unbalanced recombinants or interfere with fertility. These include small pericentric inversions of chromosome 2 and a metacentric Y chromosome inv(Y)(p11q11). An inv(2)(p12q14) appears to be more frequent among descendants of Sephardic Jews from North Africa and Spain. A common normal variant, inv(9)(p11q12), has a C band–positive block in the short arm in addition to the long arm but may not be a true inversion. It may instead have originated by amplification of a small segment of repetitive DNA in the proximal short arm.

The frequency of balanced paracentric inversions in the general population is approximately one in 1,000, based on our experience from prenatal genetic studies. Many paracentric inversions are not detected in routine studies because the altered pattern of chromosome banding can be quite subtle. Most have been identified by chance (mainly prenatal genetic studies) and only a very few from unbalanced recombinant offspring or *de novo* mutations in patients with birth defects or mental retardation. For a paracentric inversion carrier, the risk of producing a viable conception carrying a recombinant is low because, in general, the acentric and dicentric recombinant products are lost. Nevertheless, our laboratory has identified a girl with a stable dicentric recombinant chromosome whose mother carries a paracentric inv(9)(q22.1q34.3), and a girl with an unusual monocentric recombinant whose mother carries an inv(8)(p21p23) (Fig. 30.9B, Fig. 30.9C). Because only one other dicentric and five other monocentric recombinants from paracentric inversions have been reported in the literature, it is clear that viable recombinants of paracentric inversions are rare. Conversely, there may be significant underreporting in the genetics literature because we identified two of the reported cases. Although the risks for having offspring with unbalanced recombinant chromosomes are probably extremely low for most carriers of paracentric inversions, one must recognize that stable dicentric and monocentric chromosome imbalances can occur.

Uniparental Disomy

Genetic imprinting and uniparental disomy (UPD) were briefly described in Chapter 29. The normal expectation of human reproduction is that one inherits one of each chromosome from each parent. This is termed biparental disomy. However, some individuals have both copies of one chromosome from the same parent, and no contribution from the other parent (uniparental disomy). UPD can result in an abnormal phenotype in one of two principle ways.

First, if the individual inherits both copies of the same chromosome (isodisomy) from a parent and that chromosome carries a recessive mutation, then the mutation will be expressed. The first evidence of uniparental disomy in humans was found in a girl with cystic fibrosis (CF) who was homozygous for a mutation that her mother carried; her father was not a CF carrier. Additional molecular genetic studies proved that the girl had maternal isodisomy 7. Similar cases have since been found for other recessive disorders, and it has been postulated that 1% to 2% of recessive disease is owing to UPD (16).

Genetic imprinting (see Chapter 29) underlies the second class of an abnormal phenotype associated with UPD. For chromosomes that contain imprinted genes, biparental inheritance is essential to normal development. The two microdeletion syndromes, PWS and AS (see below), result from deletion of the

TABLE 30.7. CONTIGUOUS GENE SYNDROMES*a*

Autosomal	
Obesity, hyperphagia, MR	1p36
Vander Woude syndrome, MR	1q32–q41
Piebald trait, MR	2q34–q36
Albright's hereditary osteodystrophylike with obesity	2q37
Rieger's syndrome, MR	4q25–q27
Adenomatous polyposis, MR	5q21
Craniosynostosis, MR	7p21.2–p21.3
Greig's cephalopolysyndactyly, MR	7p13
Russell-Silver syndrome	upd7mat
Williams' syndrome	7q11.23
Spherocytosis I, MR	8p11.2–p21.1
Langer-Giedeon syndrome*	8q24.11–24.13
Trichorhinophalangeal, type I	8q24.12
DiGeorge syndrome	10p13
Beckwith-Wiedemann syndrome*	dup 11p15.5
Wilms' tumor, aniridia (WAGR)*	11p13
Retinoblastoma, MR*	13q14
Angelman, syndrome	15q11–q12 mat; upd15pat
Prader-Willi syndrome*	15q11–q12 pat; upd15mat
Rubinstein-Taybi syndrome	16p13.3
Miller-Dieker syndrome*	17p13.3
Smith-Magenis syndrome	17p11
Alagille's syndrome (arteriohepatic dysplasia), MR	20p11.23–p12.1
DiGeorge syndrome, MR	22q11.2
Velocardiofacial syndrome	22q11.2
X chromosomal	
Ichthyosis, chondrodysplasia punctata, and Kallmann's syndrome, MR	Xp22.3
Microphthalmia, iridoschisis, goiter, labium synechia, and craniotabes	Xp22
Choroideremia, deafness, MR	Xp21.1–21.3
Duchenne's and Becker's dystrophy, MR	Xp21
DMD, CGD, McLeod phenotype, retinitis, MR	Xp21
DMD, glycerol kinase deficiency, Aland's eye disease, MR	Xp21
Glycerol kinase deficiency, adrenal hypoplasia, hypogonadotropic hypogonadism, MR	Xp11.2–p21

[a] Asterisks indicate “classic” contiguous gene syndromes. All the contiguous gene syndromes are deletions except Beckwith-Wiedemann and Charcot-Marie-Tooth syndromes.

[b]MR, mental retardation; DMD, Duchenne's muscular dystrophy; CGD, Chronic granulomatous disease.

same segment of chromosome 15: deletion of the maternally inherited segment results in AS and deletion of the paternally inherited segment results in PWS. Likewise, paternal UPD results in PWS because the maternally imprinted chromosome 15 is absent, and conversely, maternal UPD results in AS. The girl with CF, mentioned earlier, also had small stature. Other children with UPD7, but without CF, also have short stature. This imprinting effect UPD7 is described in more detail later (see section on Russell-Silver syndrome). The origin and phenotypic effects of UPD have been the subjects of active study (cf. 17,18).

As with primary trisomy, there is also a maternal age effect in UPD, particularly maternal UPD (Fig. 30.10). A trisomic embryo can become mosaic by mitotic loss of one of the three chromosomes, thus reverting to the normal disomic state (compensating nondisjunction or trisomy rescue). If the disomic cell population has a proliferate advantage over the trisomic population, it can come to predominate in both the placenta and, more important, in the embryo proper. There is one chance in three that trisomy rescue will result in UPD. Because trisomy rescue also implies chromosomal mosaicism, an abnormal phenotype cannot be ascribed solely to UPD. Abnormalities can result even if the trisomy is limited to the placenta. For example, mosaic trisomy 16 confined to the placenta, with or without UPD, has been associated with intrauterine growth retardation (19).

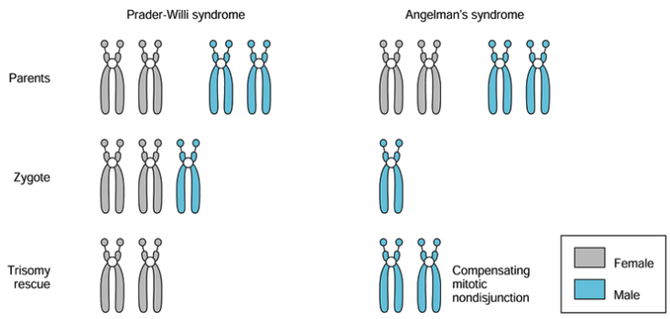


FIGURE 30.10. Illustration of common pathways to uniparental disomy. Chromosomes shaded light grey are maternally derived and dark grey are paternally derived. Paternal uniparental disomy, on the left, can be derived from a maternal meiotic nondisjunction leading to monosomy in the zygote. This is followed early in embryogenesis by a compensating mitotic nondisjunction of the paternally derived chromosome. Maternal uniparental disomy, on the right, can be derived from a maternal meiotic nondisjunction leading to trisomy in the zygote. This is followed early in embryogenesis by a compensating mitotic nondisjunction. The type of disomy that results depends on which chromosome is lost. If one of the maternal chromosomes is lost, the diploid cell population would exhibit normal biparental disomy, but if the paternal chromosome is lost as shown, the diploid cell population exhibits uniparental maternal disomy. (From Ledbetter and Greenberg, personal communication.)

AUTOSOMAL CHROMOSOME SYNDROMES

Part of "30 - Clinical Cytogenetics"

The diversity of autosomal chromosome imbalance is immense, so it follows that the variety of phenotypes associated with autosomal imbalance is great. Only a minority of chromosome syndromes can be identified by physical examination alone. Rather than compare and contrast all chromosome syndromes in detail, it is appropriate to examine the features common to many conditions. The two features found in almost all chromosome imbalances are growth retardation and mental retardation. Congenital heart disease, microcephaly, dysmorphic facial features (including hypo- or hypertelorism, broad nasal bridge, micrognathia, and low-set ears, and, in males, cryptorchidism) are also common and reflect the complexity of embryogenesis of the brain, face, urogenital tract, and heart and the sensitivity of these processes to perturbations of genetic balance. In general, any individual with mental retardation and more than two minor or major congenital anomalies is a candidate for chromosome studies unless a clearly defined nonchromosomal cause is evident.

As stated earlier, gain of a chromosome segment is tolerated better by the developing organism than is the loss of the same segment. In addition to trisomy 13, 18, and 21, several other autosomal trisomies have been reported in liveborn infants (Table 30.2). Likewise, duplication of whole arms in isochromosomes has been observed, as well as triplication through secondary trisomy. The map of reported duplications and deletions in liveborn infants (Fig. 30.6) shows visually that the amount of autosomal duplication tolerated is usually greater than that of deletion. It also shows the variety of potential chromosomal imbalance; virtually any segment can be duplicated or deleted.

Many chromosomal syndromes have been described in the *Clinical Atlas of Human Chromosomes* (20), *Smith's Recognizable Patterns of Human Malformation* (21) and the Online Mendelian Inheritance in Man (OMIM) (http://www3.ncbi.nlm.nih.gov/Omim/). In this section, we present brief descriptions of some well-known chromosomal syndromes and three categories of new syndromes: those associated with a deletion or duplication involving loss or gain of a small number of genes, those with tissue-limited mosaicism, and those related to genetic imprinting. Although each syndrome is defined by a pattern of defects, not all patients exhibit the entire constellation; rather, each defect is more common in that condition, and most of the patients will have some of the features.

The contiguous gene syndromes (Table 30.7) have several features in common (22). They result from duplications or deletions of small segments of chromosome material, containing only a few genes that are functionally unrelated, but by chance are closely linked on the chromosome. The phenotype tends to be variable because of different chromosomal breakpoints, making delineation of a clear syndrome sometimes more challenging. The occurrence of a contiguous gene syndrome is usually sporadic, but clusters within families are known. The familial clusters are owing to balanced chromosome rearrangements that carry a risk for producing unbalanced offspring with the syndrome. Before their chromosomal basis was identified, many of these syndromes were thought to represent autosomal recessive conditions or new dominant mutations. The phenotypes of some microdeletion or microduplication syndromes may involve loss or duplication of just one major gene, such as elastin in Williams syndrome (23), and there are many individual malformations that are significantly associated with loss of specific chromosomal segments (e.g., anal atresia and loss of 13q22 to 13q34 and aortic stenosis and loss of 11q23 to 11q24) (24).

For the study of contiguous gene syndromes, high-resolution cytogenetic methods or molecular cytogenetic techniques using specific DNA probes are essential. The mid to early metaphase analysis that is the accepted standard in most cytogenetics laboratories is not sufficient for reliable identification of these syndromes. This underscores the importance of good communication between the laboratory and the referring physicians. The smallest karyotype changes can appear normal, even when using high-resolution banding methods, because the imbalance is below the resolution of the light microscope. With advances in molecular cytogenetic technology, DNA probes for several contiguous gene syndromes are commercially available. For some syndromes, DNA probe technology (discussed later) has become the standard of care as an adjunct to karyotype analysis because it provides sensitive and specific assays for the detection of the smallest chromosomal deletions and duplications.

Wolf-Hirschhorn Syndrome (Wolf's Syndrome) (Deletion 4p)

Deletion of a segment of the short arm of chromosome 4 was detected by Wolf and Hirschhorn and their co-workers in 1965, before the advent of chromosome banding methods. Some cases have been difficult to identify even with routine G banding, however, and the deletion can sometimes be detected only using high-resolution banding or fluorescence *in situ* hybridization (FISH), described in Chapter 29). The specific genes deleted have not been identified yet, but the critical region for the syndrome has been narrowed to a small gene-dense segment of band

4p16.3 (25). In 10% of cases, one parent carries a balanced chromosome 4 rearrangement. The sex ratio is 2:1 female:male. This syndrome is less common than deletion 5p, which has a milder phenotype.

Phenotype

There are prenatal and postnatal growth retardation and severe hypotonia. The common major birth defects include microcephaly, cleft lip and palate, and severe heart malformation. Facial features include cranial asymmetry, prominent forehead, hemangioma, preauricular pits or tags, coloboma of the iris or other eye malformations, cleft lip with or without cleft palate, hypertelorism with a broad and prominent nasal root, micrognathia, and a long neck (Fig. 30.11). Many other birth defects have been seen, including brain and kidney malformations, hernias, abnormal external and internal genitalia, simian creases, and cutis aplasia of the scalp. Because of the severe malformations, many are stillborn or die in the first year, although survival beyond age 20 has been reported. Mental retardation is uniformly profound, and survivors have seizures and severe hypotonia. A microdeletion involving band 4p16.3 has been associated with both Wolf-Hirschhorn syndrome and Pitt-Rogers-Danks syndrome.

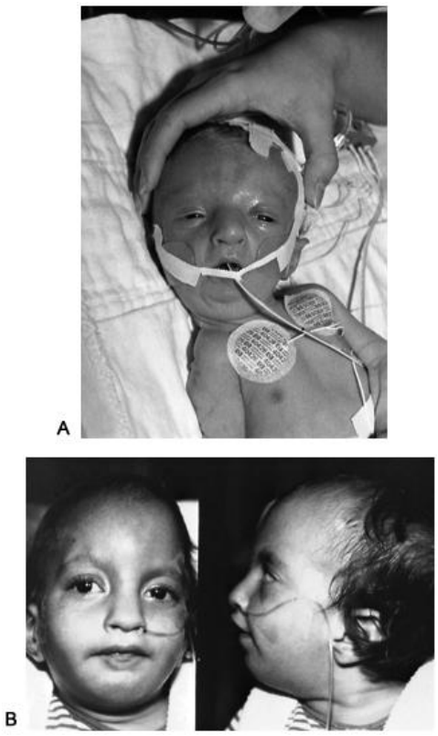


FIGURE 30.11. Two patients with a deletion involving 4p. Note the prominent forehead and nasal root and micrognathia that are common in this syndrome. (**A**, photograph courtesy of D. D. Weaver and C. A. Moore and **B**, photographs courtesy of C. A. Williams.)

Cri-du-Chat Syndrome (Deletion 5p)

In 1963, Jerome Lejeune described loss of material from chromosome 5 as the cause of cri-du-chat syndrome. As with deletion 4p, this deletion was known before chromosome banding was introduced, but some cases with very small deletions have been reported using routine G-banding and high-resolution banding methods. The smallest deletions, involving band 5p15.1, can be easily missed by a routine G-banding study. Before the introduction of chromosome banding, chromosomes 4 and 5 were not distinguishable, but autoradiographic methods revealed differences in the pattern of DNA replication. Thus, it was soon clear that Wolf's and cri-du-chat syndromes were associated with deletions on different chromosomes. The incidence of cri-du-chat syndrome is approximately one in 45,000 newborns, and 12% result from familial rearrangements involving chromosome 5. The mean parental age is normal. Among *de novo* deletions, approximately 80% represent a mutation in the chromosome 5 inherited from the father. Slightly more females than males are affected (26).

Phenotype

Cri-du-chat syndrome can be suspected in infants with a plaintive, meowing cry, low birth weight, and failure to thrive. Some

mothers have described weak fetal movements. Life-threatening birth defects are uncommon, but because of neonatal complications, approximately 10% do not survive beyond the first few months. The newborn frequently has low birth weight and small head circumference. The weak kittenlike cry that gives the syndrome its name is at least partly owing to an anatomic abnormality of the larynx. Although the cry changes with age, it does not become normal. Facial features include microcephaly, a round face, low-set ears, strabismus, broad nasal bridge, and epicanthus, which give the impression of hypertelorism, micrognathia, and occasionally facial asymmetry (Fig. 30.12). The hands and feet are small, and hands often have a simian crease, fifth digit clinodactyly, and single flexion crease. With age, premature gray hair, dental malocclusion, inguinal hernia, diastasis recti, and scoliosis are common. Infants have hypotonia, but older individuals can have normal tone or hypertonia. The IQ varies mostly between 20 and 50 but can range up to mildly retarded, with skills approaching those of 5 or 6 year olds.

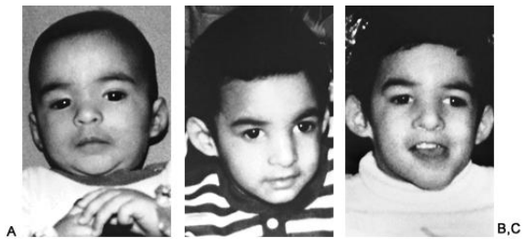


FIGURE 30.12. Three photographs of a patient with cri-du-chat syndrome depict the characteristic facial appearance at ages 5 **(a)**, 24 **(B)**, and 36 **(C)** months.

Russell-Silver Syndrome

Individuals with this syndrome have pre- and postnatal growth retardation, skeletal asymmetry, which can be mild, a normal head circumference with a small triangular-shaped face, clinodactyly, and usually normal intelligence. This is a genetically heterogeneous condition, but approximately 10% of subjects have maternal UPD7. The association with maternal UPD7 was first identified in two unrelated patients with Russell-Silver syndrome and CF who were homozygous for CF mutations that were carried only by their mothers. UPD7 has also been described in several individuals with isolated small stature (primordial dwarfism). The *GRB10* gene on chromosome 7 codes for a protein that binds to the insulin receptor, is probably imprinted (transcribed from the paternal chromosome 7 only), and is likely to be partly responsible for the growth retardation in this syndrome.

Williams' Syndrome (Deletion 7q11.23)

This microdeletion syndrome (23) can be diagnosed clinically, and the diagnosis can almost always be confirmed by FISH analysis showing deletion of the 7q11.23 subband that includes the elastin (*ELN*) gene locus and perhaps other adjacent genes. *ELN* gene mutations have been found in familial forms of supravalvular aortic stenosis, without other features of Williams' syndrome. LIM kinase gene loss may be causally associated with the spatial-cognitive defect, as Williams' syndrome patients with the smallest deletions exhibit loss of both *ELN* and *LIMK1*. The incidence of Williams syndrome is one in 20,000 to one in 50,000 livebirths.

Phenotype

The physical characteristics are somewhat variable, but the key features are supravalvular aortic stenosis (occasionally other heart defects), small stature, prominent lips, medial eyebrow flare, stellate iris pattern, anteverted nares, long philtrum, and a hoarse voice (Fig. 30.13). Hypercalcemia, if present, helps to establish the diagnosis. Mild to moderate mental retardation is typical. The behavioral phenotype is somewhat distinctive and includes an outgoing and loquacious (“cocktail party”) personality.

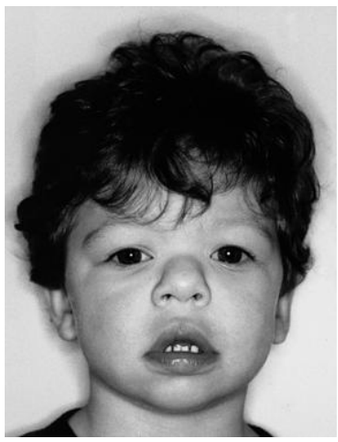


FIGURE 30.13. Patient with Williams syndrome at age 30 months. (Photograph courtesy of E. V. Bawle.)

Trisomy 8 Mosaicism (Warkany's Syndrome)

Most, perhaps all, cases are mosaic with a normal cell line. One commonly finds dramatic differences in the proportion of abnormal cells between tissues. The proportion of trisomy 8 cells appears to decrease with age, especially in the lymphocyte population. The incidence of this condition is estimated to be one in 10,000 newborns, but it is certainly underdiagnosed. The sex ratio is 3:1 male to female. The mean parental age is advanced, suggesting a meiotic nondisjunction leading to trisomy, and a mitotic compensating nondisjunction leading to the karyotypically normal cell population, but the available evidence suggests that trisomy 8 miscarriage usually results from maternal meiotic nondisjunction and that trisomy 8 in liveborn mosaics results from postzygotic error.

Phenotype

A normal birth weight with postnatal growth retardation is typical. The face is only mildly dysmorphic with a prominent forehead, malformed ears, bulbous nose with upturned nares, everted lower lip, protruding lower helix of ear, down-slanting palpebral fissures, and mild micrognathia (Fig. 30.14). Multiple skeletal abnormalities are common, including additional vertebrae or hemivertebrae, spina bifida occulta, broad dorsal ribs, hypoplastic iliac wings, and hypoplastic patellae. Other features include a long and slender body habitus, simian creases, deep plantar and palmar furrows (which may disappear in late childhood) and a similar deep furrow on the lower lip, mild joint contractures, mild congenital heart defects, and absent corpus callosum. Males often have hypospadias, cryptorchidism, or inguinal hernia. Life-threatening birth defects are unusual. Mental retardation is usually moderate but widely variable.

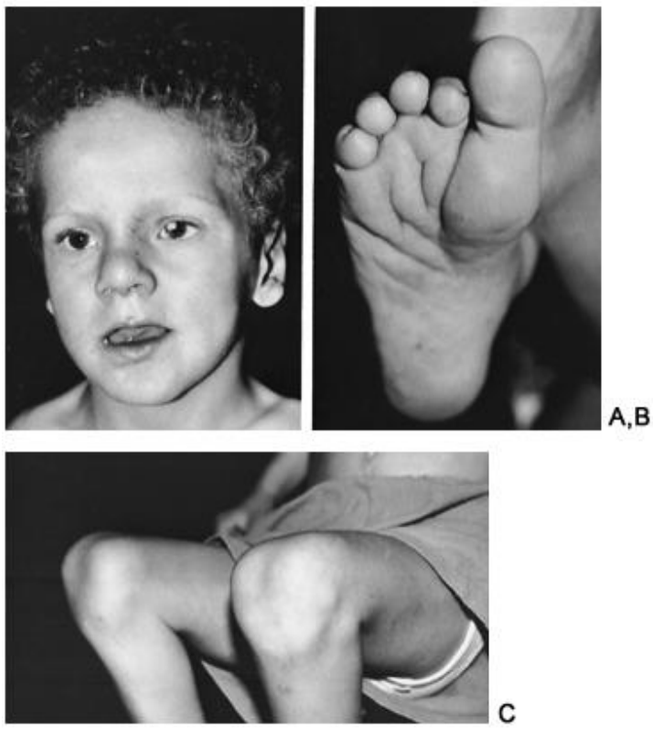


FIGURE 30.14. Trisomy 8 mosaic syndrome includes malformed ears with protruding lower helix, bulbous nose, and downward slanting palpebral fissures **(a)**. Deep creases are common in the palms and soles **(B)**, and absent or hypoplastic patellae **(C)** are characteristic. (Photographs courtesy of C. A. Williams.)

Secondary Trisomy 9p [47,+dic(9p) or 47,+i(9p)]

An extra isochromosome 9p has been identified in 13 patients, and we identified it once in a prenatal genetic study. The karyotypic interpretation can be confirmed by the observation of elevated galactose-1-phosphate uridylyl transferase (GALT) enzyme activity. The GALT gene, resides within band 9p13. The isochromosome is often dicentric. There is tissue-limited mosaicism in which most lymphocyte metaphases carry the isochromosome, whereas cultured fibroblast cells exhibit mosaicism. This reduced expression in fibroblasts may carry over to amniocytes and chorionic villus cells and make prenatal diagnosis more difficult. The average maternal age (31 years) is advanced.

Phenotype

Birth defects have included congenital heart disease, cleft lip and palate, joint dislocations, hemivertebrae, scoliosis and kyphosis, hypoplastic clavicles, kidney malformations, cryptorchidism in males, and simian creases. Facial features include hypertelorism and micrognathia. Approximately one half have died within a year. The severity of mental retardation has been variable, and one affected man has been described as semiindependent with an IQ of 63.

Beckwith-Weidemann Syndrome (Duplication of Band 11p15.5)

The overall incidence of Beckwith-Wiedemann syndrome (BWS) (or EMG syndrome: *e*xomphalos, *m*acroglossia, and *g*igantism) has been estimated as one in 14,000 livebirths. There is considerable genetic heterogeneity (27). Approximately 65% of patients have a normal karyotype and have no family history. These subjects apparently represent new mutations, perhaps some of which involve an imprinting control region on chromosome 11. Autosomal dominantly inherited Beckwith-Wiedemann syndrome (15% of cases) has been localized to 11p15.5, and all affected individuals inherit the mutation from their mother. The gene encoding insulinlike growth factor 2 (IGF2) is imprinted and expressed only on the paternally derived chromosome 11, and another gene in the BWS critical region, H19, is imprinted and expressed only on the maternally derived chromosome 11. It is likely that these two genes play a role in the disease. A visible chromosomal duplication involving band 11p15 is present in approximately 3% of cases, and a paternal origin of the duplication was identified in all the patients evaluated so far. The remaining 20% of cases appear to arise sporadically as a result of paternal UPD 11. In contrast to other disorders associated with UPD, here the UPD appears to be segmental (confined to part of chromosome 11) and appears to arise postzygotically, resulting in mosaicism for UPD and normal biparental disomy (28).

Phenotype

Patients with the visible chromosome change have mental retardation and usually congenital heart disease; some have been stillborn. Other features of the syndrome include gigantism and generalized organomegaly, a large tongue, and omphalocele (Fig. 30.15). Neonatal hypoglycemia is identified in 35% to 50% of subjects. Craniofacial features include microcephaly, linear creases on ear lobes, indentations on the posterior edge of the helix, and nevus flammeus. Cleft lip is an occasional finding. Among the dominantly inherited cases, there is approximately an 8% risk of cancer: Wilms' tumor (60% of tumors), adrenocortical carcinoma (15%), and other tumors, including gonadoblastoma, hepatoblastoma, and rhabdomyosarcoma. Presumably, the cases with a visible duplication have a similar risk. Genes that reside in this region include insulin and IGF. If they are duplicated, they may account for some of the phenotypic features, such as macrosomia and neonatal hypoglycemia. Hemihypertrophy and Wilms' tumors appear to be more common in the UPD group, and omphalocele more frequent in the patients with biparental inheritance of chromosome 11. In tumor cells, loss of heterozygosity is observed for 11p genes, including the HRAS1 locus (see Chapter 34). Mental retardation can be secondary to the associated hypoglycemia or the chromosome abnormality but is not a constant feature of the autosomal dominantly inherited syndrome.

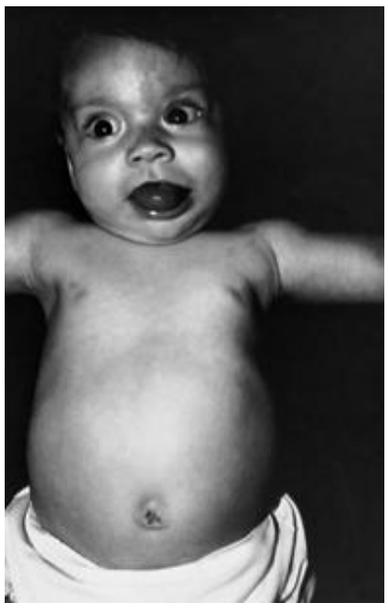


FIGURE 30.15. Patient with Beckwith-Wiedemann syndrome. (Photograph courtesy of C. A. Williams.)

WAGR Syndrome (Deletion of Band 11p13)

This contiguous gene syndrome is associated with the loss of band 11p13 and occasionally with a familial chromosome rearrangement. The incidence of WAGR (*W*ilms' tumor, *a*niridia, *g*enitourinary malformations, and mental *r*etardation) syndrome is approximately one in 50,000 live births, and the association between aniridia and Wilms' tumor was known many years before the chromosomal basis was identified. Either aniridia or Wilms' tumor alone can be inherited as a discrete autosomal dominant trait with normal intelligence and no visible deletion. Expression of the WAGR syndrome requires deletion of both the *PAX6* gene, associated with familial aniridia, and the *WT1* gene, associated with the genitourinary malformations and the increased risk of Wilms' tumor. Approximately one in 70 Wilms' tumor patients has aniridia. The sex ratio is approximately two males to one female.

Phenotype

Wilms' tumor (nephroblastoma) is identified in 35% to 50% of patients with the deletion, and males have an increased risk of gonadoblastoma. Genitourinary malformations are present in 90%, including ambiguous genitalia in most males. Aniridia is present in 97% of patients who have the other three major features. Craniofacial features include microcephaly (in 50%), prominent lips, micrognathia, and malformed ears. Moderate to severe mental retardation is present in 90% of cases.

Secondary Trisomy 12p, Killian's or Pallister-Killian Syndrome, Mosaic Tetrasomy 12p, 47,+i(12p)

This syndrome results from an extra isochromosome 12p, and all cases are mosaic. The isochromosome looks very much like an i(21q), and many cases have been misidentified as such and called tetrasomy 21. The phenotype is not similar to Down's syndrome. Increased expression of the LDHB enzyme (lactate dehydrogenase B, located within band 12p12) has confirmed the abnormality to be a +i(12p), and today FISH would be the quickest way to confirm the chromosome 12 origin. More than 60 cases have been reported (29), and we have seen at least four cases in our laboratory. Tissue-limited chromosomal mosaicism and early perinatal death have probably resulted in most diagnoses being missed, so the incidence of this syndrome is uncertain. More than 99% of lymphocytes have a normal karyotype, whereas at least 10% of fibroblasts typically have the extra i(12p). Cells with the isochromosome are probably present in the blood prenatally but disappear very soon after birth. Alternatively, patients who are identified in the newborn period may be more severely affected because of the involvement of other tissues, including peripheral blood. A high proportion of abnormal cells has been reported in bone marrow preparations from subjects who had very few cells in phytohemagglutinin (PHA)-stimulated blood cultures, suggesting poor PHA response by, or *in vitro* selection against, lymphocytes that carry the isochromosome. Parental age is more advanced than in trisomy 21, with a mean maternal age of 31 years and a mean paternal age of 34 years.

Phenotype

The birth weight is usually normal or elevated. Diaphragmatic hernia is a key feature and a common cause of perinatal death. Associated features include severe hypotonia, short limbs, imperforate anus, and talipes. Craniofacial features include a high forehead, sparse hair, small dysplastic ears, hypertelorism, irregularly shaped bushy eyebrows, flat nasal bridge with short nose and anteverted nares, and thin upper lip. Other findings include supernumerary nipples, sacral dimple, short fingers, broad hands and feet, vertebral and joint deformities, and disordered or dysplastic skin pigmentation, dental eruption, and hair distribution. The pigmentary dysplasia can suggest hypomelanosis of Ito or incontinentia pigmenti with swirls of hypopigmentation or can be expressed as sparse hypopigmented macules that are clearly seen only with a Wood's lamp. Those who live beyond the neonatal period usually have seizures and severe mental retardation. The first described cases were bedridden patients who had mental retardation, seizures, and joint contractures. It is important to distinguish this sporadic condition from autosomal recessively inherited Fryn's syndrome.

Trisomy 13 (Patau's Syndrome)

Trisomy 13 is present in one in 5,000 newborns. As with other autosomal trisomies, the average maternal age is advanced. Primary trisomy 13 is present in 75% of cases, 4% are mosaics, 10% have an unbalanced der(13;13), and 10% have an unbalanced der(13;14). Other robertsonian translocations associated with trisomy 13 are rare. At least one half of the der(13;14) but fewer than 10% of the der(13;13) cases are familial. The der(13;14) is the most common balanced rearrangement in humans, with a frequency of approximately one in 5,000 in the general population. The reproductive risk to a der(13;13) carrier parent is 100%, and for a der(13;14) carrier parent the reproductive risk is estimated to be 1% for a carrier female and less than 1% for a carrier male. The majority of der(13;14) carrier parents of unbalanced progeny are female, and as far as we know, all the der(13;13) carrier parents of unbalanced progeny have been female.

Phenotype

Many trisomy 13 conceptions result in miscarriage or stillbirth. In liveborn infants, the common and key physical features are microcephaly, localized areas of cutis aplasia on the scalp, microphthalmia, bilateral cleft lip, and polydactyly with hyperconvex and narrow fingernails (Fig. 30.16). Midline facial defects are commonly associated with holoprosencephaly and can be as severe as cyclopia. Multiple ocular malformations have been reported, with microphthalmia or anophthalmia most common. Other features include malformed low-set ears, micrognathia, short neck with extra skin folds at the nape of the neck, broad flat nose, and hemangiomata. Other malformations include heart defects (ventricular or atrial septal defect or patent ductus arteriosus), kidney malformations, single umbilical artery, cryptorchidism

in males, and a single palmar crease. Polymorphonuclear cells have characteristic nuclear projections, and persistent embryonic hemoglobin has been reported. Mental retardation is profound and seizures are common. Half of trisomy 13 liveborn infants die within the first month and 90% die within the first year. Survival longer than 5 years is unusual.

FIGURE 30.16. Fetus with trisomy 13, with bilateral clefts and postaxial polydactyly. (Photograph courtesy of J. R. Roberson and L. Weiss.)

Retinoblastoma (Deletion of Band 13q14)

A contiguous gene syndrome of retinoblastoma and birth defects is associated with an interstitial deletion of a small segment of chromosome 13 including band 13q14 that can usually be detected by G banding or FISH. The deletion usually arises as a new mutation of paternal origin. The dominantly inherited retinoblastoma gene (RB1) resides within band 13q14. The incidence of retinoblastoma is approximately one in 23,000 liveborn infants. Among patients with retinoblastoma, the tumors are bilateral in approximately 20%, and 5% of the bilateral cases have a visible chromosome 13 deletion. Retinoblastoma patients without birth defects (sporadic or dominantly inherited type) have normal 13s, although some of their tumors have a visible del(13) or loss of heterozygosity at 13q14. This is discussed in greater detail, along with the two-mutation model for carcinogenesis, in Chapter 34. The esterase D gene is closely linked to the RB1 locus and is deleted in approximately 95% of deletion cases.

Phenotype

The craniofacial features are variable but include macrocephaly (or, with larger deletions, microcephaly), prominent eyebrows, nasal bridge with a bulbous nasal tip, and a wide mouth with a thin vermilion border. There is a high risk of retinoblastoma and osteosarcoma, and most patients have been ascertained because of the association of retinoblastoma and birth defects. In approximately half of the cases with retinoblastoma, the tumor is unilateral. A few patients who have a 13q deletion that includes loss of the esterase D gene do not have retinoblastoma. Even so, these patients may remain at risk for other neoplasias. Molecular genetic studies to detect submicroscopic mutations within the retinoblastoma gene may be useful in such patients. Mental deficiency can be severe, but some patients who have a visible deletion reportedly have normal intelligence. Familial deletions and balanced chromosome 13 rearrangements have been reported, as well as duplication of band 13q14 associated with a normal phenotype in a few relatives.

Prader-Willi Syndrome (Prader-Labhart-Willi Syndrome) [del(15)(q11.2q12) or UPD15mat]

The incidence is at least one in 20,000 livebirths, and the recurrence risk is less than 1%, with very few familial cases reported. Most patients (70%) have an interstitial deletion of 15q11.2–q12 or q13 that always includes the loss of subband 15q11.2 and always involves the chromosome 15 inherited from the father. Approximately 20% of cases result from maternal UPD15; the genetic error in most UPD15 cases is maternal meiotic nondisjunction followed by a compensating mitotic nondisjunction (Fig. 30.10), and there is an increased mean maternal age in this group. A few patients (less than 5%) have some other chromosome 15 abnormality, such as a robertsonian translocation or an extra bisatellited dicentric, dic(15)(q11), associated with UPD of the structurally normal chromosome 15 pair. The cytogenetics of PWS is of more than academic interest when it bears on the clinical interpretation of chromosome 15 abnormalities in prenatal diagnostic cases (Chapter 31). Within the region 15q11.2q12 is a series of genes that are known to be imprinted and expressed only from the paternally inherited chromosome 15 (Chapter 29). The most appealing candidate gene is *SNRPN*, which is imprinted and expressed only from the paternal copy of chromosome 15 (30). A very small number of PWS patients appears to have a deletion, translocation, or point mutation affecting the function of a hypothesized “imprint control element” that controls the imprinting pattern of *SNRPN* and other genes in the 15q11.2q12 region.

The most efficient laboratory diagnosis for PWS employs a DNA methylation study as the first step because that assay will detect either a deletion or UPD (31). If the DNA results are abnormal, FISH analysis is appropriate to distinguish deletion from UPD. Karyotype analysis is appropriate in all cases and should be considered concurrently with the DNA analysis or at least if the DNA results are normal to look for other chromosome abnormalities including a distal 1p or 2q deletion (Table 30.7).

Phenotype

Birth weight and head circumference are generally normal. Prematurity, decreased prenatal movement, and breech presentation are common. As infants, these patients have hypogonadism (cryptorchidism and micropenis in males and hypoplastic labia in females), mild ocular and cutaneous hypopigmentation, strabismus, significant hypotonia, and failure to thrive; many need to have special help with feeding because of a weak suck (32). Some parents have commented about the child having cold hands and feet and skin mottling. Craniofacial features include narrow bitemporal diameter, high forehead, almond-shaped eyes, epicanthus, malar hypoplasia, and small mouth with down-turned corners. By 3 years of age, hyperphagia leading to marked obesity and small hands and feet are characteristic features (Fig. 30.17). The patients with UPD usually have a milder phenotype and tend to be identified at an older age than do the deletion cases (33). Mental retardation is usually moderate (average IQ is 65), significant behavioral problems are common, and approximately 10% have seizures.

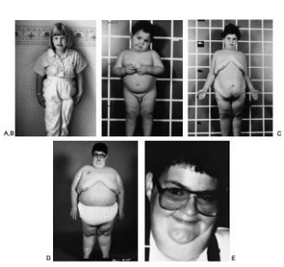


FIGURE 30.17. Although children with Prader-Willi syndrome characteristically have significant obesity, in some cases weight can be controlled through the efforts of committed parents working with dedicated health care professionals and support groups, as shown by the Prader-Willi syndrome patient depicted in **A** at age 5. Another patient shown in **B–E** at ages 4, 11, and 21 years has the more typical phenotype. (Photographs courtesy of L. Weiss and the parents of the patients.)

Angelman Syndrome [del(15)(q11.2q12) or UPD15pat]

The incidence is approximately one in 15,000 births, with most cases diagnosed after 2 years of age. This is a contiguous gene syndrome in which approximately 50% of patients have a *de novo* 15q11.2-12 microdeletion or other rearrangement visible with high-resolution G banding, and another 35% have a deletion detectable by FISH. The deletion in AS is indistinguishable from that seen in PWS, with the critical distinction that it always involves the chromosome 15 inherited from the mother. Approximately

5% of subjects have paternal UPD15, again in contradistinction to the maternal UPD15 in PWS. The AS recurrence risk is very low but is apparently higher than for PWS because several families have two or more affected members, and some families appear to transmit a defective gene at this locus or at the hypothesized imprint control element, with the syndrome only expressed when the mutation is inherited from the mother. A few families have a chromosome 15 rearrangement that is unbalanced in the affected family members. The gene responsible for most if not all the AS phenotype is UBE3A, which is expressed only from the maternal copy of chromosome 15 in the developing brain but appears to be expressed equally from both 15s in other tissues (34). Several cases have been found with point mutations within the UBE3A gene (35,36). The risk of recurrence is estimated to be <1% for the *de novo* deletion and UPD cases but as high as 50% for others. Molecular and cytogenetic laboratory diagnosis for AS employs the same steps as discussed for PWS.

Phenotype

In AS, there is microcephaly of postnatal onset, developmental delay, hypotonia with brisk reflexes, blue eyes and fair skin compared with other family members, and choroid and iris pigment hypoplasia (25% have oculocutaneous albinism). Seizures and an abnormal electroencephalogram (EEG) are typical with onset by age 10 months. The diagnosis is usually not suspected until at least age 2, after the appearance of the characteristic hyperkinetic behavior with a stiff, ataxic gait. The facial expression is alert with frequent and inappropriate smiling and laughter (Fig. 30.18). Mental retardation is severe.

FIGURE 30.18. Patient with Angelman syndrome. (Photograph courtesy of L. Weiss.)

Secondary Trisomy 15q1pter [47,+dic(15)(q1)]

A dic(15)(q1) present as a 47th chromosome has been described in numerous subjects, and it is present in approximately one in

5,000 liveborn infants. The phenotype depends largely but not entirely on the amount of material duplicated (actually triplicated) (37). In the smaller dicentrics, with a karyotype of 47,+dic(15)(q11.1), the two centromeres are very close together with little 15q material duplicated. Such bisatellited markers are often familial and consistent with an entirely normal phenotype. However, these can be associated with PWS or AS owing to UPD of the two structurally normal 15s or a 15q deletion. Larger dic(15) chromosomes that have an acrocentric appearance on G-banding are usually new mutations and are associated with mental retardation and mild dysmorphic features. The breakpoint is between 15q11.2 and 15q15, so the PWS/AS critical region is included in the dicentric. The secondary trisomy typically arises *de novo*, usually is maternal in origin, and the average maternal age is advanced (mean, 35 years).

Smith-Magenis Syndrome (Deletion of Band 17p11)

This rare contiguous gene syndrome is associated with an interstitial deletion involving band 17p11 only. In all but one case, the deletion has represented a new mutation (the mother in one case appears to be mosaic), and the deletion apparently involves the paternal chromosome 17 in most cases. One relatively mildly affected patient has a *de novo* translocation with a breakpoint within band 17p11. The deletion is usually detectable by karyotype analysis, but if the diagnosis is suspected and the karyotype appears to be normal, FISH analysis is indicated. There may be as few as two genes deleted to create this syndrome.

Phenotype

The phenotype (Fig. 30.19) is variable but includes failure to thrive, brachycephaly, prominent forehead, microcephaly, flat and broad midface, broad nasal bridge, strabismus, myopia, malformed ears, high or cleft palate, prognathism, short and broad hands and feet, scoliosis, and cryptorchidism (38). The voice is hoarse and sensorineural or conductive hearing loss is common. Heart and kidney abnormalities are present in approximately one third of subjects. Mental retardation is variable but usually severe with seizures and hyperactivity. Unusual features of the syndrome include decreased reflexes, decreased sensation, and other features of peripheral neuropathy, specific self-destructive behavior, including insertion of foreign objects into bodily orifices and pulling out finger and toe nails, and sleep abnormalities (especially disturbed rapid eye movement sleep).

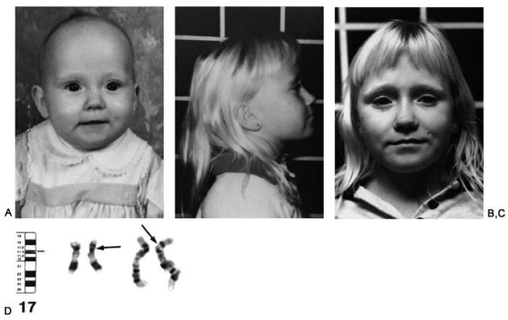


FIGURE 30.19. Patient with Smith-Magenis syndrome at age 9 months **(a)** and 8 years **(B and C)**. She has an IQ of approximately 70 and has a significant sleep disturbance. She has a *de novo* interstitial deletion involving band 17p11 only **(d)**. Two chromosome 17 pairs are shown, with arrows pointing to band 17p11 in the normal homolog. The deletion homolog is at the left in each pair. (**A–C**, photographs courtesy of J. R. Roberson and L. Weiss.)

Miller-Dieker Syndrome (Deletion of Subband 17p13.3)

This rare contiguous gene syndrome has been associated with the loss of subband 17p13.3, owing to either a simple deletion or an unbalanced chromosome rearrangement. Parental karyotypes need to be studied carefully to look for a familial translocation, which, if present, would indicate a significant risk of recurrence. Nearly one half of the patients with a 17p deletion have normal-appearing chromosomes even under high-resolution analysis, so FISH is an appropriate adjunct to high-resolution karyotype analysis if the diagnosis is suspected. In some patients with isolated lissencephaly, a submicroscopic deletion of the proximal segment of band 17p13.3 has been identified. Mutations in the *LIS1* gene are responsible for the lissencephaly in many cases of isolated lissencephaly, and in Miller-Dieker syndrome, loss of perhaps one other gene is responsible for the heart and kidney defects.

Phenotype

A key feature is microcephaly and a prominent forehead with vertical skin furrowing and bitemporal narrowing (Fig. 30.20). The phenotype includes type I lissencephaly (cerebral agyria or smooth brain with a four-layered cortex), profound EEG abnormality, seizures, profound hypotonia, severe to profound mental retardation, and pre- and postnatal growth retardation. Other facial features are ptosis, upturned nares, long philtrum with thin upper lip, mild micrognathia, and malformed ears. Heart and kidney defects are common. Most die in infancy.

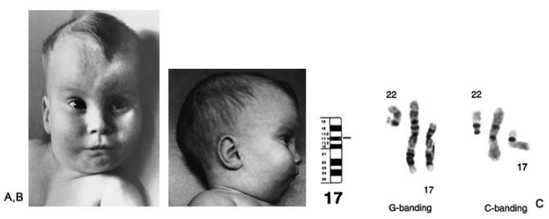


FIGURE 30.20. This Miller-Dieker syndrome patient had a *de novo* dicentric translocation between chromosomes 17 and 22. At age 12 months, she was beginning to smile socially but still had very poor head control **(A, B)**. Karyotype studies showed this patient had one normal copy of chromosomes 17 and 22. The other 17 and 22 were represented by a dicentric chromosome that contained most of 22p and appeared to have lost at least part of band 17p13. Two partial karyotypes are shown **(C)**. Molecular genetic studies confirmed the deletion of this region and demonstrated the paternal origin of the chromosome rearrangement (Unpublished results from D. Ledbetter). (**A, B**, photographs courtesy of L. Weiss.)

Trisomy 18 (Edwards' Syndrome)

Primary trisomy 18 is identified in one in 6,000 to 8,000 liveborn infants but is much more common among stillborns and represents approximately 6% of second-trimester miscarriages with an autosomal trisomy. Translocations are very uncommon. There is a maternal age effect, and the nondisjunctional event occurs

during maternal meiosis in approximately 95% of cases. The sex ratio is approximately 4:1 female to male.

Phenotype

Prenatal movement is weak, and polyhydramnios and a single umbilical artery are common. There is significant pre- and postnatal growth retardation (Fig. 30.21). Newborns are hypotonic but later become hypertonic. The head is microcephalic with a prominent occiput, narrow midface, and micrognathia. Ears have a simple helix, the nose is small with upturned nares, and the mouth is small with a high-arched palate. Heart, kidney, and genitourinary tract malformations are common. The hands are typically clenched with the second finger overlapping the third finger, and fifth overlapping the fourth. Among the many other features of trisomy 18 syndrome are single palmar crease, a simple

arch dermatoglyphic pattern on multiple fingertips, loose skin at the nape of the neck, hypoplastic nipples, dislocated hips, rockerbottom feet, and diaphragmatic hernia. Mental retardation is profound. Half of liveborn infants die within 2 months and fewer than 10% survive 1 year. Survival beyond age 5 is exceptional. Chromosomal mosaics with a normal cell line constitute approximately 10% of liveborn infants, and the phenotype can be somewhat moderated in these patients.

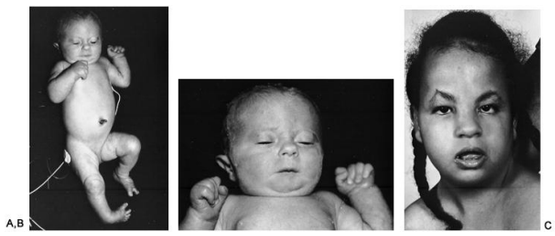


FIGURE 30.21. Infant (**A** and **B**) with trisomy 18. **C**, This 25 year-old female with nonmosaic trisomy 18 is one of the longest surviving patients with this syndrome. (Photographs **A** and **B** are courtesy of C. A. Williams; **C**, photograph courtesy of L. Weiss.)

Deletion 18p

Deletion of part of this small arm was identified before the use of banding techniques and in fact was the first published deletion syndrome. Some ring chromosome 18 patients have a deletion 18p phenotype. Most 18p deletions are new mutations, but the parents should be karyotyped to look for a balanced rearrangement. The average parental age is advanced, which is atypical for chromosome deletions and is unexplained. Slightly more females than males are affected.

Phenotype

Approximately 10% to 15% of patients have gross craniofacial malformations including holoprosencephaly, but in the patients without major brain malformations, there are usually no life-threatening birth defects. The physical features include growth retardation and facial features of mild microcephaly, low-set, soft, and malformed ears, strabismus, ptosis, epicanthus, flat nasal bridge, upturned nares, Cupid's bow of upper lip, and micrognathia. Dentition is poor and severe caries are very common. A short neck, broad chest, and mild edema of hands and feet can be reminiscent of Ullrich-Turner or Noonan's syndrome. Other features include hernias and IgA deficiency. Mental retardation is typical but severity varies.

Trisomy 20

This unusual trisomy is found in miscarriage specimens, and mosaic trisomy 20 is found in approximately one in 2,500 amniotic fluid chromosome studies (39). There is no syndrome of birth defects associated with trisomy 20 mosaicism, and most newborns have been phenotypically normal. The mosaicism has been confirmed in foreskin fibroblasts and other tissues but never in peripheral blood cells. The risk of malformation or mental retardation seems low for cases identified prenatally, but this tissue-limited mosaicism is still poorly understood.

Trisomy 21 (Down's Syndrome)

One infant in 800 to 1,000 has Down's syndrome (40). Primary trisomy 21 accounts for approximately 93% of all cases. Another 2% to 3% are mosaics, 3.5% to 5% have a robertsonian translocation, and fewer than 1% have other rearrangements. The sex ratio among Down's syndrome liveborn infants is approximately three males to two females. In 1960, 11% of births were to women over age 35, whereas this frequency is only approximately 5% today, and in the United States, nearly one half of pregnant women over age 35 obtain prenatal diagnosis. This has resulted in a small decrease in Down's syndrome births overall but a relative increase of Down's syndrome babies born to younger women. The availability of serum screening the pregnant population (Chapter 31) is decreasing the incidence of Down's syndrome further. Rare patients with duplication of small segments of chromosome 21 have been instructive in the development of a phenotype map of chromosome 21 in which duplication of band 21q22.3 is essential for the expression of the Down's syndrome phenotype (Fig. 30.22A).

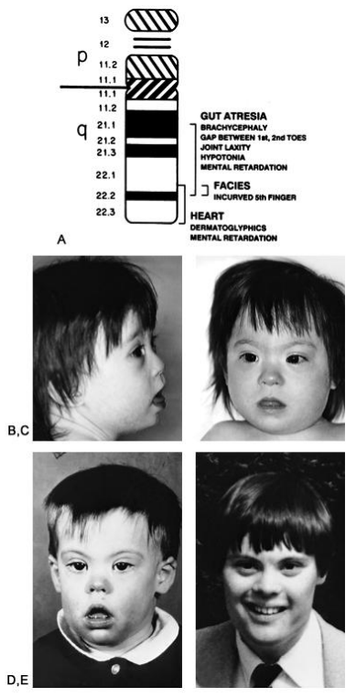


FIGURE 30.22. The phenotype of Down's syndrome. Individual stigmata of the Down's syndrome phenotype have been mapped by molecular analysis of small duplications involving chromosome 21 **(a)**. **B–D:** the childhood phenotype is illustrated, showing many of the features detailed in the text. The male shown in **D** is shown in **E** at age 18 years. (**B, C**, photographs courtesy of D. D. Weaver and C. A. Moore; **D, E**, photographs courtesy of L. Weiss and J. R. Roberson.)

The maternal age effect in trisomy 21 is illustrated in Fig. 30.23, and rates of Down's syndrome and other trisomies are

provided in Table 30.3. There is also a maternal age effect in mosaic trisomy 21; this finding and other studies suggest that most mosaics begin as trisomic conceptions, and anaphase lag gives rise to the diploid cell line. The normal cell line may have a selective growth advantage that results in an increase in the proportion of normal cells with age, especially in lymphocytes.

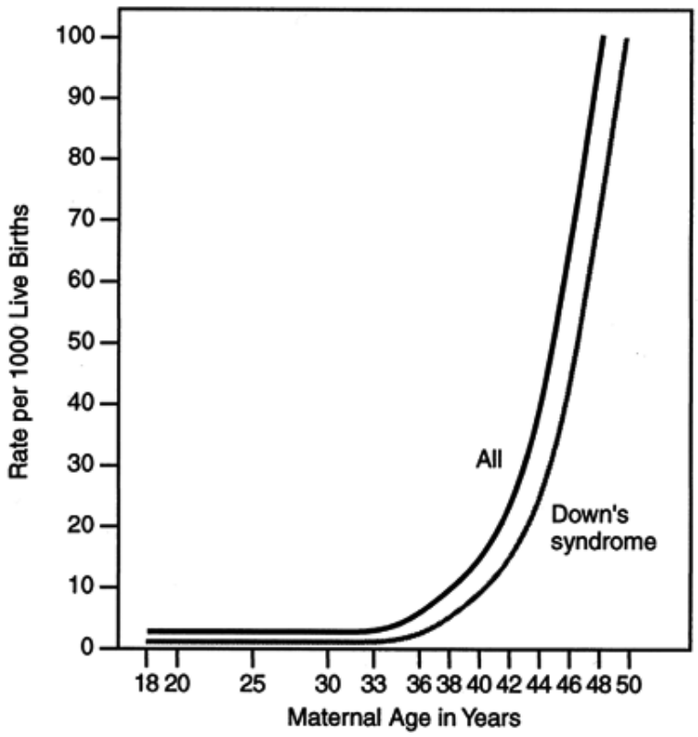


FIGURE 30.23. Frequency of Down's syndrome, and of all trisomies, among liveborn infants, by maternal age. [Statistics from Hook (3).]

In a few trisomy cases (3% in one study, 0.3% in another study), one parent appears to be a trisomy 21 mosaic, and presumably the trisomy 21 offspring resulted from secondary nondisjunction. The incidence of parental mosaicism is probably greater among parents with more than one child with the same trisomy and may be lower in older parents. The mosaic parents in general have normal intelligence and no obvious Down's syndrome stigmata. It has not become standard practice to study the chromosomes of parents of trisomy 21 patients, partly because it is impossible to disprove that a parent is a low-level mosaic and partly because even if one or two +21 cells are observed among 100 cells, the chance remains that this is a technical artifact rather than a clear demonstration of +21 mosaicism. Instead, it seems reasonable based on the available data to advise the parents of a risk of recurrence of at most 1% more than the maternal age-associated risk. Experience shows approximately a 0.5% recurrence risk for trisomy 21 plus 0.5% for other chromosome abnormalities (the latter risk does not appear to be significantly greater than the general population risk). It has also been proposed that a woman who has a trisomy 21 miscarriage has an increased risk of having a trisomy 21 liveborn infant, but the evidence for this, independent of maternal age effects, is weak. In general, other relatives of a trisomy 21 individual (e.g., siblings) do not have an increased risk beyond their maternal age-associated risk of having a trisomy 21 child.

The majority of translocations associated with Down's syndrome are der(14;21) or der(21;21). Approximately one half of der(14;21) cases are inherited, but only 5% of der(21;21) cases are inherited (Table 30.8). This low frequency of inherited der(21;21) cases, together with recently presented molecular genetic studies, is consistent with the hypothesis that these are often isochromosomes rather than robertsonian translocations, but another factor in their high mutation rate is that a der(21;21) cannot be passed through multiple generations. In most inherited

cases, the mother is the carrier; there appears to be reduced fertility in some carrier males and lower empirical risk of Down's syndrome in the offspring of carrier males. Recurrence has been reported in several instances of apparently *de novo* der(21;21) or other robertsonian translocation Down's syndrome cases, suggesting germ line mosaicism in a parent, and as with trisomy 21, it is reasonable to advise parents of *de novo* translocation cases that the risk of recurrence is approximately 1% more than the maternal age-associated risk.

TABLE 30.8. ROBERTSONIAN TRANSLOCATIONS IN DOWN'S SYNDROME*a*

			Risk to Carrier[c] (%)	
	% of Cases[a]	% Inherited[b]	Female	Male
der (14;21)	51.7	50	14	4
der (13;21)	2.9	-[d]	-	-
der (15;21)	4.6	-	-	-
der (21;21)	37.2	5	100	100
der (21;22)	3.5	-	7	≤3

[a] Based on data from Therman et al. (41) and cases from our laboratory.

[b] Based on a sample of 4,760 cases of Down's syndrome from Giraud and Mattei (42).

[c] Risk of bearing an offspring with Down's syndrome owing to an unbalanced form of the translocation. Statistics for der (14;21) are from Daniel et al. (43). Statistics for der (21;22) are from Chapman et al. (44).

[d] A – indicates that the data are insufficient to make an estimate. The proportion of inherited der (13;21), der (15;21), and der (21;22) may be similar to that of der (14;21). The estimate for der (21;21) also includes our experience of 10 *de novo* and zero inherited cases. However, the 5% seems to be an overestimate.

The increase in trisomy with advancing maternal age is reflected in a relative decrease in translocation cases with advancing maternal age. This is helpful in genetic counseling situations when the karyotype of the affected person and the parents is unknown and they are unavailable for evaluation. If the mother's age was under 30 when the affected person was born, the chance of a familial translocation is approximately 1.5%, whereas that chance is 0.4% if maternal age was 35 years, and 0.04% if maternal age was 40 years.

Phenotype

As many as 75% of trisomy 21 conceptions result in miscarriage (Table 30.2), and 1% of stillborns have trisomy 21. Among trisomy 21 pregnancies identified by midtrimester amniocentesis and not electively interrupted, approximately one fourth abort spontaneously or are stillborn. There is significant pre- and postnatal growth retardation with reduced birth weight. Heart defects, mainly endocardial cushion defects (atrioventricular septal defect, ventricular septal defect, patent ductus arteriosus, or tetralogy of Fallot) are present in 40% of subjects. The facial features in Down's syndrome commonly include a round and flat face, ear malformations such as a folded upper helix, epicanthus, upward-slanting palpebral fissures, Brushfield spots, flat nasal bridge, and open mouth with a protruding tongue (Fig. 30.22B,Fig. 30.22C and Fig. 30.22D). The hands frequently have a transverse palmar crease, short fifth finger with a single flexion crease, and clinodactyly. There is often a gap between toes 1 and 2. Many have an umbilical hernia. Radiologically, there is a small pelvis with hypoplastic iliac wings and small acetabular angle. A multitude of less common birth defects in Down's syndrome include strabismus, cataracts, duodenal atresia, imperforate anus, and an increased risk of leukemia. Hundreds of genes are duplicated in trisomy 21, and those that play the greatest role in Down's syndrome are beginning to be understood (45).

Newborns are hypotonic, especially if a heart defect is present, and have poor reflexes. Many have loose skin on the back of the neck that is a remnant of prenatal lymphatic edema. They do not have a significant adolescent growth spurt. Adolescents and adults have short stature [standard growth charts for Down's syndrome are available in Cronk et al. (46)]. A tendency to be overweight is typical. Hypothyroidism is common (and treatable) among infants and children with Down's syndrome. There is also an increased risk of atlantoaxial instability, diabetes mellitus, depression, premature aging of the skin, and dementia. There are numerous other health issues that arise more commonly throughout life (47).

Only one male is known to have reproduced (one chromosomally normal offspring). Females have reproduced, and approximately 40% of their offspring have had Down's syndrome owing to secondary nondisjunction. (The expected 50% risk is not met, perhaps because of reduced viability of the +21 embryos.) Congenital heart disease and heart failure, respiratory tract infection, intestinal obstruction, and leukemia are major factors in mortality. Of patients without heart defects, 90% survive 1 year and 80% survive to age 30. Those who have heart disease are smaller on average, with 75% surviving 1 year and 50% to age 30. Only 45% of Down's syndrome patients are expected to survive to age 60, compared with 86% of the general population.

Developmental delay is not always evident until after 1 year of age because the infants tend to be content and only modestly delayed in milestones that are normally attained by 12 to 18 months. Mental retardation is usually mild to moderate, and children are often placed in education programs for the trainably or educably mentally impaired. In institutionalized populations, there is an excess of males among the Down's syndrome patients, probably because affected males present more severe behavior problems. Whereas most were institutionalized years ago, today most adults with Down's syndrome live with family or in group homes, and many work in sheltered situations. As marketplace demands have increased for unskilled labor, an increasing number of adults with Down's syndrome find employment.

There have been no prospective studies of the phenotype and development in trisomy 21 mosaics, so the likelihood that a mosaic will have a moderated phenotype is unknown. Nevertheless, as is true of other mosaics, individuals with a milder or atypical

Down's syndrome phenotype appear more likely to have mosaicism with some karyotypically normal cells (48). As stated earlier, a few individuals with low-level +21 mosaicism have been identified after having a Down's syndrome child.

Cat Eye Syndrome (Secondary Trisomy 22q11pter)

Patients with this phenotype often have a *de novo* extra bisatellited chromosome derived from chromosome 22 (an example of a bisatellited chromosome 15 is provided in Fig. 30.3J). *In situ* hybridization using a chromosome 22q11 DNA probe has shown that this band is consistently duplicated. Although the defect usually arises *de novo*, in two cases, the parent carried the extra isochromosome but had normal intelligence and a diagnosis of Duane's syndrome (radial ray abnormalities and deafness). Maternal origin of the *de novo* extra chromosome was shown in at least two cases. Tissue-dependent mosaicism is common: one patient had 3% normal cells in lymphocytes, but 70% of cultured fibroblasts had normal chromosomes.

Phenotype

The common facial features are coloboma of the iris, down-slanting palpebral fissures, preauricular tag or pit, and imperforate anus. All these features are not seen in every instance, however. Mental retardation is usually mild to moderate, but congenital heart disease or severe kidney defects can cause death in infancy.

22q11 Deletion Syndrome, DiGeorge Syndrome, and Velocardiofacial Syndrome

Patients with this common deletion (one in 4,000 liveborn infants) involving band 22q11.2 exhibit a wide phenotypic spectrum. A 22q11 deletion accounts for approximately 8% of patients with cleft palate (without cleft lip), 5% of newborns with heart defects, and 20% of familial heart defects. The phenotypic heterogeneity among 22q11 deletion patients cannot be explained by genetic heterogeneity because family members can be affected very differently. Individuals whose phenotype is consistent with the 22q11 deletion phenotypic spectrum should have a routine G-banded karyotype study and FISH analysis for the 22q11 deletion. A karyotype is important because similar phenotypes are exhibited by several less common chromosome abnormalities, such as a 4q, 8p, 10p, or 18q deletion. If an abnormality is identified, the parents should be studied because 10% to 15% of cases are familial (49). The risk of recurrence is very low if neither parent carries the deletion. The deletion maps within a 250-kb region of proximal 22q, and probably numerous genes are lost (50).

Phenotype

The phenotypic spectrum of the 22q11 deletion syndrome includes patients with features of DiGeorge syndrome, velocardiofacial syndrome (also known as Schprintzen syndrome, CHARGE association, Opitz GBBB syndrome, Robin sequence, isolated cleft palate, and isolated conotruncal heart defect) (51). The abnormalities may include conotruncal heart defects, and thymic and parathyroid hypoplasia with associated hypocalcemia, seizures, and cellular immune deficiency. Approximately 8% patients die in infancy because of heart defects or immune deficiency. Giant platelets appear to be a very common finding (52). Affected individuals often have a “pixie” face with low-set ears, a prominent nose, and hypertelorism (Fig. 30.24). Approximately 10% have renal abnormalities, and 10% to 30% have conductive hearing loss. Many of the malformations result from flawed embryonic development involving the

third and fourth branchial arches. Intelligence ranges from normal (sometimes with delayed speech development) to moderate-severe mental retardation. Behavioral problems are identified in approximately 9% of children and 15% of adults and include psychotic episodes in 1% of children and 8% of adults.

FIGURE 30.24. Three family members with 22q11 deletion syndrome (velocardiofacial syndrome). **A:** Mother; **B:** daughter at age 8 and son at age 4. Fluorescent *in situ* hybridization studies confirmed a deletion of the 22q11 critical region in all three family members. (Photographs courtesy of E. Bawle.)

Tertiary Trisomy from t(11;22)(q23;q11.2)

A European collaborative study (53) identified more than 100 families with this particular translocation. Our laboratory has identified five unrelated families, three through prenatal genetic studies and two through studies of couples with multiple miscarriages. Many families have been identified through an abnormal proband who carries an extra small chromosome (Fig. 30.7), and the risk of 3:1 meiotic segregation is 5%, with the majority of the affected individuals having 47 chromosomes with two normal 11s and 22s plus the small derived 22 from the translocation. The net result is duplication of proximal 22q and distal 11q. Many patients are mosaic, and extra care must be exercised in prenatal studies of these families to identify affected mosaics. The carrier parent is the mother in approximately 90% of cases, and many carrier males have reduced fertility. A peculiar pattern of segregation has also been observed in female carriers: nearly 70% of the normal offspring of carriers are themselves carriers, and more than half of the carrier progeny are females. This deviation from randomness is not yet understood. The pattern of segregation to offspring of male carriers appears to be normal.

Phenotype

The phenotype includes hypotonia, microcephaly, malformed ears with preauricular pits and tags, cleft or high-arched palate, micrognathia, dislocated hips, anal atresia or anal stenosis, and cryptorchidism. Two thirds have congenital heart defects (atrial or ventricular septal defect or patent ductus arteriosus). Mental retardation is moderate to severe.

Triploidy and Diploid/Triploid Mosaicism

Triploidy is identified in 7% of miscarriages and is found rarely in stillborns, with 60% being 69,XXY, 35% being 69,XXX, and only 5% being 69,XYY. Triploids with 69,XYY are less common and tend to be found in earlier miscarriages, suggesting less viability compared with XXY or XXX triploids. The liveborn incidence may be as great as one in 2,500 births, but most escape detection because of neonatal death. The mean parental age is not advanced. Probably all liveborn infants with triploidy have a mixture of diploid and triploid cells. The diagnosis of diploid/triploid mosaicism is typically made by karyotyping cultured skin fibroblast cells. This syndrome is surely underdiagnosed because the lymphocyte population is primarily diploid. It is possible that triploid lymphocytes are nonviable or do not respond to the mitogen.

Several plausible origins have been proposed for diploid/triploid mosaicism. A relatively common origin of this mosaicism may be fertilization of both the egg and a first or second polar body, one with a single sperm and the other with two sperm. Subsequently, the two zygotes develop as a single organism. This sequence of events actually results in chimerism, not mosaicism. (A chimera is an organism that results from fusion of more than one zygote.) An alternative origin is fertilization of separate ova, with postzygotic fusion during the earliest stages of embryogenesis.

Phenotype

Triploid miscarriages are often associated with cystic villi or a hydatidiform mole. The embryos often have retarded limb development and generalized growth retardation, facial malformations, open spine defects, syndactyly, and subectodermal hemorrhage. We identified several triploids in amniotic fluid cell cultures studied because of an elevated maternal serum α-fetoprotein level or anatomic malformations observed on ultrasonography. Key physical features of triploid stillborns and liveborn infants include significant intrauterine growth retardation, disproportionately small body compared with the head size, syndactyly, and ambiguous genitalia. A history of polyhydramnios or oligohydramnios and a large placenta with hydatidiform changes are common. Open spine defects are present in 25% and a few have omphalocele. Most affected males have abnormal external genitalia. Liveborn infants usually die in the neonatal period, although one patient was 21 years old at diagnosis. Liveborn infants, as might be expected, frequently also have a diploid cell line and can present with bodily and facial asymmetry. Mental retardation is variable but usually severe.

SEX CHROMOSOME SYNDROMES

Part of "30 - Clinical Cytogenetics"

As reviewed in Chapter 29, our understanding of the sex chromosomes has improved greatly over the past few years. Males have an X and a Y chromosome and females have two X chromosomes, one of which is inactivated to compensate for the single X dosage in males. Most of the genes on the X and the Y chromosomes are unique to each, but some are shared. Many of the genes that are shared by the X and the Y chromosomes are confined to the pseudoautosomal region on distal Xp and Yp and are not inactivated on the inactive X in females. Some of the genes that reside in proximal Xp and Xq also appear to have counterparts on the Y and are expressed on both the active and the inactive X, thus maintaining the equivalence of gene expression between the sexes. A small number of expressed genes on the Y chromosome, including SRY (sex-determining region of Y), appear to be responsible for male sexual differentiation. As our understanding of these various groups of X- and Y-linked genes improves, so will our understanding of their effects in the sex chromosome syndromes (54).

Monosomy X and Structural Abnormalities of the X Chromosome

The incidence of monosomy X is approximately one in 3,000 newborns. A mosaic chromosomal constitution is identified in 50% of girls with Ullrich-Turner syndrome: 15% have some 46,XX cells (45,X/46,XX mosaicism), 20% have an isochromosome Xq, 10% have mosaicism with a Y chromosome

(45,X/46,XY), and most of the remainder have X or Y rearrangements (e.g., deletions, rings, isochromosomes). Among nonmosaic 45,X cases, the nondisjunctional event results from loss of the paternal X in the majority (70% to 80%). This is a striking difference from the maternal meiotic error associated with autosomal trisomy.

Phenotype

The phenotype of Ullrich-Turner syndrome patients is variable and depends in large measure on the karyotype (Fig. 30.25). The nearly universal features are decreased birth weight and short stature (in the range of the fifth percentile on a growth curve), and gonadal dysgenesis. Common features (25% to 75% of cases) include high-arched palate, visual defects (usually strabismus), webbed or short, broad neck, low posterior hairline, shield chest, pigmented nevi, cubitus valgus, short fourth metacarpal, and thyroid disease. Lymphedema is common and is severe in abortuses and stillborns. One third of patients have bicuspid aortic valves, and 12% have other severe heart defects such as a ventricular septal defect or coarctation of the aorta. More than one half have kidney malformations (including double collecting system, malrotation, or malformation), making recurrent urinary tract infections common, although life-threatening renal malfunction is uncommon.

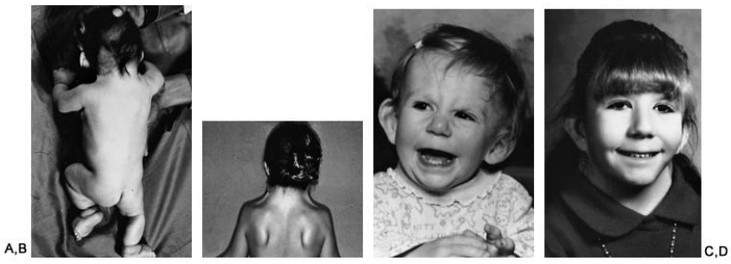


FIGURE 30.25. A: Ullrich-Turner syndrome in infancy is sometimes suspected on the basis of lymphedema involving the hands and feet and loose skin at the back of the neck. **B:** Neck webbing is shown in an older patient. The girl depicted in **C** and **D** at ages 16 months and 11 years, respectively, is now a well-adapted young adult. (Photographs are courtesy of L. Weiss and J. R. Roberson.)

In a population survey of consecutive newborns, 13 subjects with a 45,X chromosomal constitution were identified. Three of the 13 died neonatally and one died at age 11. All four of these had serious congenital heart disease. Of the nine survivors, five had multiple stigmata of Ullrich-Turner syndrome, two had pedal edema only, and two had high-arched palate only.

Most Ullrich-Turner syndrome patients have streak gonads at birth; their ovaries appear normal at 12 weeks of gestation, but the ovarian follicles degenerate after that time. One fifth of Ullrich-Turner syndrome patients have spontaneous menstruation, and most of these are 45,X/46,XX mosaics. The available data on their pregnancy outcomes suggest a high risk for miscarriage and 45,X liveborn infants, and possibly an increased risk of trisomy 21.

Mosaic individuals have a better prognosis, including lower risk of Ullrich-Turner syndrome stigmata and lower risk for gonadal dysgenesis, whereas 45,X without obvious mosaicism is more often associated with Ullrich-Turner syndrome. Ring X and isochromosome X patients usually have features of Ullrich-Turner syndrome, but these patients typically have a 45,X cell line as well. Many females with Xq deletions are less severely affected. A few patients with Xq deletion have been fertile, and the chance of fertility may depend on the breakpoint or perhaps to a greater extent on whether a 45,X or 46,XX cell line mosaicism is present. Those with Xp deletions often have features of Ullrich-Turner syndrome, including short stature and primary or secondary amenorrhea, although a few have been fertile.

The incidence of mental retardation among patients with Ullrich-Turner syndrome is not dramatically increased, and the majority have normal intelligence. They are more likely to have problems with spatial relations and numerical identification. There is often a deficiency in perceptual motor organization or in fine motor execution. Therefore, the nonverbal IQ is lower than the verbal IQ. Personality traits commonly include inertia to emotional arousal, a high capacity to deal with stress, and strong traditional femininity. Mosaicism with a very small ring X chromosome has a high risk of mental retardation, and this is usually owing to lack of normal inactivation of the ring chromosome (55). Some females with X;autosome translocations have mental retardation or express an X-linked mendelian disorder owing to interruption of gene function at the translocation breakpoint.

45,X/46,XY Mosaicism

The phenotype of 45,X/46,XY mosaics is extremely varied, ranging from typical Ullrich-Turner syndrome to ambiguous genitalia

to normal male. Females with a 45,X/46,XY karyotype or who have a structurally abnormal Y-derived chromosome have a 15% to 25% risk of gonadoblastoma. Their gonads should be removed before school age because of this risk.

Many phenotypically normal males with 45,X/46,XY mosaicism have been identified in prenatal studies, but their fertility has not been evaluated because they have not yet reached reproductive age. However, this karyotype does not appear to be over-represented in studies of male infertility nor is there an increased risk of gonadoblastoma in phenotypically normal males who have 45,X/46,XY mosaicism.

46,XY Females

Females with nonmosaic 46,XY typically have gonadal dysgenesis and little secondary sexual development but few other features of Ullrich-Turner syndrome. There is a high risk of gonadoblastoma. Most 46,XY females have a structurally normal X and Y chromosomes and have a mendelian inherited disorder of sexual differentiation, such as testicular feminization syndrome (X-linked recessive) or 5α-reductase deficiency (autosomal recessive). As mendelian traits, such disorders carry a significant risk of recurrence (56).

A few 46,XY females were identified as having del(Y)(p11) who are probably affected because their Y chromosome has lost the SRY. Females with a Y chromosome deletion are more likely to exhibit some of the features of Ullrich-Turner syndrome, although most appear to have normal stature. As with the nondeletion cases, there is a high risk of gonadoblastoma. In distinction to the nondeletion cases, the Y deletions are not known to have an increased risk of recurrence in a family. Molecular genetic studies showed that Yp-specific DNA probes did not hybridize with the deleted Y chromosome. The deletion breakpoints differed slightly, but the deletions overlapped in the sex-determining region. One patient had a more complex deletion involving two noncontiguous DNA segments. The normal stature and the gonadoblastoma support the hypothesis that proximal Yp or the long arm of the Y contains some of the genes for stature and spermatogenesis.

Triple X Syndrome 47,XXX

The incidence of 47,XXX is approximately one in 1,000 liveborn females. There is a maternal age effect, with maternal meiosis I errors in 50% of cases; paternal errors account for fewer than 10% of cases. Triple X women do not appear to have an increased risk for having XXX or XXY offspring. The phenotype typically is within normal limits, with development of secondary sexual characteristics and normal fertility. Many but not all triple X females are developmentally normal, although the exact risk of mild or more severe mental retardation is uncertain. Among females with mild mental retardation, the frequency of triple X is increased to approximately one in 200. Ten 47,XXX females who were identified in unselected newborn surveys have been followed to age 20 to 22 (57). As children, they had less motor coordination and more speech and language problems than their peers, but standard speech therapy seems to have been sufficient for most of them. One was in a school program for the educationally mentally impaired, three were mainstreamed with some remedial education, and six were in regular classes. Four of the 10 have an IQ of 60 to 80, so in this small, unselected sample 40% had mild mental retardation. At age 20 to 22 years, nine of the 10 appear to cope well in society and hold a job or have a family.

Klinefelter's Syndrome 47,XXY

The incidence of 47,XXY is approximately one in 700 male newborns. Because of the maternal age effect in 47,XXX and 47,XXY, these karyotypes are seen more frequently in prenatal genetic studies. For 47,XXY, maternal and paternal meiotic errors are responsible in approximately equal proportions.

Phenotype

The phenotype of Klinefelter's syndrome includes tall stature, small testicles and prostate, and infertility. Half have a eunuchoid habitus and half have gynecomastia. A few have a varicocele, undescended testes, or a small penis.

Because of the relatively mild phenotype, most subjects with Klinefelter's syndrome are never identified. The diagnosis is usually made during adolescence or adulthood because of small testicles, gynecomastia, or infertility. Mosaics, mostly 47,XXY/46,XY, constitute approximately 10% of Klinefelter patients, and as a group have milder symptoms. Most have azoospermia and most of those who have oligospermia are chromosomal mosaics. As children, many have mild speech and language deficits, which respond to standard speech therapy. They also tend to have a poor attention span and memory and low self-esteem. Early educational intervention appears to be effective in overcoming these potential liabilities. Twelve men with Klinefelter's syndrome who were identified in unselected newborn surveys have been followed to age 20 to 22 (57). Eight had a normal IQ and have coped well in society. Four had IQs of 80 to 90 and had learning problems in school. Two of these men were said to cope well as adults and have found employment. The other two had psychiatric problems as well as learning difficulties and were said not to cope well. Nevertheless, the XXY men as a group were said to be not as severely affected as the XXX women.

More severely affected variants of XXY include 48,XXYY; 49,XXXYY; 49,XXXXY, and other karyotypes. Their cumulative incidence is probably approximately one in 2,500 males. Hypogonadism is present, and cryptorchidism is more common. Adult height is more than 6 feet (180 cm) in 80% of variants, and skeletal abnormalities are not uncommon. Mental retardation is more likely than in 47,XXY.

46,XX Males

The incidence of 46,XX males is approximately one in 20,000 liveborn males. These are phenotypic males who have normal or slightly shorter stature and normal external genitalia but who have reduced testosterone levels, reduced facial and body hair, dysgenetic testes, and azoospermia, and may have gynecomastia. There is strong cytogenetic and molecular genetic evidence favoring an aberrant paternal meiotic exchange between the X and

Y chromosomes, such as unequal meiotic exchange, resulting in an X;Y rearrangement. Some of these are cytogenetically visible. Many other XX males have molecular evidence of a rearrangement between the X and Y that results in the SRY being moved to one X chromosome, i.e., Yp-specific DNA is present. A minority of XX males (10%) do not have evidence of Yp-specific DNA, and their phenotype includes gynecomastia and hypospadias or a small penis.

47,XYY Males

The frequency of XYY is approximately one in 800 males, with no maternal age effect because the nondisjunction occurs during male meiosis or as a postzygotic error. As is true for XXX and XXY, most individuals with an XYY constitution are never identified. Stature is taller than average; one in 200 males taller than 6 feet (180 cm) have an XYY constitution. Fertility appears to be normal in most but not all cases because approximately 1% of men with oligospermia have XYY. The personality is described as impulsive in nature but variable in how that is expressed. Some XYY males are overtly aggressive and others are timid and have autistic features. For reasons that are not entirely clear, approximately one in 200 prisoners have 47,XYY. As children, XYY boys tend to have a high activity level, poor emotional control, and a weak self-concept and tend to be clumsy and easily frustrated. There may be a higher frequency of XYY boys with severe behavior disorders, and on average, the IQ appears to be slightly lower. However, most XYY males have intelligence and personality traits within normal limits. The frequency of tall stature, lower IQ, and criminal records is similar in XYY and XXY males, so the phenotype cannot be attributed directly to the extra Y chromosome. Early-intervention education programs may be useful in 47,XYY as well as in 47,XXX and 47,XXY.

INDICATIONS FOR CHROMOSOME STUDIES (SUBPOPULATIONS)

Part of "30 - Clinical Cytogenetics"

Miscarriages, Stillbirths, and Liveborn Infants

Miscarriage occurs in approximately 15% of all recognized pregnancies, with stillbirth representing another 1%. A chromosome abnormality is observed in 11% of stillbirths and 5% of perinatal deaths, compared with less than 1% of liveborn infants. Although the frequency of chromosome abnormalities is highest in macerated or malformed fetuses, many anatomic defects are not obvious, particularly those associated with trisomy 18, which constitutes approximately 30% of chromosome defects among stillborns. Therefore, it is appropriate to karyotype all stillbirths as part of a formal protocol that should also include family and obstetrical history, photographs of the face and any malformations, a whole-body radiograph, bacterial cultures, and autopsy. When malformation or a chromosome abnormality is identified, genetic counseling is warranted. The resources are unavailable to karyotype all miscarriage specimens, but for couples with two or more pregnancy losses, it is appropriate to carry out chromosome studies of the abortus as well as peripheral blood karyotype studies of the couple.

Approximately 50% of first-trimester, 25% of second-trimester, and 11% of third-trimester miscarriages have a chromosome abnormality (Table 30.9). For the earliest recognized miscarriages (less than 8 weeks from last menstrual period), however, there is some evidence that the frequency of chromosome abnormalities may be as low as 10%, although this may be an underestimate because 20% of normal-appearing *in vitro* fertilization embryos have a chromosome abnormality. Ectopic pregnancies do not have a substantially increased frequency of chromosome abnormalities.

TABLE 30.9. INCIDENCE OF CHROMOSOME ABNORMALITIES IN MISCARRIAGES AND STILLBORNS COMPARED TO NEWBORNS*a*

	Miscarriages[b] (%)	Stillborns (%)	Newborns (%)
Normal karyotype	50	95	99.3
Abnormal karyotype	50	5	0.7
Tetraploid	2.5	<<1%	-
Triploid	7	0.2	-
Trisomy (autosomal)	30	3.0	0.14
+13	1.5	0.5	0.01
+15	2.3	-	-
+16	9.9	-	-
+18	1.8	1.4	0.01
+21	2.6	0.8	0.11
+22	2.8	-	-
All other	9.1	0.3	0.01
Monosomy X	8.6	0.05	0.01
XXX, XXY, XYY, variants	0.6	0.5	0.2
Other	3	1.4	0.4

[a] Statistics compiled from Angell et al., (58), Simpson and Bombard (59), and Hook and Hamerton (6).

[b] Mostly first-trimester miscarriages.

Polyploidy is observed in 10% of abortus specimens, with 3% being tetraploid and 7% being triploid. These are often associated with an empty sac or a severely disorganized embryo, but mosaics are often less severely affected and some liveborn infants are known with polyploidy mosaicism.

Hydatidiform degeneration of the placenta is frequent in triploid conceptions, and although some hydatidiform moles are triploid, the majority of complete moles (which carry a risk of malignant degeneration) are diploid with two paternal and no maternal contribution to the zygote. Diploid moles appear to arise from fertilization of an empty (degenerating) egg by a haploid sperm with subsequent duplication of the paternal chromosomes or by dispermy (fertilization by two sperm). The absence of embryonic development in diploid moles is consistent with the hypothesis that a paternal chromosome set is required to form the placenta and a maternal set required to form the embryo.

Trisomy for the autosomes (chromosomes 1 to 22) as a group comprises approximately 30% of all miscarriages (Table 30.2). By far the most common is trisomy 16, which accounts for approximately one third of cases. Trisomy 16 is frequently associated with an empty sac and hydatidiform degeneration of the placenta. Trisomy 1 has not yet been seen in a miscarriage specimen. As with liveborn trisomy, there is a maternal age effect, and most trisomic miscarriages result from nondisjunction during maternal meiosis I.

In general, women who have one or more miscarriages have an increased risk of miscarriage in future pregnancies compared with women who have no history of miscarriage (25% versus 12%). However, this information is based on unkaryotyped miscarriages, so it is unclear whether there is a maternal age-independent increased risk of pregnancy loss subsequent to a trisomic miscarriage.

After a couple experiences a trisomic miscarriage, there is an increased likelihood that any subsequent miscarriage will also be trisomic, although not necessarily for the same chromosome (Table 30.10). Some centers recommend prenatal diagnosis in future pregnancies for women who have had a trisomic miscarriage, but there is only weak evidence that such women have a maternal age-independent increased risk for having a liveborn trisomic.

Monosomy for an autosome is extremely rare, and most monosomies have never been observed. One plausible reason is that loss of a chromosome or chromosome segment has a much more severe effect on phenotype than does gain of the same chromosome or chromosome segment. In contrast, monosomy X accounts for approximately 9% of miscarriages.

Although 45,X is compatible with full-term development, the

TABLE 30.10. FREQUENCY OF RECURRENT TRISOMIC AND CHROMOSOMALLY NORMAL MISCARRIAGES AMONG 125 PAIRS OF ABORTUSES*a*

	Karyotype of Second Loss		
Karyotype of First Loss	**Normal**	**Trisomic**	**Other**
Normal	55	6	7
Trisomic	8	24	2
Other	9	8	6

[a] Adapted from Hassold TJ. A cytogenetic study of repeated spontaneous abortion. *Am J Hum Genet* 1980; 32:723–730.

great majority do not survive to term. In one study, only 1% to 6% of 45,X miscarriages had cytogenetic or molecular evidence of mosaicism, whereas at least 50% of liveborn infants are mosaics.

Other chromosome abnormalities seen in miscarriages and stillborns include unbalanced rearrangements, and the balanced translocation or inversion is sometimes identified in a parent.

The reasons for miscarriage or stillbirth in chromosomally normal conceptions are legion, but there are major categories that can be explored in women who have more than one chromosomally normal miscarriage. Smoking, alcohol, cocaine, and infections are major environmental factors in chromosomally normal miscarriages. Conversely, the identification of a chromosome abnormality in an abortus is generally sufficient to exclude a teratogenic or other nonchromosomal cause. This can be useful to some couples who are concerned about environmental exposures or other factors as a cause for their miscarriage. If one of two miscarriages is aneuploid and the other euploid, it is also safe to conclude that the multiple miscarriages are not causally related.

Multiple Miscarriages

Because 15% of recognized pregnancies end in miscarriage or stillbirth, simple probability dictates that 2% of couples with two pregnancies will have two miscarriages, and 6% of couples with three pregnancies will have two miscarriages. Nonetheless, in 4% of couples (2% of subjects) who experience multiple pregnancy losses (more than one miscarriage or stillbirth), one member of the couple carries a balanced chromosome rearrangement. By comparison, the incidence of balanced chromosome rearrangements in the general population is only 0.6% (Table 30.5). This increased frequency of rearrangements justifies a chromosome study in all couples who experience more than one pregnancy loss.

The rearrangements found in couples with multiple miscarriages include reciprocal translocations (50%), robertsonian translocations [35%, mostly der(13;14)], inversions (13%, mostly pericentric), and 2% others (e.g., X deletion or normal/ring mosaic). The majority (75%) of the identified carriers are female. The nature of the chromosome rearrangement in a specific family influences the risk of miscarriage or malformed liveborn, but it is reasonable to suggest family studies and offer prenatal diagnosis to the carriers. Couples with three or more consecutive miscarriages may have a lower frequency of rearrangements as obstetric, immunologic, and endocrinologic risk factors become more likely.

Many women who have multiple miscarriages have a small proportion of 45,X or 47,XXX cells in their peripheral blood chromosome preparations. This finding is age related and does not indicate an increased risk for subsequent miscarriages or conceptions with an abnormal karyotype beyond the patient's age-related

risk. The frequencies of 47,XXX and 47,XYY constitutions are probably no greater in multiple miscarriage subjects than in the general population. There have been suggestions of an association between miscarriages and chromosome variants, especially 9qh and Yqh, but the evidence favoring this association is weak.

Male Infertility

A chromosome abnormality is detected in 15% of men with azoospermia and in 6% with low sperm counts (less than 10 million/mL) (Table 30.11). Among men with sperm counts ranging from 10 to 20 million, approximately 2% have a chromosome abnormality. Men with a 47,XXY karyotype are in the majority in the azoospermic group, but balanced rearrangements are more important among the oligospermic men. 47,XYY is not overrepresented among azoospermic men but is more frequent in oligospermia.

TABLE 30.11. CHROMOSOMES IN AZOOSPERMIA AND OLIGOSPERMIA*a*

Azoospermia: 15% abnormal (73 of 489 men)
56 47,XXY
5 47,XYY mosaics
1 47,XYY mosaic
5 46,XX males
1 45, X/46,XY mosaic (30:70 ratio)
5 Balanced rearrangements
Oligospermia (less than 10 million/mL sperm count): 6% abnormal (58 of 959 men)
7 47,XXY
4 47,XXY mosaics
7 47,XYY
1 47,XYY mosaic
32 Balanced rearrangements
1 46,X, del (Y)(q12) new mutation
6 46,XY/47,XY, +marker

[a] Data from Retief et al. (61) and Bourrouillou et al. (62).

A 45,X/46,XY mosaicism does not appear to be a significant risk factor for infertility. Most prenatally diagnosed cases have normal-appearing male external genitalia at birth. Males with obstructive infertility (e.g., congenital absence of the vas deferens) should be tested for mutations at the CF gene locus. Many males with nonobstructive infertility have a submicroscopic deletion that spans the *DAZ* gene locus on Yq. These possibilities need to be considered before intracytoplasmic sperm injection (63).

Female Infertility

A chromosome abnormality is detected in 25% of primary amenorrhea and gonadal dysgenesis patients, with approximately 10% numerical and 10% structural changes of the X, and 5% 46,XY or structural changes involving the Y chromosome. In one study, 33% of patients with secondary amenorrhea had an X abnormality, and in our experience, 10% of women karyotyped because of premature menopause have an isochromosome Xq or other structural abnormality of the X chromosome. Some women with premature ovarian failure are fragile X carriers (64). There are many genetic but nonchromosomal causes of female infertility (56).

Ambiguous Genitalia

Ambiguous genitalia in a newborn becomes a social emergency for the family, whether it appears as an isolated birth defect or as part of a syndrome. Abnormalities of the external genitalia can be associated findings in several sex chromosome abnormalities. For an infant with hypotonia as well as hypogonadism, Down's syndrome and PWS should be considered, and high-resolution cytogenetic studies should be done. The biochemistry of abnormal sexual differentiation in chromosomally normal 46,XX and 46,XY newborns and the many nonchromosomal causes of ambiguous genitalia are beyond the scope of this chapter (65,66,67 and 68).

The karyotype is essential and should be done as a stat specimen. It is standard practice to order serum 17-hydroxyprogesterone and testosterone measurements, either at the same time or depending on the cytogenetic result. Because many hospitals send these tests to reference laboratories, the urgency of the karyotype result is accentuated. Until the question of gender is answered, anxiety is tremendous. The physicians and other medical staff generally prefer to avoid mention of gender, and use the ambiguous term gonads rather than testes or ovaries.

Mental Retardation and Multiple Congenital Anomalies

Estimates of the frequency of chromosome abnormalities among the mentally retarded vary widely, but we find that the following estimates have been workable. Among the moderately to severely retarded, Down's syndrome is present in 10% to 15%, and 3% to 5% have other chromosome abnormalities, including other trisomies, and balanced and unbalanced structural changes. The fragile X syndrome can be identified in 3% to 6% of males and in 3% to 4% of females whose IQ is between 55 and 75 (69,70). The frequencies of *de novo* balanced reciprocal translocations and inversions are increased, but inherited balanced rearrangements and *de novo* robertsonian translocations are not more frequent among the mentally retarded. The likelihood of finding a chromosome abnormality is greater if multiple minor or major malformations are present. Of patients with congenital heart disease, 13% have a chromosome defect, mainly Down's syndrome, and to a lesser extent trisomy 8, 13, 18, monosomy X, and unbalanced chromosome rearrangements.

Tissue-limited mosaicism is present in 3% to 10% of subjects with multiple malformations, mental retardation, a normal blood karyotype, and no clear nonchromosomal syndrome.

Pigmentary Dysplasia

Several recent studies have drawn attention to an association between swirls and lines of hyperpigmentation or hypopigmentation of the skin and chromosomal mosaicism (71). In some patients, the pigmentary changes can be seen better using an ultraviolet lamp. The terms *hypomelanosis of Ito* and *incontinentia pigmenti chromiens* have been used to describe these pigmentary abnormalities, which traverse the trunk in a generally horizontal

pattern and linearly along the arms and legs. The lines may reflect migration of melanocytes from the neural crest region during early embryogenesis. The associated chromosome abnormalities include XX/XY chimerism, diploid/triploid mosaicism, secondary trisomy 12p, and a variety of rings with normal/ring mosaicism, including patients who have some features of Ullrich-Turner syndrome, seizures, mental retardation, and a small ring X. Several X;autosome translocation cases are represented in this group, and the X breakpoint is often near the locus of a hypothesized gene for incontinentia pigmenti, although the patients do not have classic incontinentia pigmenti.

In patients with pigmentary dysplasia and mental retardation, seizures, or other birth defects, chromosome studies should be extended to cultured fibroblasts if the blood karyotype is normal, and biopsies should be taken from skin with each pigmentary expression or from the borders between areas of different pigment.

Arthrogryposis (Congenital Contractures)

Of patients with arthrogryposis and delayed development or mental retardation, approximately 15% have a karyotype abnormality, and half of these are mosaics (72). Trisomy 8 mosaicism is perhaps most common, but contractures have been associated with a variety of other chromosome abnormalities. A blood karyotype should be studied if multiple joint contractures are associated with developmental delay, and especially if other birth defects are present. If the blood karyotype is normal, a cultured skin fibroblast chromosome study should be studied.

METHODS

Part of "30 - Clinical Cytogenetics"

Chromosome preparations for cytogenetic analysis can be obtained from many tissue sources. Peripheral blood is the most frequently used specimen for routine chromosome analysis, as it is the easiest tissue to obtain. Blood karyotypes are usually ordered because of the presence of birth defects, mental retardation, infertility, or multiple miscarriages. Lymphocytes are incubated under appropriate conditions, during which blast transformation is induced by one of several lectins, such as PHA. Cell harvest begins with the addition of an agent (e.g., Colcemid or vinblastine) to arrest cell division at the metaphase stage. Other treatments may be employed to inhibit chromosome condensation or prepare the cells for specialized banding patterns such as red-blue-green (RBG) staining (Chapter 29). The cultures are then treated with a hypotonic solution to swell the cells and allow dispersion of the chromosomes within the cell membrane. Fixative is used as a wash solution to lyse the remaining red cells and remove some chromosomal proteins. Microscope slides are prepared and treated with one of several staining or banding methods. Chromosome preparations are then analyzed through the microscope for numerical and structural aberrations, and findings are documented with photographs or computer-generated images.

Chromosome analysis in newborns requires special consideration because a smaller volume of specimen is usually received and decisions regarding life-support systems or surgical intervention require rapid or stat analysis and interpretation. Direct preparations can be obtained from bone marrow cells harvested from specimens immediately after being received in the laboratory, although the quality is inconsistent. Whole blood specimens or bone marrow aspirates can be cultured as briefly as overnight. Thus, analysis of numerical and major structural changes can be performed within 6 hours, but normal results from marginal preparations must be confirmed with better material. This technique or a 48-hour method can be employed for prenatal percutaneous umbilical blood samples.

Fibroblast monolayer cultures can be established from miscarriage material to help identify the cause of the loss and from skin biopsies to look for sex-chromosome mosaicism or tissue-limited mosaicism. Cultures are established using small tissue pieces (explants) or dissociated cell suspensions obtained by treatment of tissue pieces with enzymes that break collagen fibers in tissues. Harvest techniques follow the basic principles used in harvesting lymphocytes.

The use of other tissues for chromosome analysis has been described in the literature but is not routinely done. Culture of urine sediment cells from newborns has been employed to confirm a prenatally detected trisomy mosaicism. Sperm chromosome analysis is being used by a few laboratories to detect germ cell mosaicism.

Many methods for cell culture, banding techniques, and differing protocols for cell analysis have been published. Each institution must modify procedures to suit their laboratory conditions and needs. The methods described here are those currently used in our laboratory and meet the requirements of accrediting agencies.

Cytogenetic Procedure for Peripheral Blood

Reagents

Blood culture media: 500 mL RPMI 1640 (GIBCO, Rockville, MD) with 100 mL fetal bovine serum, 6.5 mL penicillin-streptomycin, 7 mL L-glutamine. Dispense 10 mL aliquots into sterile T-25 flasks and add 0.33 mL PHA to each flask. Store at 4°C for as long as 2 weeks.

Fetal bovine serum (Hyclone, Logan, UT).

Penn-Strep (GIBCO): 10,000 IU/mL penicillin and 10,000 µg/mL streptomycin.

L-Glutamine (GIBCO): 200 mmol/L.

PHA (GIBCO).

Ethidium bromide (BioRad, Hercules, CA): 1 mg/mL. Warning: Mutagenic.

Colcemid (GIBCO): 10 µg/mL.

Acetic-methanol fixative: 3 parts anhydrous methanol: 1 part glacial acetic acid. Prepare fresh for each use.

Hypotonic: 0.075 mol/L KCl (5.59 g/L).

Microscope slides can be washed in a sonicator using methanol. Sonicate for 10 to 15 minutes, rinse in distilled water, and store at 4°C in distilled water.

Except where noted, all centrifugation steps are performed under the same conditions: 10 minutes at approximately 165*g* (e.g., 900 rpm using a 175-mm radius head).

Culture

This method is used to obtain routine mid to early metaphase preparations. To obtain prometaphase chromosomes for high-resolution analysis, a modification of the method is provided in Chapter 29.

1. Obtain 10 mL heparinized blood (usually drawn in a Vacutainer) and allow the blood to settle at room temperature. For small volumes of blood, cultures can be established using whole blood.
2. With a sterile pipette, collect the buffy coat layer between the plasma and red blood cells and transfer to a sterile tube. Obtain 1.5 mL plasma and white cells and mix well in the tube.
3. Using sterile technique, establish two cultures per patient by inoculating blood culture medium with 0.75 mL buffy coat suspension or 0.5 mL whole blood.
4. Incubate in 5% CO_2 at 37°C.

Harvest

5. On the third day, after 66 to 71 hours of incubation, add 0.1 mL ethidium bromide and 0.1 mL Colcemid to each flask. Return flasks to the incubator for 2 hours. *Variation*: Rush or stat cultures can be harvested after 48 hours of incubation.
6. Gently swirl the culture to loosen cells that have adhered to the flask and transfer culture to a conical centrifuge tube.
7. Centrifuge for 5 minutes at 1,400 rpm and aspirate the supernatant.
8. Add 10 mL of prewarmed (37°C) hypotonic solution and mix gently with a pipette to a smooth suspension. Incubate in 37°C water bath for 20 minutes, resuspend cells at 5-minute intervals.
9. Add 2 mL of freshly prepared cold (4°C) fixative and mix well with pipette.
10. Centrifuge and aspirate the supernatant.
11. Resuspend cell pellet in residual supernatant (approximately 0.5 mL) using a pipette. Add 10 mL fresh cold fixative, mix well, and allow to stand for 20 minutes at room temperature.
12. Centrifuge and aspirate the supernatant.
13. Continue the wash procedure with fresh fixative until the cell pellet is clean and white. After the final wash, resuspend the cells in fresh fixative to form a cloudy suspension. The cells are now ready to be dropped onto slides. If the cells are to be stored, add fresh fixative but do not resuspend the pellet. Store in a tightly capped tube at 4°C.

Slide Preparation

14. If the pellet was stored, rewash cell pellet with fixative and resuspend cells with fixative to produce a slightly cloudy suspension.
15. Using a glass pipette, place two to four drops of the cell suspension on a cold, precleaned, wet slide. Blow, drain, or shake the excess fluid from the slide and dry on 65°C warming plate.
16. Examine slides with phase optics and adjust dropping technique for future slides. Ideally, metaphases should be black and chromosomes well spread and nonrefractory. Changes in room humidity, height at which the suspension is dropped, and steaming of slides over boiling water will affect the spreading. Generally, four to six slides per patient are adequate to complete the analysis.
17. Place slides in 90°C drying oven for 60 minutes.
18. Allow slides to come to room temperature before banding and staining. (See Chapter 29 for chromosome banding methods.)

Cytogenetic Procedure for Tissue Specimens and Skin Biopsies

Reagents

Ham's F-10 Complete medium: 100 mL bottle (GIBCO) with 25 mL fetal bovine serum, 0.25 mL gentamicin, and 1.5 mL L-glutamine. Store at 4°C for as long as 2 weeks. Complete F-10 can also be used as a transport medium.

Minimum Essential Medium-Alpha Complete Medium (MEM-alpha): 100 mL bottle (GIBCO) with 25 mL fetal bovine serum, 0.25 mL gentamicin, and 2.0 mL L-glutamine. Store at 4°C for as long as 2 weeks.

Chang's Complete Medium: 100 mL bottle (Irvine, Santa Ana, CA) with 14 mL Supplement C, 0.25 mL gentamicin, and 2.0 mL L-glutamine. Store at 4°C for 3 days.

Gentamicin (GIBCO): 10 mg/mL.

L-Glutamine (GIBCO): 200 mmol/L.

Nystatin (GIBCO): 10,000 IU/mL.

Penn-Strep (GIBCO): 10,000 IU/mL penicillin; 10,000 µg/mL streptomycin.

Fetal bovine serum (Hyclone, Logan, UT).

Hanks' buffered saline solution (HBSS) (GIBCO).

1×: 100 mL bottle with 0.25 mL gentamicin.

10×: 100 mL bottle with 2.5 mL gentamicin.

Colcemid (GIBCO): 10 µg/mL.

Hypotonic: 0.4% KCl (0.054 mol/L KCl).

Acetic-methanol fixative: 3 parts anhydrous methanol: 1 part glacial acetic acid. Prepare fresh for each use.

Trypsin solution (GIBCO): 0.25% 1× with ethylenediaminetetraacetic acid (EDTA)

Collagenase 10× stock solution: 100 mg collagenase V (Sigma, St. Louis, MO) in 10 mL HBSS. Filter sterilize and dilute with F-10 medium to make up working solutions of 1×, 2×, and 5×. Store at 4°C.

Culture

Cell cultures are initiated in a laminar flow hood using sterile technique, and cultures are incubated at 37°C in a 5% CO_2 atmosphere. Methods vary somewhat for fetal tissues, placental villi, and skin biopsies.

Fetal Tissues

1. Tissue samples are delivered to the laboratory in a sterile container with normal saline, Ringer's solution, or transport medium. Various sources of tissues from miscarriage material can be used to establish cultures. Fetal villi, membrane, cord, or skin are the most likely to produce fibroblast cultures. Frozen tissue samples or those exposed to formalin or formaldehyde will not be viable.
2. Using sterile forceps, place each piece of tissue in a large Petri dish containing 10× HBSS. Transfer pieces to another Petri dish containing 1× HBSS and gently swirl to rinse.
3. Sort each different tissue type into a separate dry dish and keep moist with a small volume of HBSS.
4. Trim away any necrotic areas, cartilage, decidua, or blood clots. Limit the total size of each sample to 5 mm^3 or smaller to ensure adequate enzyme dissociation. Because some tissue types dissociate poorly, such as umbilical cord, they should be cut into smaller pieces.
5. By “scissoring” two scalpels blade to blade, mince each tissue piece into cuttings of 1 mm^3.
6. Using fine sterile forceps, transfer cuttings to a sterile tube containing 5 mL of prewarmed working enzyme solution and incubate the tissue in the appropriate enzyme solution at 37°C in a dry bath as follows: for early gestation products of conception or macerated fetal parts use 1× collagenase for 1 to 3 hours; for foreskin, late-gestation cord, skin, tendon, artery, and other fibrous tissues, use 2× or 3× collagenase for 2 to 4 hours.
7. After enzyme treatment, pipette the suspension vigorously using a 5-mL pipette to triturate.
8. Centrifuge and aspirate the supernatant.
9. Resuspend the pellet in 5 mL of prewarmed complete F-10 medium or MEM-alpha.
10. Centrifuge and aspirate the supernatant.
11. Resuspend the cell pellet in an appropriate volume of fresh complete MEM-alpha. Usually 5 mL is used to establish three T-25 flask cultures.
12. Set the tube upright for 10 to 30 seconds to allow the larger tissue fragments to settle to the bottom. Establish *in situ* cultures by aliquotting 0.5 mL of the cell suspension of mostly single cells and small cell aggregates onto four to six sterile 22-mm coverslips in 35-mm culture dishes. Establish one culture with 1–2 mL each of the bottom suspension including the larger tissue fragments. Incubate cultures at 37°C.

The cultures established from the small cell aggregates can be fed after 1 day, but the explant cultures should not be disturbed until the explants are firmly attached to the flask, usually 2 to 3 days. If little dissociation is observed after short-term enzyme treatment, establish the single-cell suspension in one flask and transfer the larger fragments to fresh 1× collagenase solution and incubate overnight. Rinse and establish another culture the next day.

Placental Villi

This is a variation of the chorionic villus sample method for prenatal diagnosis, detailed in Chapter 31.

1. After the villi have been rinsed in 10× and 1× HBSS as with other tissues, examine carefully through a dissecting microscope and discard fragments of maternal decidua.
2. Carefully remove 10 to 20 mg of the most viable-appearing villi and place in a 60-mm sterile culture dish.
3. Add 5 mL of 1× trypsin-EDTA solution and incubate at 37°C for approximately 90 minutes. Macerated villi often digest more quickly and should be monitored.
4. After trypsin incubation, gently swirl the softened villi using a pair of sterile forceps and transfer the entire suspension to a sterile tube containing 5 mL 1× collagenase working solution.
5. Incubate at 37°C for 1.5 hours and establish cultures as previously described.

Skin Biopsies

Because most skin biopsies are obtained under aseptic conditions, it is not necessary to rinse in HBSS.

1. Suspend the biopsy in the transport medium with gentle agitation and carefully pour it into a 100-mm sterile culture dish.
2. Transfer the tissue and a few drops of fluid to a dry dish.
3. Obtain four to six 1-mm^3 pieces by scissoring two scalpels blade to blade through the specimen.
4. Transfer one or two fragments into each of three or four flasks. Gently create a puddle with 0.5 mL of fresh MEM-alpha over each explant and incubate undisturbed for 2 to 3 days or until growth is observed. Usually epithelial growth is first noted, followed by fibroblast growth.

Feeding

The cells in the suspension cultures usually attach after 1 day and can be fed at that time. *In situ* cultures are fed with 1.5 mL of prewarmed MEM-alpha every 4 days until harvest. To feed cultures, aspirate media carefully (do not dislodge cells) and add 3 mL of prewarmed MEM-alpha. Explant flasks need to be handled with care until fibroblast growth is well established. When initial epithelial growth is noted (3 to 5 days), add approximately 1 mL of fresh Ham's F-10 media and leave undisturbed for another 5 to 7 days.

If the cultures become contaminated with yeast or fungus, aspirate the medium and rinse the cultures with 1× HBSS. Add to the flask 3 mL of fresh media and 0.01 mL Nystatin. Return the cultures to the incubator and monitor carefully.

Subculture of Tissue Specimens

Subculture of a flask or coverslip may be necessary if additional cells are required or if the culture becomes confluent.

For coverslip subculture, follow the protocol described in Chapter 31 for subculture of amniotic fluid cell cultures. Use the following protocol for T-25 flask subculture:

1. Rinse each flask with 3 mL 1× HBSS.
2. Add 2 mL trypsin-EDTA solution to each flask and incubate at 37°C for approximately 5 minutes. Monitor cell release using an inverted microscope.

3. Gently tap the base of the flask against a large rubber stopper to dislodge the cells and obtain a single-cell suspension.
4. Add 2 to 4 mL of complete medium to inactivate the trypsin.
5. Establish four to six coverslips and retain the rest of the cell suspension in the original flask.
6. Incubate cultures at 37°C and maintain subcultured flasks until harvest.

Harvest

For harvest of *in situ* cultures, follow the protocol described in Chapter 31 for amniotic fluid cell *in situ* cultures. For flask cultures, use the following protocol. Sterile technique is not required unless the culture is to be carried for additional harvests. Cultures are examined using an inverted microscope. Cultures are ready for harvest when they are approximately 60% confluent. Slower growing cultures may give a better yield of mitotic cells if harvested at 80% confluency.

1. Add Colcemid, 0.01 mL per milliliter of medium in flask.
2. Return the culture to the incubator for 2.5 hours or until rounded up (mitotic) cells are evident.
3. Pour medium from the flask into a 15-mL centrifuge tube.
4. Add 2 mL 1× HBSS to the flask, swirl gently, and decant into tube containing the medium.
5. Add 2 mL 1× trypsin-EDTA solution to the flask. Incubate at 37°C for 5 minutes.
6. When most of the cells are free floating, gently tap the base of the flask against a large rubber stopper to loosen any cells still adhered to the flask.
7. Add 2 mL of F-10 medium to each flask to inactivate the trypsin.
8. Gently swirl and immediately decant into the tube. If the culture is to be maintained, add 3 mL of fresh media and return the flask to the incubator.
9. Centrifuge and aspirate the supernatant.
10. Add approximately 8 mL of prewarmed 0.04% KCl hypotonic solution and mix well. Incubate in 37°C water bath for 15 minutes.
11. Add 4 mL of freshly prepared cold (4°C) fixative and mix well with a pipette.
12. Centrifuge and aspirate the supernatant.
13. Resuspend the cell button by flicking the bottom of the tube with a finger. Add 4 mL of fresh cold fixative and mix well with a pipette.
14. Centrifuge and aspirate the supernatant.
15. After the final wash, the cells are ready to be dropped onto slides, or the pellets can be stored as a cell-fixative suspension in a tightly capped tube at 4°C.

Slide Preparation

Slides are prepared in much the same way as lymphocyte metaphase slides, but because of the small pellet, extra care is taken to conserve material.

Technical Variations

Fluorescence In Situ Hybridization

The mapping of the human genome has made it possible to extend the applications of molecular biology to cytogenetics. FISH techniques allow the highly sensitive detection of specific nucleic acid sequences in specimens on fixed microscope slides.

Several types of DNA probes have been developed and each has unique applications. A repetitive sequence probe (usually centromere-specific alphoid sequence probes) has unique DNA sequence repeats in multiple copies and is useful in determining ploidy and in identifying small marker chromosomes. Chromosome paints consist of a cocktail of different repeat and unique sequence probes, all located on the same chromosome. The unique sequences are homologous to DNA sequences along the entire length of the chromosome (Fig. 30.26A). Chromosome painting is useful in identifying translocation segments and in studies using somatic cell hybrids. Locus-specific probes are isolated for the identification of a specific gene or unique DNA sequence. For example, a locus-specific probe from the long arm of chromosome 7 can be used to look for a microdeletion associated with Williams syndrome (Fig. 30.26B).

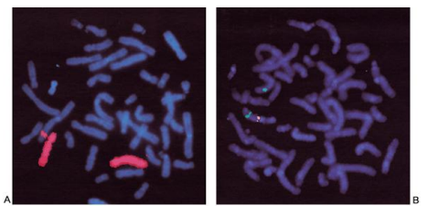


FIGURE 30.26. Fluorescence *in situ* hybridization in **A** employed a chromosome 5 painting probe on a cell from a patient with severe mental retardation and self-abusive behavior. On G-banded analysis, an abnormal dark staining band was seen in the p arm of chromosome 1. Using a 24-chromosome painting probe device (Cytocell, Ltd., Oxfordshire, UK), the abnormal segment was identified to be derived from chromosome 5 and confirmed using another chromosome 5 painting probe (Vysis, Inc, Downer's Grove, IL.). A hybridization signal is seen on both normal 5 homologs as well as an inserted segment in the p arm of chromosome 1. His karyotype is 46,XY,ins(1;5). In **B**, a locus-specific probe cocktail (Vysis, Inc., Downer's Grove, IL) used in the identification of patients with Williams syndrome was applied to a cells from a patient with clinical features of the syndrome. The orange-pink signal identifies the hybridization of a probe containing DNA sequences (including the elastin gene) known to be deleted in Williams syndrome (7q11.23). The green hybridization signal identifies control probes for 7q31. The absence of a signal with the critical region probe on one chromosome 7 identifies a deletion of the elastin gene.

For most FISH protocols, slides are prepared using standard protocols. If the slides are not aged, they are often pretreated in 2×SSC, which artificially ages the preparations. The chromosomal DNA and labeled probes are denatured, the probe is applied to the slide and incubated to allow hybridization of probe and target DNA sequences. After hybridization, excess probe is removed by stringent washing. The slides are counterstained, and the hybridization signal is analyzed using a fluorescence-equipped microscope.

With the increasing availability of commercially produced probes, FISH has become a valuable tool for cytogenetic analysis. In recent years, probes have been developed for several microdeletion syndromes including velocardiofacial syndrome and Williams syndrome, and their use has become standard in most cytogenetics laboratories. Telomere-specific probes for each arm of all chromosomes have been developed and are used to identify subtelomeric rearrangements that are estimated to be present in perhaps 6% of patients with mental retardation. Probes to identify rearrangements characteristic of specific leukemias [e.g., t(9;22) in CML and t(15;17) in AML-M3] as well as other cancer probes [e.g. her2-neu for breast cancer testing] are available and are most helpful in identifying abnormalities where sample volume is limited or where cell culture is impractical. New technologies using 24 different colored probes (one for each chromosome) are being developed and will prove most beneficial in the analysis of complex chromosome rearrangements in cancers.

Analysis of Chromosome Preparations

Cells in mid or early metaphase are usually suitable for analysis. Shorter chromosomes with fewer bands are easier to analyze but may mask subtle deletions or inversions. Selection of cells should be appropriate for each specimen type, and each laboratory should establish criteria for its own standards. A chromosome band resolution similar to that depicted in Fig. 30.6 is reasonable for most applications other than high-resolution analysis for the

contiguous gene syndromes. We specifically look for bands 17p13, 17p11.2, and 17q23 as markers of adequate resolution for routine analysis.

Preparations are analyzed using a light microscope. There are several fine manufacturers; we use a Zeiss Axioscope equipped with a 12-V, 100-W halogen lamp, mechanical, stage with verniers, 10×, 20× phase 2, and 100×/1.30 oil plan neofluar objectives, and a basic top-mounted 35-mm camera. A complete system cost approximately $18,000 in 1999.

Suitable metaphase cells have a sharp banding pattern and a small number of overlapping chromosomes, with the chromosome spread having a roughly circular shape. Excessively spread, teardrop-shaped metaphase cells are frequently broken, with one or more chromosomes lost. The location of each analyzed cell is recorded. The number of chromosomes in each spread is counted. The counts are pooled to determine the modal number (the count representative of most cells).

After the chromosomes are counted, the banding pattern of each chromosome is carefully inspected. In a sufficient number of cells, band-to-band comparison studies are done on each pair, and any structural differences are noted. Any abnormality should be confirmed in other cells because morphologic changes can occur within any cell as an artifact. Because banding intensity and condensation of chromosomal material varies between cells, thorough analysis of an adequate number of cells is essential for the proper interpretation of results. A structural or numerical abnormality found in one cell should be sought when scoring other cells. It is reasonable to assume the presence of a second clone if, among 30 cells scored, at least two cells have the same extra chromosome or the same chromosome rearrangement, or at least three cells have lost the same chromosome. In an abnormal karyotypic background, the presence of one normal cell usually indicates that a normal cell line is present.

The number of cells to be studied depends on the specimen type and the indication for the study (Table 30.12). If there is any question of chromosomal mosaicism, we analyze at least 30 cells (e.g., abnormal sex differentiation, suspected sex chromosome syndrome, or infertility). For situations in which mosaicism is occasional but is not a strong consideration, we analyze at least 20 cells (e.g., multiple congenital anomalies, mental retardation, possible Down's syndrome, or stillborns).

TABLE 30.12. BLOOD AND TISSUE ANALYSIS

Blood karyotype analysis, number of cells to study		
20 metaphase cells in most cases		
30 cells if mosaicism is possible; rule out (r/o) sex chromosome abnormality, ambiguous genitalia, male or female infertility, trisomy 8 mosaicism, or parent of a chromosomally abnormal offspring		
r/o Prader-Willi syndrome: also focus high-resolution at 15q1, 1p36, and 2q37; molecular studies are usually also ordered.		
r/o Turner syndrome: also focus high-resolution at 2q37		
Heart defect or r/o DiGeorge or VCFS; also focus high resolution at 22q11, 4q, 10p, 17p, and 18q. FISH studies are often ordered if DiGeorge or VCFS is suspected		
Tissue karyotype analysis		
20 metaphase cells in most cases		
30 cells from multiple cultures if tissue-limited mosaicism is suspected		
10 cells if karyotype is nonmosaic abnormal		
5 cells to confirm an abnormal prenatal karyotype		
Looking for mosaicism:		
When the 20- or 30-cell study is completed		
if one cell is found with	then analyze an additional	and if normal, interpretation is
Monosomy	0 cells	Random loss
A missing sex chromosome (if sex mosaic is possible)	20 cells	Random loss
+8, 9, 12–15, 18, 20–22, X, Y	20	Artifact
Other trisomic cell	0	Artifact
Balanced rearrangement	0	Artifact
if multiple miscarriages	20	Artifact
Unbalanced structural change	20	Artifact
≥50 chromosomes	0	Artifact
Normal cell in otherwise abnormal karyotype	20	Mosaic possible
+Marker	20	Artifact
if two cells are found with		
The same missing chromosome, look for mosaic or a ring	20	Random loss
–Y in male over age 50	0	Not clinically significant
-X or +X in female over age 23	0	Not clinically significant
If true mosaicism is found, consider UPD studies		

UPD, uniparental disomy.

These guidelines are used in our laboratory and are provided as an example of protocol development to make the most efficient use of personnel working at the microscope and to assist the staff in the preliminary interpretation of findings that are usually not clinically significant. All abnormal cells are also brought to the attention of the person responsible for clinical interpretation of the karyotype findings.

After the analysis, photographs are taken of metaphase cells representative of the analyzed cells (73,74). Photographs serve as a permanent record of what was observed, and we do not recommend their use as an alternative to routine analysis through the microscope. Black and white enlargements are made, and individual chromosomes are cut to arrange a karyotype. The final report includes at least two full karyotypes.

Guidelines and recommendations for culture and analysis for all specimen types have been established by the American College Medical Genetics (75) and the Association of Genetic Technologists (76) and are pending from the American College of Medical Genetics. Chromosome nomenclature should conform to standards established by the Standing Committee of Human Cytogenetic Nomenclature (1).

As an alternative to photographic documentation, computerized

systems for cytogenetic analysis have been developed by several companies. Basic systems allow the technologist to image digitally metaphase cells and prepare a karyotype. The karyotype can be printed in hard copy format and can be archived on a disk or other memory system. Sophisticated systems are able to locate automatically metaphase cells on slides, count the number of chromosomes, and produce a karyotype of the cells, using the technologists to isolate and identify some or all of the individual chromosomes in an interactive process.

Sex Chromatin Studies (Barr Bodies)

The Barr body, formed during interphase from the inactivated X chromosome, can be observed in cells from normal females and males with Klinefelter's syndrome, but not in normal males or females with 45,X or 46,XY constitutions. Females with more than two X chromosomes have some cells with multiple Barr bodies, and some structurally abnormal X chromosomes show smaller or larger Barr bodies. Scoring Barr bodies in buccal mucosal cells is less costly than karyotype analysis, but the technical limitations of the test are considerable, and the presence of the Y chromosome is not addressed without application of a second stain for the Y body (the quinacrine-brilliant segment of Yq). Because of the obvious benefit of a full cytogenetic study, sex chromatin studies have fallen out of favor.

REFERENCES

1. Mitelman F, ed. ISCN (1995). *An international system for human cytogenetic nomenclature*. Basel: Karger, 1995:114.

2. Bond DJ, Chandley AC. *Aneuploidy*. Oxford: Oxford University Press, 1983:13.

3. Hook EB. Cytogenetics: incidence, risks, and recurrence of chromosome abnormalities. In: Brock DJH, Rodeck C, Ferguson-Smith MA, eds. *Prenatal diagnosis and screening*. London: Churchill-Livingstone, 1992:351–392.

4. Koehler KE, Hawley SR, Sherman S, et al. Recombination and nondisjunction in humans and flies. *Hum Mol Genet* 1996;5:1495–1504.

5. Falik-Borenstein TC, Pribyl TM, Van Dyke DL, et al. A stable ring chromosome 21: molecular and clinical definition of the lesion. *Am J Med Genet* 1992;42:22–28.

6. Hook EB, Hamerton JL. The frequency of chromosome abnormalities detected in consecutive newborn studies—differences between studies—results by sex and by severity of phenotypic involvement. In: Hook EB, Porter TH, eds. *Population cytogenetics studies in humans*. New York: Academic Press, 1977:63–79.

7. Van Dyke DL, Weiss L, Roberson JR, et al. The frequency and mutation rate of balanced autosomal rearrangements in man estimated from prenatal genetic studies for advanced maternal age. *Am J Hum Genet* 1983;35:301–308.

8. Bourrouillou G, Colombies P, Dastugue N. Chromosome studies in 2136 couples with spontaneous abortions. *Hum Genet* 1986;74: 399–401.

9. Boue A, Gallano P. A collaborative study of the segregation of inherited chromosome structural rearrangements in 1356 prenatal diagnoses. *Prenat Diagn* 1984;4:45–68.

10. Daniel A, ed. *The cytogenetics of mammalian autosomal rearrangements*. New York: Alan R Liss, 1988.

11. Forsythe MG, Walker H, Weiss L, et al. Duplication and deletion 11q23–q24 recombinants in two offspring of an intrachromosomal insertion ("shift") carrier. *Henry Ford Hosp Med J* 1988;36:183–186.

12. deGrouchy J, ed. Special issue on pericentric and paracentric inversions in man. *Ann Genet* 1986;29:126–214.

13. Worsham MJ, Miller DA, Devries JM, et al. A dicentric recombinant 9 derived from a paracentric inversion: phenotype, cytogenetics, and molecular analysis of centromeres. *Am J Hum Genet* 1989;44:115–123.

14. Feldman GL, Weiss L, Phelan MC, et al. Inverted duplication of chromosome 8p: ten new patients and review of the literature. *Am J Med Genet* 1993;47:482–486.

15. Jaarola M, Martin RH, Ashley T. Direct evidence for suppression of recombination within two pericentric inversions in humans: a new sperm-FISH technique. *Am J Hum Genet* 1998;63:218–224.

16. Engel E. Uniparental disomies in unselected populations. *Am J Hum Genet* 1998;63:962–966.

17. Ledbetter DH, Engel E. Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. *Hum Mol Genet* 1995;4:1757–1764.

18. Kotzot D. Abnormal phenotypes in uniparental disomy (UPD). *Am J Med Genet* 1999;82:265–274.

19. Kalousek DK, Barrett J. Genomic imprinting related to prenatal diagnosis. *Prenat Diagn* 1994;14:1191–1201.

20. deGrouchy J, Turleau C. *Clinical atlas of human chromosomes*. 2nd ed. New York: John Wiley & Sons, 1984.

21. Jones KL. *Smith's recognizable patterns of human malformation*, 4th ed. Philadelphia: Saunders, 1988.

22. Schmickel RD. Contiguous gene syndromes: a component of recognizable syndromes. *J Pediatr* 1986;109:231–241.

23. Wu Y-Q, Sutton VR, Nickerson E, et al. Delineation of the common critical region in Williams syndrome and clinical correlation of growth, heart defects, ethnicity, and parental origin. *Am J Med Genet* 1998;78:82–89.

24. Brewer C, Holloway S, Zawalnyski P, et al. A chromosomal deletion map of human malformations. *Am J Hum Genet* 1998;63:1153–1159.

25. Altherr MR, Wright TJ, Denison K, et al. Delimiting the Wolf-Hirschhorn syndrome critical region to 750 kilobase pairs. *Am J Med Genet* 1997;71:47–53.

26. Niebuhr E. The cri du chat syndrome. Epidemiology, cytogenetics, and clinical features. *Hum Genet* 1978;44:227–275.

27. Dutley F, Baumer A, Kayserili H, et al. Seven cases of Wiedemann-Beckwith syndrome, including the first reported case of mosaic paternal isodisomy along the whole chromosome 11. *Am J Med Genet* 1998;79:347–353.

28. Catchpoole D, Lam WWK, Valler D, et al. Epigenetic modification and uniparental inheritance of H19 in Beckwith-Wiedemann syndrome. *J Med Genet* 1997;34:353–359.

29. Mathieu M, Puissan C, Thepot F, et al. Collaborative study of mosaic tetrasomy 12p or Pallister-Killian syndrome (nineteen fetuses or children). *Ann Genet* 1997;40:45–54.

30. Kushnick CD, Kobori JA, Mohapatra G, et al. Prader-Willi syndrome is caused by disruption of the SNRPN gene. *Am J Hum Genet* 1999;64:70–76.

31. Monaghan KG, Van Dyke DL, Feldman G, et al. Diagnostic testing: a cost analysis for Prader-Willi and Angelman syndromes [Letter]. *Am J Hum Genet* 1997;60:244–247.

32. Cassidy SB. Prader-Willi syndrome. *J Med Genet* 1997;34:917–923.

33. Gunay-Aygun M, Heeger S, Schwartz S, et al. Delayed diagnosis in patients with Prader-Willi syndrome due to maternal uniparental disomy 15. *Am J Med Genet* 1997;71:106–110.

34. Jiang Y, Lev-Lehman E, Bressler J, et al. Genetics of Angelman syndrome. *Am J Hum Genet* 1999;65:1–6.

35. Kushino T, Lalande M, Wagstaff J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* 1997;15:70–73.

36. Matsuura T, Sutcliffe JS, Fang P, et al. De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. *Nat Genet* 1997;15:74–77.

37. Cheng SD, Spinner NB, Zackai EH, et al. Cytogenetic and molecular characterization of inverted duplicated chromosome 15 from 11 patients. *Am J Hum Genet* 1994;55:753–759.

38. Greenberg F, Lewis RA, Potocki L, et al. Multidisciplinary clinical study of Smith-Magenis syndrome (deletion 17p11.2). *Am J Med Genet* 1996;62:247–254.

39. Van Dyke DL, Roberson JR, Babu VR, et al. Trisomy 20 mosaicism identified prenatally and confirmed in foreskin fibroblasts. *Prenat Diagn* 1989;9:601–602.

40. Epstein CJ. Down syndrome (trisomy 21). In: Scriver CR, Beaudet AL, Sly WS, et al., eds. *The metabolic basis of inherited disease*, 7th ed. New York: McGraw-Hill, 1995:749–794.

41. Therman E, Susman B, Denniston C, et al. The nonrandom participation of human acrocentric chromosomes in Robertsonian translocations. *Ann Hum Genet* 1989;53:49–65.

42. Giraud F, Mattei JF. Aspects epidemiologiques de la trisomie 21. *J Genet Hum* 1975;23:1–30.

43. Daniel A, Hook EB, Wulf G. Risks of unbalanced progeny at amniocentesis to carriers of chromosome rearrangements: data from United States and Canadian laboratories. *Am J Med Genet* 1989;33:14–53

44. Chapman CJ, Gardner RJM, Veale AMO. Segregation analysis of a large t(21q22q) family. J Med Genet 1973;10:362–366.

45. Hernandez D, Fisher EMC. Down syndrome genetics: unravelling a multifactorial disorder. *Hum Mol Genet* 1996;5:1411–1416.

46. Cronk C, Crocker AC, Pueschel SM, et al. Growth charts for children with Down syndrome: 1 month to 18 years of age. *Pediatrics* 1988;81:102–110.

47. Pueschel SM. Clinical aspects of Down syndrome from infancy to adulthood. *Am J Med Genet* 1990;7(Suppl):52–56.

48. Mikkelsen M. Down's syndrome—current stage of cytogenetic research. *Humangenetik* 1971;12:1–28.

49. Swillen A, Devriendt K, Vantrappen G, et al. Familial deletions of chromosome 22q11: the Leuven experience. *Am J Med Genet* 1998;80:531–532.

50. Holmes SE, Riaze MA, Gong W, et al. Disruption of the clathrin heavy chain-like gene (CLTCL) associated with features of DGS/VCFS: a balanced (21;22)(p12;q11) translocation. *Hum Mol Genet* 1997;6: 357–367.

51. Ryan AK, Goodship JA, Wilson DI, et al. Spectrum of clinical features associated with interstitial chromosome 22q11 deletions: a European collaborative study. *J Med Genet* 1997;34:798–804.

52. Van Geet C, Devriendt K, Eyskens B, et al. Velocardiofacial syndrome patients with a heterozygous chromosome 22q11 deletion have giant platelets. *Pediatr Res* 1998;44: 607—611.

53. Iselius L, Lindsten J, Aurias A, et al. The 11q;22q translocation: a collaborative study of 20 new cases and analysis of 110 families. *Hum Genet* 1983;64:343–355.

54. Ratcliffe SG, Paul N, eds. *Prospective studies on children with sex chromosome aneuploidy*. New York: Alan R. Liss, 1986.

55. Migeon BR, Luo S, Stasiowski BA, et al. Deficient transcription of XIST in individuals with tiny ring X chromosomes and severe phenotypes: implications for the role of XIST in X inactivation. *Proc Nat Acad Sci U S A* 1993;90:12025–12029.

56. Adashi EY, Hennebold JD. Single-gene mutations resulting in reproductive dysfunction in women. *N Engl J Med* 1999;340:709–718.

57. Stewart DA, ed. *Children with sex chromosome aneuploidy: followup studies*. [March of Dimes Birth Defects: Original Article Series 18(4)]. New York: Alan R. Liss, 1982:251.

58. Angell RR, Sandison A, Bain AD. Chromosome variation in perinatal mortality: a survey of 500 cases. *J Med Genet* 1984;21:39–44.

59. Simpson, Bombard. In: Bennett MJ, Edmonds DK, eds. *Spontaneous and recurrent abortion*. Oxford: Blackwell Scientific Publications, 1987:51–76.

60. Hassold TJ. A cytogenetic study of repeated spontaneous abortion. *Am J Hum Genet* 1980;32:723–730.

61. Retief AE, Van Zyl JA, Menkveld R, et al. Chromosome studies in 496 infertile males with a sperm count below 10 million/ml. *Hum Genet* 1984;66:162–164.

62. Bourrouillou G, Dastugue N, Colombies P, et al. Chromosome studies in 952 infertile males with a sperm count below 10 million/ml. *Hum Genet* 1985;71:366–367.

63. Kim ED, Bischoff FZ, Lipschultz LJ, et al. Genetic concerns for the subfertile male in the era of ICSI. *Prenat Diagn* 1998;18:1349–1365.

64. Allingham-Hawkins DJ. Fragile X premutation is a significant risk factor for premature ovarian failure. *Am J Med Genet* 1999;83:322--225.

65. Blethen SL, Weldon VV. Disorders of external genital differentiation. In: Kelley VC, ed. *Practice of pediatrics*, rev. ed. Philadelphia: Harper and Row, 1987;7:1–23.

66. Lanes RL. Ambiguous genitalia, micropenis, and cryptorchidism. In: Lifshitz F, ed. *Pediatric endocrinology: a clinical guide*, 2nd ed. New York: Marcel Dekker, 1990.

67. Moore DC. Disorders of gonadal differentiation. In: Kelley VC, ed. *Practice of pediatrics*, rev. ed. Philadelphia: Harper & Row, 1987;7:1–30.

68. Styne DM. The testes: disorders of sexual differentiation and puberty. In: Kaplan SA, ed. *Clinical pediatric endocrinology*. Philadelphia: Saunders, 1990:367–425.

69. Sutherland GR, Mulley JC. Fragile X syndrome and fragile XE mental retardation. *Prenat Diagn* 1996:16:1199–1211.

70. De Vries BBA. The fragile X syndrome. *J Med Genet* 1998;35:579–589.

71. Flannery DB. Pigmentary dysplasia, hypomelanosis of Ito, and genetic mosaicism. *Am J Med Genet* 1990;35:18–21.

72. Reed SD, Hall JG, Riccardi VM, et al. Chromosomal abnormalities associated with congenital contractures (arthrogryposis). *Clin Genet* 1985;27:353–372.

73. Bradbury S. *An introduction to the optical microscope*. Oxford: Oxford University Press, 1984.

74. Delly JG. *Photography through the microscope*. Rochester, NY: Eastman Kodak, 1988.

75. American College of Medical Genetics. *Standards and guidelines: clinical genetics laboratories*, 2nd ed. Bethesda, MD: American College of Medical Genetics, 1999.

76. Barch MJ, ed. *The Association of Cytogenetics Technologists cytogenetics laboratory manual*, 3rd ed. New York: Raven Press, 1997:666.

31
Prenatal Cytogenetic Diagnosis

Daniel L. Van Dyke

Jacquelyn R. Roberson

Anne Wiktor

With advancing maternal age, there is an increased risk of having a child with Down's syndrome. This fact has led to the development and wide acceptance of prenatal diagnosis of chromosomal and nonchromosomal birth defects and inherited conditions. The methods used to obtain fetal cells for prenatal diagnosis include amniocentesis, chorionic villus sampling (CVS), and umbilical blood sampling (1). Methods to obtain fetal cells from the maternal blood circulation may not be perfected for years, but several biochemical measures of maternal serum are used to screen pregnancies at increased risk. Including all accepted reasons (Table 31.1), prenatal diagnosis is indicated in 5% to 8% of all pregnancies. Other than screening for neural tube and abdominal wall defects, nonchromosomal prenatal diagnostic studies are beyond the scope of this chapter. For a detailed discussion of these topics, the reader is referred to Milunsky (2).

TABLE 31.1. INDICATIONS FOR PRENATAL CYTOGENETIC OR AFP DIAGNOSIS

Advanced maternal age
One parent carries a chromosome rearrangement
Previous offspring with a chromosome abnormality
Previous offspring with a neural tube defect or hydrocephalus
High maternal serum AFP
Low maternal serum AFP (or similar screening result)
Fetal abnormality detected by ultrasound examination
Fetus at risk for X-linked or prenatally diagnosable inherited disease

AFP, α-fetoprotein.

Amniocentesis for prenatal genetic diagnosis is typically performed at 16 to 20 weeks. The procedure has been performed as early as 9 weeks, although its safety seems questionable before 15 weeks of gestation. Genetic amniocentesis has become widely available since it was introduced in the late 1960s, it is regarded as safe to the mother and fetus, and the culture success rate is 99.5% or better. Methods are also available to screen amniotic fluid interphase cells for the common trisomies and sex chromosome abnormalities, using fluorescence *in situ* hybridization (FISH) of probes for chromosomes 13, 18, 21, X, and Y.

Although CVS for prenatal diagnosis got off to a rocky start, its acceptance is increasing again. The reason for concerns in the early 1990s was an association between CVS and limb reduction defects, but this was primarily limited to procedures done before 10 weeks of gestation. The risks of the procedure to the mother and the embryo are now considered acceptable if performed after 10 weeks of gestation, with more than 96% of procedures providing a satisfactory cytogenetic result (3,4,5 and 6). CVS for first-trimester prenatal diagnosis is routinely done between gestational age of 10 and 12 weeks using an ultrasound-guided biopsy catheter passed through the cervix (transcervical CVS) or abdominal wall (transabdominal CVS). Transabdominal CVS (placental biopsy) can also be performed during the second and third trimesters and is sometimes useful to evaluate the karyotype when oligohydramnios or fetal defects are found on ultrasound examination.

For CVS, two general chromosome preparation methods have been developed: direct preparation and short-term cell culture. Short-term culture of cells, mainly originating from the mesenchymal core of the chorionic villus, has the disadvantage of a longer reporting time. However, this is offset by the advantages of better quality preparations, the opportunity to analyze independent colonies of cells harvested *in situ*, no greater frequency of false-positive results (0.5% to 1% of specimens), and a lower frequency of false-negative results. (A false-positive result would mean the karyotype interpretation is abnormal but the fetal karyotype is actually normal. A false-negative result would mean that the karyotype interpretation is normal, but the fetal karyotype is actually abnormal.)

The alternative CVS method is a direct preparation of uncultured, rapidly dividing cytotrophoblast cells that comprise the outer sheath of the villus, after 24 to 48 hours of incubation of these cells. This has the advantage of more rapid reporting time but the disadvantages of a higher false-negative rate and poorer quality preparations. The false-negative rate of 0.03% to 0.1%, missing approximately 1% of the true fetal chromosome abnormalities, has been a major factor leading to broader acceptance of CVS culture techniques. For second- and third-trimester CVS, the direct method, or percutaneous umbilical blood sampling (PUBS), may be preferable to culture of amniocytes or CVS cells if a very rapid result is the first consideration, although false-negative CVS direct preparation results have been reported for all three trimesters. Thus, cultures should be employed for any CVS chromosome studies. The false-negative rate is unknown for second- and third-trimester CVS studies.

PUBS has also been used to obtain fetal cells for chromosome analysis, although it is not widely used for this purpose, being more applicable to studies of fetomaternal infections and Rh incompatibility.

INDICATIONS FOR PRENATAL CHROMOSOME ANALYSIS

Part of "31 - Prenatal Cytogenetic Diagnosis"

Maternal Age

In our center, approximately 55% of prenatal diagnostic procedures are done primarily because the mother will be age 35 years or older at birth. The risks of trisomy 13, 18, 21, X, and 47,XXY all increase with maternal age (7) (Table 31.2). There is an increased risk of some other chromosome defects with maternal age as well, including deletion 18p, extra isochromosomes, and other small accessory marker chromosomes. In our experience, such nontrisomic chromosome imbalances that are associated with maternal age constitute 10% of prenatally diagnosed chromosome abnormalities. In the United States, a maternal age of 35 years at the expected date of confinement (EDC) is the accepted age at which physicians routinely offer prenatal diagnosis Some centers have suggested lowering the age criterion to age 32 at EDC. In fact, any criterion is to some extent arbitrary. Some women would like to have prenatal diagnosis no matter what their risk; others think it wholly unnecessary. Increasingly, women over age 35 are basing their decision to undergo amniocentesis on the estimated risk provided by maternal serum screening programs.

TABLE 31.2. RATE PER THOUSAND OF CHROMOSOME ABNORMALITIES AT CVS AND AMNIOCENTESIS, AND IN LIVEBIRTHS, BY MATERNAL AGE

	Trisomy 21			All Abnormalities		
Maternal Age	CVS	Amnio	Livebirth	CVS	Amnio	Livebirth
35	4.2	4.0	2.6	8.8	8.0	5.2
36	5.7	5.2	3.4	11.5	9.9	6.4
37	7.5	6.7	4.4	15.1	12.2	7.9
38	10.0	8.7	5.7	19.8	15.2	9.8
39	13.4	11.2	7.3	26.0	19.0	12.1
40	17.9	14.5	9.4	34.1	23.8	15.2
41	23.8	18.7	12.2	44.7	29.9	19.0
42	31.7	24.1	15.7	58.6	37.6	23.9
43	42.3	31.1	20.2	76.8	47.2	30.2
44	56.4	40.1	26.1	100.8	57.8	37.3
45	75.1	51.8	33.7	132.1	71.3	46.6
46	˜100	66.8	43.4	˜170	88.6	58.4
47	˜130	86.2	56.0	˜230	110.9	73.6
48	˜180	111.2	72.3	˜300	139.9	90.4

CVS, chorionic villus sampling; amnio, amniocentesis. Statistics from Hook (7).

A chromosome abnormality is identified in 5% to 8% of CVS from women age 35 or older, compared with 1.5% to 3% of amniotic fluid samples from the same age group. The abnormality rate is higher in earlier gestational weeks. However, if detailed ultrasound examination is employed to exclude deceased embryos, the rate of chromosome abnormalities at CVS is approximately the same as for amniocentesis. Most centers do not attempt CVS if the embryo is dead, but perhaps it is worth considering because nearly all this subset have an abnormal karyotype and the cell culture success rate is much better for CVS than for evacuated products of conception.

Positive Family History

Between 1% and 3% of prenatal chromosome studies are performed because one parent carries a balanced chromosome rearrangement or the couple has had a child or stillborn with a chromosome abnormality. The risk of a liveborn infant with an unbalanced chromosome rearrangement varies with the specific rearrangement. In general, the risk is higher if the carrier is female. The risk of recurrence for a structural chromosome abnormality is probably much less than 1% if the parental karyotypes are normal.

For recurrence of a numerical chromosome abnormality such as trisomy 21, the empirical risk is almost never more than 1% greater than the maternal age-adjusted risk.

Prenatal Screening

Prenatal screening of maternal serum α-fetoprotein (MSAFP) and related tests (described below) has become part of routine obstetric care. Abnormal serum screen results account for approximately 30% of amniocentesis procedures at our institution. The proportion of women tested and the proportion of tests for these indications vary somewhat among centers. Most centers study the karyotype for screen-positive maternal serum multiple markers whether it is screen positive for neural tube defects or chromosome abnormalities as an indication, even though a high MSAFP is not generally considered a risk factor for a chromosome abnormality.

In our experience, a clinically significant karyotype abnormality is identified in 0.6% to 1% of amniotic fluid chromosome studies done because of abnormal maternal serum screening.

Ultrasound-Identified Abnormalities

Approximately 10% of prenatal cytogenetic studies are done because an ultrasound examination identified anatomic defects, obstructive uropathy, intrauterine growth retardation (IUGR), nonimmune hydrops, oligohydramnios, or polyhydramnios (Table 31.3). Of these, 10% to 20% (in our laboratory) have a chromosome abnormality: 20% of these were 45,X, 27% trisomy 21, 18% trisomy 18, 8% trisomy 13, 5% triploid, and 22% involved structural or other abnormalities. Fetal nuchal skin thickening observed between 15 and 20 weeks of gestation is used in some centers as a risk factor for trisomy 21 (11). Cystic hygroma is associated with 45,X in 50% to 75% of cases. There is usually more generalized fetal edema or fetal hydrops, and if so, spontaneous fetal death is very likely. Cystic hygroma is less commonly associated with trisomy 21 and other chromosome abnormalities, although nonimmune hydrops without cystic hygroma has approximately a 25% rate of chromosome abnormalities with approximately half of these being trisomy 21, one fourth 45,X, and one fourth other abnormalities. If amniotic fluid is inaccessible, it is possible to obtain cells for both phytohemagglutinin-stimulated culture and for standard monolayer culture from a sample of any bodily fluid or chorionic villi.

TABLE 31.3. KARYOTYPE AND FETAL ABNORMALITIES DETECTED ON ULTRASOUND EXAMINATION

Structural Abnormality	Most Common Chromosome Abnormality
Intrauterine growth retardation	Triploid, +18, +13, +21, 45,X
Brain	
Ventriculomegaly	+18, +21, triploid, +13
Holoprosencephaly	+13, +18, triploid
Microcephaly	+13
Choroid plexus cysts	+18
Dandy-Walker cyst	+18, +13
Agenesis of corpus callosum	+18, +13, triploid
Strawberry-shaped head	+18
Neck and face	
Cleft lip	+13, +18
Micrognathia	+18, triploid, +13
Nuchal cystic hygroma	45,X, +21
Nuchal thickening	+21, +13, +18
Thorax and abdomen	
Heart defect, structural	22q11 deletion, +21, +13, +18, triploid, 45,X
Nonimmune hydrops	45,X, +21
Diaphragmatic hernia	+18, +i(12p)
Omphalocele	+18, +13
Duodenal atresia	+21
Echogenic bowel	+21 (also cystic fibrosis)
Bladder outlet obstruction	+18, +13
Multicystic/dysplastic kidney	+13
Limbs	
Short femur	+21, +13, +18, triploid, and 45,X
Polydactyly	+13
Syndactyly	+18, triploid
Placenta, amniotic fluid	
Partial mole	Triploid
Complete mole	Androgenic 46,XX
Polyhydramnios	+21, +18M

Frequency estimates are from our laboratory experience. The risk of a karyotype abnormality associated with an isolated defect is usually low and sometimes confers no increased risk. The observation of one defect is an indication for a detailed ultrasound examination. Sources include Eydoux et al. (8), Snijders and Nicolaides (9), Stoll et al. (10), Van Dyke and Wiktor, unpublished data.

Fetal cardiac defects merit special mention because of their association with the velocardiofacial (22q11 deletion) syndrome. The classic phenotype (Chapter 30) includes a conotruncal heart malformation. However, the phenotypic spectrum associated with 22q11 deletion is very wide, and almost all structural heart defects have been reported in affected patients (12). We therefore use a FISH probe for the critical region associated with 22q11 deletion syndrome in addition to routine karyotype analysis in all prenatally diagnosed structural heart defects (excluded are hypoplastic left heart, echogenic focus in the heart, and fetal arrhythmia). Using this protocol, we identify a 22q11 deletion in 9% of such pregnancies.

Chromosome Instability Syndromes

Prenatal diagnosis of the mendelian chromosome instability syndromes are described in Chapter 29.

X-Linked Mendelian Disorders

A large number of X-linked conditions such as fragile X syndrome, Duchenne's and Becker's muscular dystrophy, hemophilia A and B, and Hunter's syndrome can be prenatally diagnosed with confidence (http://www.genetests.org/). Fetal sex can be determined by chromosome analysis or FISH analysis of X and Y chromosome-specific DNA sequences.

GENETIC COUNSELING

Part of "31 - Prenatal Cytogenetic Diagnosis"

Before any prenatal diagnostic procedure, it is important to obtain a family history. This would include, at a minimum, all first-degree relatives of both parents for birth defects, mental retardation, heritable diseases. It is also important to look for other risk factors that can be addressed in a constructive way (e.g., maternal diabetes or seizures, exposure to alcohol, cocaine, or other teratogenic agents). If the indication is a family history of a specific birth defect, a known syndrome, or an inborn error of metabolism, it is imperative that the specific diagnosis be confirmed. Failure to do so can cause a prenatal diagnostic error. For routine obstetric use, several brief questionnaires have been devised to assist in identifying genetic and teratogenic risk factors.

Informed consent for a prenatal diagnostic procedure should include discussion of the risks to the mother and the fetus and accuracy of the tests being done on the sample. If the patient is trying to choose between CVS and amniocentesis, a discussion of the advantages and limitations of each technique should be provided (Table 31.4). For amniocentesis, there is approximately a one in 500 risk of miscarriage or infection. Infection is a potentially serious risk to the mother and her future fertility. For CVS, the risk of miscarriage is estimated at 1/125 to one in 300. Although the risk of severe maternal infection after transcervical CVS has been estimated at approximately one in 500, as experience with the procedure has increased, this risk seems to have decreased dramatically. Because of the potential of a vascular disruption sequence (e.g., limb reduction or hypoglossia-hypodactyly

sequence), CVS is no longer recommended before 10 weeks of gestation. It is uncertain whether the relatively small increased risks of miscarriage owing to CVS or amniocentesis procedures are statistically different from each other. It is important that patients understand that prenatal diagnosis employs a specific set of tests and does not identify all causes of birth defects or mental retardation and that some diagnostic results are difficult to interpret or need further studies before interpretation is possible.

TABLE 31.4. COMPARATIVE ADVANTAGES OF PRENATAL CYTOGENETIC DIAGNOSIS USING THE CHORIONIC VILLUS SAMPLING AND AMNIOCENTESIS TECHNIQUES

Chorionic villus sampling
Earlier gestational age provides more personal privacy
Pregnancy termination at an earlier gestational age is quicker and safer
Preferred by many women, assuming similar risks to fetus and mother
Anxieties relieved sooner, although amniocentesis follow-up, necessary in 1% to 3%, is a source of new anxiety
Amniocentesis
Usually less expensive
Risks to fetus and mother are slightly lower
Wide availability
Available after maternal serum screening
Ability to assay amniotic fluid α-fetoprotein and acetylcholinesterase
Fewer confusing results (e.g., mosaicism 0.3% vs. 2% in chorionic villus sampling)

A follow-up study, usually including amniocentesis, is necessary in 1% to 2% of CVS cases because of an inadequate sample or an indeterminate cytogenetic result (usually mosaicism). In addition, approximately 1% of women who have CVS also need amniocentesis because of an elevated MSAFP result.

An abnormal or unexpected cytogenetic result is obtained in approximately 3% of all prenatal studies (Table 31.5). Half of these are autosomal or sex chromosomal trisomies. Many of the rest require follow-up studies and can be difficult to interpret.

TABLE 31.5. CHROMOSOME ABNORMALITY IN PRENATAL STUDIES AND PARENTAL DECISIONS REGARDING CONTINUATION OF THE PREGNANCY IN OUR PROGRAM

Abnormality	% of All Results[a]	% of All Abnormals	% Who Continued
47,+21	0.7	31	8
47,+13	0.1	5	0
47,+18	0.1	5	0
46/47,+20	0.05	2	75
45,X[b]	0.1	4	(None has survived)
47,XXX	0.1	4	25
47,XXY	0.1	5	66
47,XYY	0.07	3	0
45,X/46,XX or XY[b]	0.1	4	100
Inherited balanced rea	0.6[c]	26	100
De novo balanced rea	0.1	5	100
Inherited extra marker	0.07	3	100
De novo extra marker	0.05	2	50
Other	0.02	1	—

[a] Approximately 2% to 3% of prenatal studies have an abnormal cytogenetic result, and 0.3% are true mosaics (or roughly 10% of the abnormal results).
[b] The nonmosaic 45,X cases were associated with ultrasound-diagnosed abnormalities. The 45,X mosaics all appeared normal on ultrasound and had a normal outcome at birth.
[c] Between 75% and 85% are inherited from the mother.

At the time of CVS, approximately 10% to 13% fetal nonviability or an abnormal gestational sac is discovered (this is the frequency in the CVS population, not the entire pregnant population). Approximately 6% of women of ages 35 to 39, and approximately 20% of women age 40 or over miscarry between weeks 7 and 12. Among women age 35 or over who schedule amniocentesis, approximately 4% miscarry before their appointment date. When fetal demise is discovered at ultrasound, the prenatal clinic staff must be prepared to help the family deal with their grief.

Parents do not need to commit themselves to discontinue a pregnancy should a chromosome abnormality be identified. One cannot be expected to make such decisions in the abstract. These are complicated, personal decisions in which the genetics counseling team can and should play a supporting and nondirective role. The burdens of a developmentally disabled child can be tremendous on a family and may lead to high divorce rates and other less obvious outcomes, including annual costs of care of

more than $100,000. The decision whether to continue the pregnancy depends to some extent on the perceived burden of the defect (Table 31.5), but some parents will continue the pregnancy regardless of the diagnosis, and others will elect to terminate the pregnancy in relatively low-risk situations.

When a trisomic pregnancy is continued, the remaining chance of spontaneous fetal demise is approximately 30% for trisomy 21, 40% for trisomy 13, and 70% for trisomy 18. For monosomy X, the chance of a demise approaches 100%.

For laboratory and clinic quality assurance, our clinic employs postcard or telephone follow-up of every continued pregnancy with a normal karyotype result. One cannot assume that a discrepant result will be automatically reported. Long-term follow-up of continued pregnancy after an abnormal or unclear cytogenetic result is recommended.

SCREENING FOR NEURAL TUBE DEFECTS

Part of "31 - Prenatal Cytogenetic Diagnosis"

Prenatal diagnosis for spina bifida defects is the second most common reason for amniocentesis. Open neural tube defects, including anencephaly, meningocele, and meningomyelocele, occur in approximately one to two newborns per thousand in the United States.

Anencephaly is a lethal condition, and most affected fetuses are stillborn. The prognosis for meningocele and meningomyelocele can be quite variable, depending on the anatomic site of the defect, the extent of spinal cord damage, and the presence or absence of associated hydrocephalus.

The familial nature of spina bifida has been well documented, although many genes influence the risk (multifactorial inheritance). For couples who have one child with spina bifida, the empirical risk for recurrence is in the range of 2% to 5%. However, 95% of infants with spina bifida are born to couples with no family history.

Some cases of spina bifida are associated with other anomalies as part of a genetic syndrome such as Meckel or Roberts syndrome, both autosomal recessive conditions with a 25% risk of recurrence. Some can be associated with other risk factors such as maternal diabetes or prenatal medication exposure (i.e., isotretinoin or valproic acid).

Prior to availability of MSAFP screening, only patients with a positive family history of spina bifida were offered prenatal diagnosis. Although ultrasound examination can identify between 50% and 75% of spina bifida defects in the second trimester of pregnancy (13), its high cost has prevented its adoption as a screening test.

α-Fetoprotein (AFP) is the major protein in fetal plasma in early pregnancy. Peak levels of AFP are attained in fetal plasma at 12 to 14 weeks of gestation. The concentration falls rapidly thereafter. AFP levels in amniotic fluid parallel fetal serum levels at 1/100th concentration. In maternal serum, however, AFP levels rise throughout the first and second trimesters and peak at 28 to 30 weeks of gestation (Fig. 31.1). In 1972, Brock and Sutcliffe (14) were the first to diagnose spina bifida defects prenatally by elevated amniotic fluid AFP levels.

In 1973, Brock and co-workers first reported the use of MSAFP assays to identify pregnancies at increased risk for spina bifida, introducing the concept of mass screening for open neural tube defects. Patients with elevated MSAFP values could then be offered diagnostic testing such as amniocentesis and ultrasound evaluation.

Antenatal Screening for Spina Bifida

Although neonatal screening for some genetic disorders such as phenylketonuria was commonplace by the early 1960s, the idea of prenatal screening evoked considerable anxiety. The anxiety was partly related to the fact that an elevated MSAFP level was not always associated with a birth defect in the fetus. Furthermore, an elevated MSAFP did not identify all the fetuses that had a neural tube defect. However, the element of false-positive and false-negative test results is inherent in the concept of screening tests. A valid screening test identifies a subgroup of patients that has a sufficient likelihood of an abnormality to warrant further diagnostic investigation. Over time, MSAFP screening gained broad acceptance, and during 1990, approximately one half of all pregnancies in the United States were screened.

The United Kingdom was the first to apply population-based screening for spina bifida defects (Table 31.6). Maternal serum samples obtained between 15 and 20 weeks of gestation were found to have the best predictive value. Approximately 85% of affected fetuses were identified with MSAFP cutoffs of 2.5 multiples of the median (MoM). Approximately 5% of all pregnant women would have values above this level, but only one woman in 50 in this high-risk group would actually carry a fetus with anencephaly or spina bifida. One half of the women with initially elevated MSAFP values had an inaccurate estimate of gestational age, fetal demise, or twin gestation as the cause of the elevation. Because an elevated MSAFP result is not diagnostic of a fetal abnormality, MSAFP screening can be effective only within a well-integrated program offering genetic counseling, ultrasound evaluation, and amniocentesis (15). The United Kingdom's nationalized health care system offered an ideal setting for such a coordinated program. After encouraging data were presented concerning MSAFP screening programs in the United Kingdom and after extensive confirmatory studies, MSAFP test kits were released in the United States in 1983.

MSAFP Methods

MSAFP determinations are usually made by radioimmunoassay or immunometric assay techniques. Values may be expressed in either nanograms or International Units per milliliter. Commercial MSAFP kits have been tested in clinical trials and “kit” normal ranges are published for each week of gestation, but there is considerable variability among laboratories. Therefore, “kit normal values” are not recommended in establishing upper (or

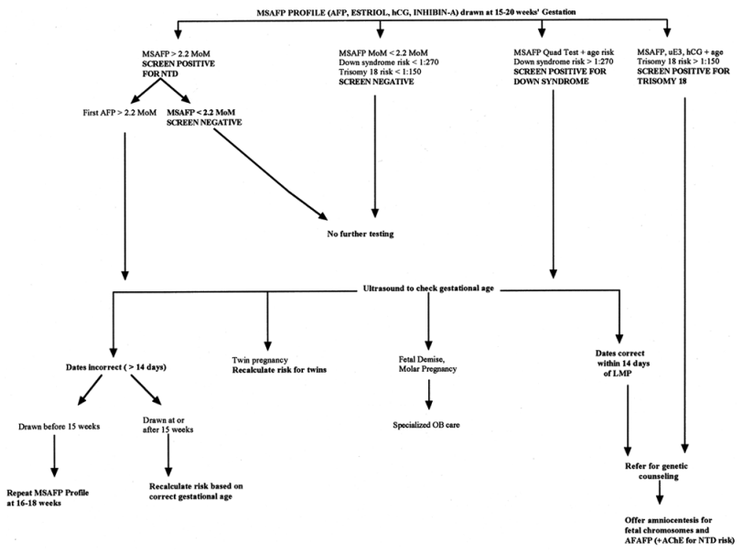


FIGURE 31.1. Protocol used at Henry Ford Hospital for the management of patients with abnormal maternal serum screening results.

TABLE 31.6. REQUIREMENTS FOR A WORTHWHILE SCREENING PROGRAM

Aspect	Requirement
Disorder	Well defined
Prevalence	Known
Natural history	Medically important disorder for which an effective remedy is available
Finance	Cost-effective
Facilities	Available or easily installed
Ethics	Procedures following a positive result are generally agreed on and acceptable to screening authorities and patients
Assay	Simple and safe
Test performance	Distributions of test values in affected and unaffected individuals are known. The extent of overlap is small. A suitable cutoff level is defined.

From Cuckle HS; Wald NJ. Principles of screening. In: Wald NJ, ed. *Antenatal and neonatal screening*. London: Oxford University Press, 1984;1–22, with permission.

lower) limits of normal. Each laboratory must establish its own median for each gestational week between 15 and 20 weeks of gestation. To facilitate interlaboratory communication, AFP values are expressed as MoM. For example, if laboratory A has a median value of 30 IU/mL at 16 weeks and laboratory B's median is 39 IU/mL, both values would be called 1.0 MoM at 16 weeks of gestation.

Determination of the appropriate cutoff for AFP depends on several criteria (Table 31.7). Most laboratories set a cutoff in the range of 2.0 to 2.5 MoM. MSAFP cutoffs in this range usually result in an 85% to 90% detection rate for open spina bifida, with a false-positive rate of 3% to 5%.

TABLE 31.7. PREVALENCE, DETECTION RATE, AND SPECIFICITY IN SCREENING FOR SPINA BIFIDA

The *birth prevalence* of spina bifida varies in a specific geographic region. In the United States, the prevalence is greater in the Northeast than in the Southwest. The prevalence in the United States also varies by ethnic group: approximately 1/700 in whites compared with 1/1,500 in blacks. The overall incidence in the United States is 1/1,000 livebirths.

The *detection rate* (sensitivity): The proportion of affected individuals with a positive test result.

The *false-positive rate*: The proportion of unaffected individuals with a positive test result.

Steps in determining an AFP screening policy:

1. What is the local birth prevalence of open spina bifida?
2. What risk of open spina bifida is acceptable after screening?
3. The likelihood ratio (LR) is determined by considering the factors in 1 and 2.
4. 
5. There is only one AFP cutoff that corresponds to a given LR.

AFP, α-Fetoprotein. From Wald NJ: Principles of screening. *Antenatal and neonatal screening*. London; Oxford University Press, 1984; 1–22, with permission.

Issues in the Interpretation of MSAFP Screening Results

By midtrimester, the primary source of AFP in amniotic fluid is fetal urine. If a fetus has an open body wall defect, additional AFP leaks into the amniotic fluid as a transudate from the fetal serum.

Gestational Age

MSAFP normally increases by approximately 15% per week from 14 to 21 weeks of gestation. Before 15 weeks of gestation, fetal urine is not the primary source of amniotic fluid. After 22 weeks of gestation, the fetal liver gradually switches from production of AFP to production of other serum proteins such as albumin. Therefore, outside this gestational range, AFP is an unpredictable marker for spina bifida. In addition, discrepancies in gestational dating can have a significant impact on AFP interpretation. The first step in interpretation of abnormal AFP results is usually an ultrasound examination to check gestational dates (Fig. 31.1). Ultrasound gestational dating in the second trimester by biparietal diameter (BPD) is usually considered accurate to within 7 to 10 days. Weeks of gestation are calculated through the last completed week (e.g., 16.1 to 16.9 weeks of gestation are within the 16th week). MSAFP interpretations are usually revised if the ultrasound examination findings change the gestational dating by at least 10 to 14 days.

Race

MSAFP values tend to be 12% to 15% higher in black women compared with white, and blacks have a lower prevalence of spina bifida defects. Because of these differences, application of a uniform cutoff without regard to race would label a larger proportion of black patients as elevated, leading to a higher amniocentesis rate in blacks. For screening programs with a significant black patient population, racially pooled samples would also raise the medians for the entire population, and this in turn could lead to a lower detection rate in nonblacks. This problem can be addressed in one of two ways. Some laboratories establish separate medians for black and nonblack patients. Others calculate medians for nonblacks and use a correction factor to provide race-adjusted medians for black patients.

Diabetes

Pregnant women with insulin-dependent diabetes mellitus have an increased risk for birth defects, including open and closed central nervous system defects and congenital heart defects. Some studies suggest a 15% to 20% rate of birth defects compared with the general population incidence of 3% to 5%. For this reason, some programs use a lower cutoff MoM for open neural tube defects in diabetic patients. Many geneticists recommend more thorough diagnostic evaluation in pregnancies of insulin-dependent diabetic women, including amniotic fluid AFP, acetylcholinesterase (AChE), and detailed fetal ultrasound examination. In addition, diabetic mothers have 70% to 80% lower MSAFP levels at each week of gestation. To compensate for this, many laboratories use a correction factor to adjust the MoM.

Twins

MSAFP cutoffs for singleton pregnancies (2.0 to 2.5 MoM) label 56% of twin gestations as elevated (Fig. 31.2). This is not

surprising because each normal fetus contributes approximately 1.0 MoM of AFP to the maternal circulation. Therefore, many laboratories set the upper limit of normal for twin gestations at double the singleton cutoff (4.0 to 5.0 MoM). However, because twins are not usually concordant for spina bifida, a more important screening concern is the expected MSAFP results for discordant twins, i.e., when one twin is affected and one is not. In singleton pregnancies, the median MSAFP value for a fetus with open spina bifida is 3.7 MoM, so screening programs that set their twin cutoff below 4.7 MoM (1.0 MoM + 3.7 MoM) can expect at least a 50% detection rate for discordant twins.

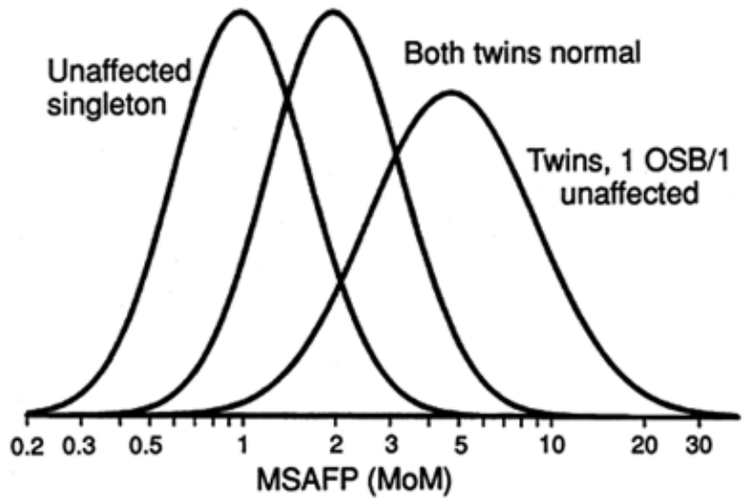


FIGURE 31.2. Distribution of maternal serum α-fetoprotein values for singleton and twin pregnancies. Median multiple of median (MoM) for normal singletons is 1.0 MoM. Median MoM for normal twins is 2.0 MoM. Median MoM for twins—one twin normal and one with open spina bifida (OSB) is 4.8 MoM. (From Palomaki G, Foundation of Blood Research, Scarborough, ME, with permission.)

Procedure-Related Elevations of MSAFP

Both amniocentesis and CVS are associated with fetomaternal bleeding, which causes a transient elevation in MSAFP of 2 to 4 weeks' duration. Genetics programs that obtain blood samples for MSAFP on amniocentesis patients must obtain samples before performing the procedure.

First-trimester CVS or early amniocentesis (11 to 14 weeks of gestation) offers the advantage of earlier fetal chromosomal or biochemical diagnosis, but spina bifida prenatal diagnosis by assay of amniotic fluid AFP is not yet available at that gestational age. Therefore, all such patients should be offered the option of MSAFP screening at 16 weeks of gestation. Approximately 1% of CVS patients have an elevated MSAFP at 16 weeks and require amniocentesis and detailed ultrasound evaluation.

AFP elevations are not specific for neural tube defects. Other open body wall defects, such as gastroschisis and omphalocele, also allow excess AFP transudation. Many other fetal anomalies may also be associated with elevated MSAFP (Table 31.8). Because many of these can be identified prenatally, complete diagnostic evaluation is recommended, including amniocentesis for amniotic fluid AFP, acetylcholinesterase, fetal karyotype, and ultrasound examination.

Multiple Marker Screening for Trisomies

An association between low MSAFP values and chromosome abnormalities was first reported in 1984 (17). The MSAFP median for Down's syndrome fetuses is approximately 0.7 MoM for normal singleton pregnancies. Some reports suggested that trisomy 18, triploidy, and sex chromosome abnormalities also have lower median MSAFP values. The etiology of low MSAFP in chromosomal syndromes is unclear, although placental immaturity has been suggested (18).

Before MSAFP screening, maternal age was shown to be the strongest indicator of risk for Down's syndrome and other trisomies. However, only 20% of Down's syndrome fetuses were born to women age 35 or older. Screening for low MSAFP provides an independent indicator of increased risk for Down's syndrome. Studies showed that approximately 20% of Down's syndrome fetuses had MSAFP values less than 0.4 MoM, whereas only 5% of all MSAFP values fall below this level (19).

Most laboratories have adopted multiple marker screening for Down's syndrome (Table 31.9). The most common combination of markers used is triple screening with MSAFP, estriol (uE3), and human chorionic gonadotropin (hCG). Researchers documented that Down's syndrome fetuses tended to have low MSAFP and uE3 with elevated levels of hCG. The median MoM for a Down's syndrome fetus was 0.7 MoM for MSAFP, 0.7 MoM for uE3, and 2.05 MoM for hCG (21). Using a published algorithm that assigns a likelihood for Down's syndrome based on the multiple markers and the woman's age at EDC, triple screening detects approximately 68% of all Down's syndrome cases when a 1 in 270 cutoff for Down's syndrome is used. In our program, women whose risk for Down's syndrome is at least one in 270 (the mid trimester risk for Down's syndrome in a 35-year-old woman) are offered further diagnostic evaluation including an ultrasound examination to confirm gestational dates and amniocentesis. It remains the standard of obstetric care to offer amniocentesis to all women over age 35 at the EDC. However, many women over age 35 choose to have maternal serum multiple marker screening to reassess their risk for Down's syndrome before deciding whether to have invasive prenatal testing.

Since 1987, investigators have identified several fetoplacental chemicals in maternal serum, which have a predictive association with Down's syndrome. Bogart and co-workers (22) reported a 56% detection rate for chromosomally abnormal fetuses using a simple cutoff of 2.5 MoM for maternal serum hCG. Canick and

co-workers (18,23) found that the median value of maternal serum estriol in 22 Down's syndrome fetuses was approximately 0.79 MoM. Cuckle and co-workers found that the median MoM for dimeric inhibin-A was 1.59 in 19 Down's syndrome fetuses (24).

Most maternal serum screening programs that use multiple markers also screen for trisomy 18. Studies have demonstrated that MSAFP, uE3, and hCG tend to be low when the fetus is affected with trisomy 18. Trisomy 18 occurs in approximately one in 2,400 pregnancies at mid trimester and one in 8,000 pregnancies at birth. The risk for trisomy 18 increases with maternal age and is approximately one fourth of the second-trimester risk for Down's syndrome at mid trimester. By 1990, a screening protocol that used fixed cutoffs of 0.75 MoM for MSAFP, 0.60 MoM for uE3, and 0.55 MoM for hCG had an estimated 60% detection rate for trisomy 18. Subsequently, a risk-based algorithm that included maternal age in the calculation was shown to have a 70% detection rate for trisomy 18. Less than 0.5% of screened pregnancies have values in this range. Our program employs a one in 150 second-trimester risk cutoff for trisomy 18 screening using MSAFP, uE3, and hCG.

Many multiple marker screening programs are adding a fourth maternal serum marker to their screening panel. The addition of maternal serum inhibin-A to the triple screen has been shown to decrease the false-positive rate from 6.6% to 5.0% and increases the detection rate for Down's syndrome from 70% to 75% (20). Inhibin-A is a placental marker that tends to be elevated in Down's syndrome pregnancies. The median MoM for dimeric inhibin-A in a Down's syndrome fetus is 1.95 MoM (25). Our program recently added inhibin-A to the screening panel for Down's syndrome.

Maternal Weight

Shortly after the introduction of low MSAFP screening, genetics centers noted a disproportionately high number of obese patients in the low MSAFP group. This association was owing to dilution of fetal AFP in the greater serum volume in heavier women. Most laboratories now use a weight correction factor for each marker to lessen this bias.

Multiple marker screening for Down's syndrome is also dependent on accurate gestational dating. Very low values are often associated with a gestational age less than 14 weeks, molar pregnancy, or missed abortion. MSAFP values are reinterpreted if an ultrasound examination demonstrates a change in dates of at least 10 to 14 days (Fig. 31.1). Recalculation of screening positive for trisomy 18 results is not recommended based on ultrasound gestational dating because trisomy 18 fetuses often have severe growth retardation.

Amniotic Fluid AFP

Prenatal diagnosis of open spine defects by amniotic fluid AFP analysis has been routinely available since the early 1970s. More than 98% of cases of anencephaly and open spine defects have amniotic fluid AFP levels higher than 2.5 MoM compared with less than 1% of normal fetuses that have values higher than 2.5 MoM.

Other anomalies associated with defects in the fetal skin, such as omphalocele and gastroschisis, also result in elevated amniotic fluid AFP.

False-positive amniotic fluid AFP values may be associated with fetal blood admixture in the amniotic fluid sample. A Kleihauer-Betke test on the sample can confirm the presence of fetal red cells.

Amniotic fluid AChE is a second biochemical marker that is positive in 95% of fetuses with an open neural tube defect (26). The likelihood of an open spine defect is greater than 99.5% when amniotic fluid AFP and AChE are both positive. Some fetuses with ventral wall defects may also have positive AChE results. Wald and co-workers (27) found that the ratio of pseudocholinesterase to AChE was a useful tool to differentiate open spina bifida from ventral wall defects.

Some patients may present with the diagnostic dilemma of elevated amniotic fluid AFP and negative AChE assay. This biochemical pattern may be seen in one third of fetuses with ventral wall defects, which can usually be identified by an ultrasound examination. The most common reason for positive amniotic fluid AFP and negative AChE is fetal hemorrhage into the amniotic fluid, and this pattern is observed in 90% of fetal blood contamination cases because the concentration of AChE is lower in fetal blood than in fetal cerebrospinal fluid. In the remaining 10%, the hemorrhage is so extensive that AChE is also elevated. Positive amniotic fluid AFP associated with negative AChE has also

TABLE 31.8. OTHER CAUSES OF ELEVATED MATERNAL SERUM α-FETOPROTEIN VALUES

Oligohydramnios
Renal anomalies
Bilateral renal agenesis
Congenital nephrosis
Urinary tract obstruction
Infantile polycystic kidney disease
Sacrococcygeal teratoma
Chorioangioma of the placenta
Fetal demise
Chromosome abnormality (e.g., triploidy or autosomal trisomy)

TABLE 31.9. DOWN'S SYNDROME DETECTION AND FALSE-POSITIVE RATE USING COMBINATIONS OF AFP, uE3, hCG, AND INH-A

Maternal Age + Screen Markers	Detection Rate (%)		False-Positive Rate (%)[a]		OAPR (1:n)	
	LMP	BPD	LMP	BPD	LMP	BPD
AFP and hCG	68	71	7.3	7.2	78	74
AFP, uE3, and hCG	70	74	6.6	6.5	69	64
AFP, uE3, hCG, and INH-A	75	78	5.0	5.1	48	48

[a] When gestational age is estimated by last menstrual period (LMP) or fetal ultrasound biparietal diameter (BPD).
Adapted from Haddow JE, Palomaki GE, Knight GJ, et al. Second trimester screening for Down's syndrome using maternal serum dimeric inhibin A. *J Med Screen* 1998; 5:115–119.
OAPR, odds of being affected given a positive result; AFP, α-fetoprotein; hCG, human chorionic gonadotropin; uE3, estriol; INH-A, inhibin-A.

been reported in congenital nephrosis. Most pregnancies with a positive AFP and negative AChE pattern end in the delivery of a normal full-term infant.

Detailed ultrasound evaluation in midtrimester is useful to identify the level of the spinal defect and to detect early changes of fetal hydrocephalus in most cases (Fig. 31.3).

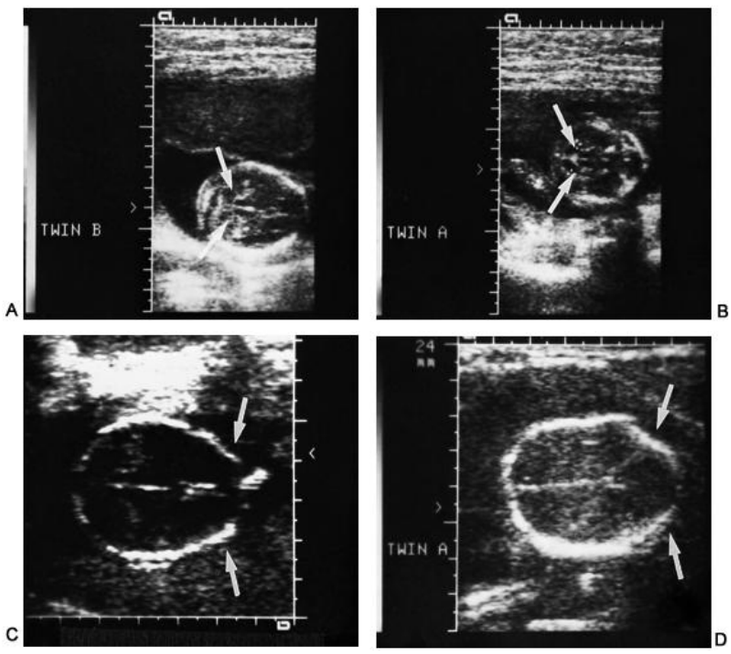


FIGURE 31.3. Comparison of mid-trimester cranial ultrasound findings in fetuses with and without open spina bifida. **A:** Normal transverse scan of cranium at the level of the cerebellum (*arrows*). **B:** Scan at the same level of a fetus with open spina bifida, demonstrating a positive "banana sign," compression of the cerebellum owing to Arnold-Chiari malformation. **C:** Transverse scan of a normal fetus at the level of the lateral ventricles. **D:** Scan at the same level of a fetus with open spina bifida, demonstrating the positive "lemon sign," reflecting a narrowing of the frontal calvarium.

SPECIMEN COLLECTION, CELL CULTURE, AND SLIDE PREPARATION

Part of "31 - Prenatal Cytogenetic Diagnosis"

Amniocentesis is the most commonly used procedure to obtain fetal cells at 14 to 20 weeks of gestation. Most laboratories culture amniocyte cells on coverslips in small Petri dishes and harvest the mitotic cells *in situ* for analysis of chromosomes from cells of distinctly different colonies. Harvesting of amniocyte cultures follows the same principles used in harvesting lymphocytes and other cell types or tissues. Colcemid is added to arrest cell division at metaphase, and treatment with a hypotonic solution disperses chromosomes within the cell membrane. The cells are then fixed *in situ*, chromosome preparations are stained usually using GT banding, and metaphase spread are analyzed microscopically.

CVS is generally performed during the first trimester. The villi can be used to obtain chromosome preparations by direct harvest or monolayer culture methods. By using a direct harvest method, metaphase spreads can be prepared within a few hours after sampling. For monolayer cultures, cells in the chorionic villi are enzymatically and mechanically dissociated, and the cell suspension is used to establish *in situ* cultures. *In situ* harvest methods are similar to those for amniotic fluid cultures and permit analysis of independent colonies of cells.

PUBS provide a fetal tissue source from pregnancies at 20 weeks of gestation or beyond, and up to term depending on the indication for the study. Fetal blood samples drawn directly from the umbilical cord can provide chromosome preparations within 48 hours. Specimens are cultured and harvested using standard lymphocyte culture methods.

Methods to Obtain Chromosome Preparations for Prenatal Testing

Reagents

Chang complete media: 90 mL (Irvine Scientific, Santa Ana, CA) with 14 mL frozen supplement C, 0.25 mL penicillin-streptomycin and 2.0 mL L-glutamine. Store at 4°C for as long as 3 days.

Amnio-Max: 90 mL (GIBCO, Rockville, MD) basal media with vial of 14 mL frozen supplement. Store at 4°C for as long as 3 days and protected from light.

RPMI complete media: 100 mL RPMI 1640 (GIBCO) with 18 mL fetal bovine serum, 1.5 mL L-glutamine, 0.25 mL penicillin-streptomycin. Store at 4°C for as long as 7 days.

CVS media (Chang 60/RPMI 40): 60:40 mix of Chang complete and RPMI complete media. Store at 4°C for as long as 7 days.

CVS transport media: 100 mL RPMI 1640 (GIBCO) with 1.5 mL L-glutamine, 1.0 mL heparin, 5.25 mL fetal bovine serum, and 0.25 mL gentamicin. Store at 4°C for as long as 14 days.

Gentamicin (GIBCO): 10 mg/mL.

L-Glutamine (GIBCO): 200 mmol/L.

Penicillin-streptomycin (GIBCO): 10,000 IU/mL penicillin and 10,000 IU/mL streptomycin.

Sodium heparin: 1,000 USP/mL

Colcemid (GIBCO): 10 µg/mL.

Fetal bovine serum (Hyclone, Logan UT).

Hanks' buffered salt solution (HBSS) (GIBCO) 1×: 100-mL bottle with 0.25 mL gentamicin.

Collagenase V (Sigma, St. Louis, MO): 100 mg in 10 mL HBSS. Filter sterilize. Store at 4°C for as long as 7 days.

Hypotonic: 0.8% sodium citrate or 1.0% sodium citrate.

Fixative: 3:1 anhydrous methanol:glacial acetic acid. Prepare fresh before each use.

Trypsin solution (GIBCO): 0.05% 1× with ethylenediaminetetraacetic acid (EDTA).

Ethidium bromide (SIGMA): 0.3 mg/mL. Warning: mutagenic.

Sterilized 22-mm^2 coverslips are prearranged in sterile 35-mm culture dishes.

All centrifugation steps are performed under the same conditions: 10 minutes at 140*g* (e.g., 800 rpm using a 185-mm radius centrifuge head.)

All tissue culture work is performed in a laminar flow hood using sterile technique. All cultures are incubated in 5% O_2/5% CO_2/90% N_2 incubators. The cultures grow more quickly in a low-oxygen atmosphere, but traditional incubators with 5% CO_2 can be used.

Culture of Amniotic Fluid Cells

1. Two tubes of amniotic fluid (Fig. 31.4), with 20 mL total, are generally sent to the laboratory. If AFP and/or AChE assays are requested, aliquot 1 to 2 mL per test into a sterile tube. Centrifuge both tubes.

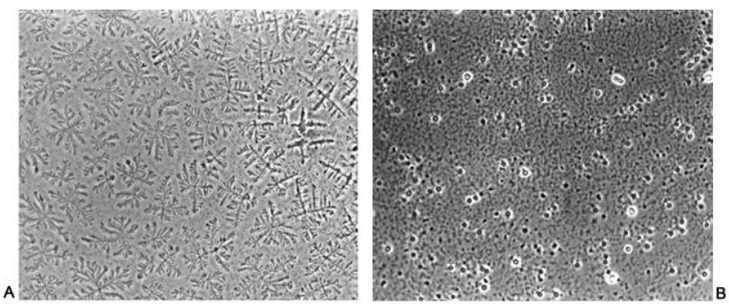


FIGURE 31.4. Urine and amniotic fluid can be distinguished by microscopic examination of a droplet of the fluid spread and dried on a microscope slide. The proteins in amniotic fluid give the appearance of ferning **(a)** that is not observed with urine **(B)**.

2. Remove all but 0.5 mL of supernatant from each tube. Add 2 to 3 mL media to each of the original tubes and resuspend the cells. Note: Cultures are initiated using both Chang complete and Amnio-Max media. Generally, each tube of amniotic fluid received is set up with different media.
3. Distribute the cell suspension from each tube among a total of eight coverslip cultures. When specimens are grossly bloody or when a large number of cells is required for special testing protocols, establish cultures in T-25 flasks. Add another 3 mL media to each flask culture. *Note*: Extreme care must be taken in labeling specimens for prenatal testing. A name and/or identification number should appear on each culture as well as identifying which amniotic fluid tube was used to establish the culture. Analysis should be performed on cells cultured from each tube.
4. Incubate cultures at 37°C.

5. After 12 to 24 hours of incubation, gently add 1 to 1.5 mL prewarmed media to each coverslip, taking care not to dislodge cells that have adhered to the coverslip.
6. On days 3 to 5 of culture, aspirate the media from the culture dish and flask and replenish with 1.5 mL prewarmed media (4 mL media for flask cultures). Maintain flask cultures by replenishing the media every third day with 4 mL prewarmed media. Coverslip cultures should be fed every third day until they are ready to harvest.
7. Cultures are examined on days 4 to 6 of culture and every day thereafter until all cultures are harvested. An inverted microscope is used to evaluate cell growth. A coverslip is ready to harvest when there are three or more feathery colonies with 40 to 65 cells per colony.

Culture of Chorionic Villus Cells

1. A 5- to 15-mg sample of chorionic villi (Fig. 31.5), or more, is transported to the laboratory in transport media. A sample of less than 5 mg of villi is only marginally satisfactory.
2. Transfer the specimen to a large culture dish, and using two tuberculin syringes, dissect away any blood clots and tissue that do not appear to be villus in origin.

Direct Preparation

1. Select six pieces of the most active-looking (thicker, bushlike in appearance) and place the pieces in a 35-mm culture dish with 2 mL of Chang media and incubate overnight.
2. Add 30 µL Colcemid and incubate for 3 hours.
3. Carefully remove the media from the culture dish. Add 2 to 3 mL of 1.0% sodium citrate and incubate at room temperature for 5 minutes.
4. Carefully aspirate the hypotonic solution and add 2 mL of fixative. Incubate at room temperature for 10 minutes.
5. Aspirate the fixative and add 2 mL of fixative. Place the culture dish in a -20°C freezer for 20 minutes. Aspirate the fixative and replace with 2 mL of fixative.
6. Transfer a piece of the villus to the lid of the culture dish and allow to dry briefly. Using a Pasteur pipette, add two to three drops of 30% glacial acetic acid to the villi and allow small bubble to appear on the sample (approximately 20 seconds). The longer the villus is in the 30% acid solution, the worse the morphology but the higher the mitotic index.
7. Pipet the sample and acid mixture to a dry microscope slide on a 42°C hot plate.
8. Place the slide in a 90°C drying oven for 1 hour.
9. Chromosomes are G-banded using trypsin as described in Chapter 29.

Collagenase Culture Initiation

1. Select villi with buds and visible capillaries and rinse several times in HBSS.
2. Transfer villus pieces into a 60-mm culture dish containing 5 mL trypsin-EDTA solution. Swirl the dish gently and incubate at 37°C for at least 2 hours.
3. With sterile forceps, pick each piece and shake vigorously in the trypsin solution. Place the pieces in a 60-mm culture dish containing 0.5 mL Chang media. Using two sterile tuberculin syringes or two scalpels, mince the tissue into small pieces, approximately 1 to 2 mm in diameter. Add 4 mL Chang media and 0.5 to 0.7 mL collagenase. Incubate the collagenase-cell suspension for 2 hours.
4. Resuspend the villus/collagenase mixture vigorously and transfer the mixture to a sterile 15-mL tube.
5. Centrifuge and aspirate the supernatant.
6. Add 5 mL Chang media and mix well.
7. Centrifuge and aspirate the supernatant.
8. Resuspend the cell pellet in 2 mL of fresh CVS media.
9. Establish *in situ* cultures as for amniocytes and maintain the cultures using CVS media. Harvest cultures following protocol for amniotic fluid.

Subculture of Flasks and Coverslips

Subculture of a flask or coverslip culture may be necessary if a culture is needed for noncytogenetic studies or the culture becomes confluent. Harvests from subcultures are usually better quality if carried out either 1 day after feeding or 2 days after subculture. For flask cultures, follow the protocol for subculture of fibroblast cultures (Chapter 30). For coverslip cultures:

1. Remove media and replace with HBSS (3.0 mL for flask, 1.5 mL for coverslip culture). Swirl the culture to rinse and remove the HBSS.
2. Add trypsin EDTA (2 mL for flask, 1 mL for coverslip culture).
3. Incubate at 37°C for 5 minutes.
4. Lightly tap the culture vessel to dislodge some of the cells.
5. Add complete media to inactivate the trypsin and split the culture into one or more vessels.
6. Feed with media on the following day.

Harvest

Sterile technique is not required. For harvest of tissue culture flasks, follow the protocol described in Chapter 30 for fibroblast cultures. For coverslip cultures:

1. Add 50 µL ethidium bromide to each coverslip culture and return it to the incubator for 50 minutes. Add 10 µL Colcemid and return the culture to the incubator for an additional 30 minutes.
2. Using a Pasteur pipette held against the side of the dish, aspirate all of the media, taking care not to touch the coverslip.
3. Gently add 2 mL of hypotonic solution and incubate at room temperature for 23 minutes.
4. Gently add 3 mL of fixative to the dish and incubate at room temperature for 3 minutes.
5. Aspirate most of the liquid from the dish but do not allow the coverslip to dry out.
6. Add 3 mL of fresh fixative and let stand for 20 minutes.
7. Aspirate fixative and repeat fixative washes twice with incubation times of 20 and 10 minutes.
8. After final wash, aspirate the fixative until a thin layer remains

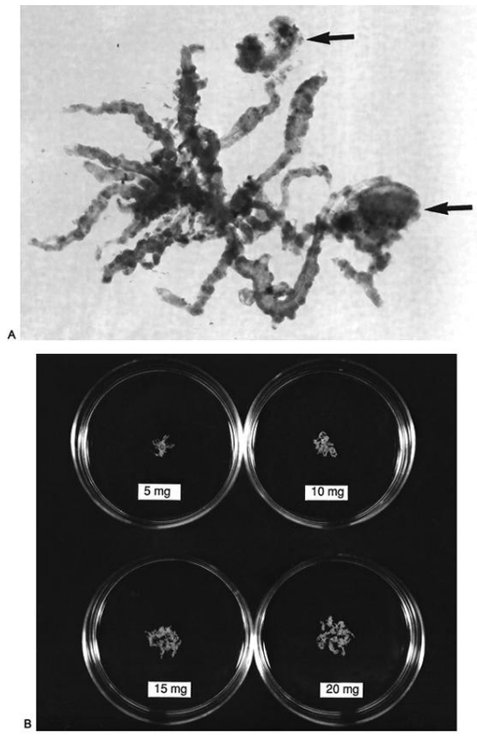


FIGURE 31.5. Chorionic villi **(a)** have a characteristic appearance of fronds of coral, whereas decidua (*arrows*) appear spongiform. **B:** The appearance of 5, 10, 15, and 20 mg of villi in 60-mm Petri dishes.

on the coverslip and the edges around the coverslip begin to dry.

9. Immediately invert the culture dish over an alcohol lamp flame until the cell side of the coverslip is nearly dry. The intensity of heat required for good spreading varies with room humidity.
10. Write the coverslip identification number on the back of the coverslip and place in drying oven at 65°C overnight or at 90°C for 1 hour.
11. Chromosomes are G-banded using trypsin as described in Chapter 29.

Robotic Harvesting of Cultures

In situ cultures can be processed using a robotic harvesting system. This is a time-saving, consistent tool for harvesting and can improve the quality and quantity of metaphases. After incubation in Colcemid, cultures that are ready to harvest are placed on the base of the instrument. The aspiration and dispensing of media, hypotonic, and fixative are all done automatically through a computer-programmed process. The rate of aspiration, amount of solution dispensed, and incubation time between each step can be individually programmed to adapt to each laboratory's individual procedure. After the harvesting procedure is completed, cultures can be placed under an airflow hood to control evaporation of fixative and spreading of the chromosomes.

Culture of PUBS

Lymphocyte cultures are established following the protocol for peripheral blood (Chapter 30). The amount of blood sent is usually 1 mL or less, and whole blood (rather than a buffy coat suspension) is used to inoculate the cultures. Two cultures are established, one to be harvested after 48 hours of incubation and one generally incubated for a standard 66- to 72-hour harvest.

Fluorescent In Situ Hybridization (FISH)

Depending on clinical indication, trisomies 13, 18, 21, and the sex chromosome aneuploidies account for between 85% and 95% of cytogenetic abnormalities associated with mental retardation and birth defects. On this basis, a screening assay to detect aneuploidy of these chromosomes on uncultured amniocyte interphase cells is offered by many cytogenetics laboratories. FISH using DNA probes, either singly or in a cocktail with each probe labeled with a different fluorochrome, is used to screen for aneuploidy by scoring the hybridization signals in interphase cells (Fig. 31.6). An 85% to 95% detection rate is not adequate to employ interphase trisomy screening as a stand-alone test, therefore routine chromosome studies must be done concurrently with interphase trisomy screening.

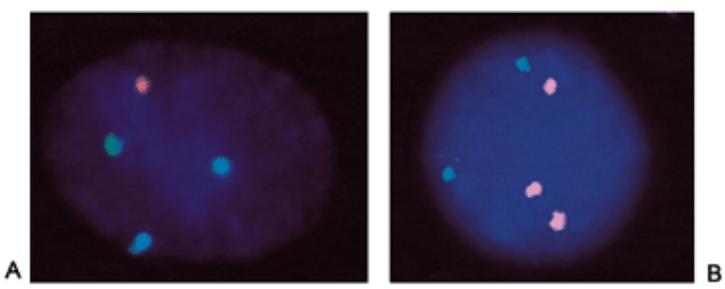


FIGURE 31.6. Fluorescent *in situ* hybridization on uncultured amniocytes. In **A**, using centromere probes for chromosomes X (green), Y (orange) and 18 (aqua), one green and orange signal indicate XY with two aqua signals indicating two copies of chromosome 18. In **B**, region-specific probes for chromosomes 13q (green) and 21q (orange) are used. Three orange signals indicate three copies of chromosome 21 (trisomy 21) and two green signals indicate two copies of chromosome 13.

Including balanced and unbalanced structural rearrangements and chromosomal mosaicism, many of which have clinical implications for other family members as well as for the tested pregnancy, only approximately 65% of chromosome abnormalities are detectable by interphase aneuploidy screening. This weakness, the added cost of the assay, and other technical limitations in aneuploidy screening (28,29) probably account for its narrow acceptance. Although aneuploidy screening is widely available, only approximately 20 laboratories in the United States process more than 50 interphase FISH studies per year. In an era of severe cost constraints in medicine, we discourage most aneuploidy screening by FISH. Nevertheless, situations arise in which the test has added value, such as a late gestational age pregnancy with defects identified on ultrasound examination that are consistent with a known aneuploidy syndrome and in which intervention may be a medical necessity before the results of routine cytogenetic studies become available.

FISH testing using a probe for the critical region associated with velocardiofacial syndrome (deletion 22q11 syndrome) is appropriate for any fetus with a structural heart defect detected on ultrasound. This test can be done after a routine G-banded karyotype analysis if the karyotype is otherwise normal. FISH testing using other individual probes or probe kits can also be useful to characterize rearrangements and marker chromosomes when parental chromosome studies are not informative.

CHROMOSOME ANALYSIS FOR PRENATAL DIAGNOSIS

Part of "31 - Prenatal Cytogenetic Diagnosis"

The analysis of prenatal specimens is done in much the same way as for blood or tissue samples, as described in Chapter 30. The major difference with this specimen type is that cells for analysis can be selected from multiple distinct colonies. Each colony represents an independent sample of a cell or cells taken from the amniotic fluid or chorion. A laboratory ideally should be able to provide a reporting time of 3 to 6 days for PUBS, 6 to 12 days for CVS cultures, and 7 to 14 days for amniotic fluid cell cultures.

For *in situ* harvests, the College of American Pathologists and the American College of Medical Genetics (30) recommend analysis of at least 15 metaphase cells taken from 15 different colonies and representing at least two independent cultures. In general, one cell per colony is analyzed, but other cells in the colony should be analyzed if an abnormality is found. We try to sample the greatest variety of cells to increase the likelihood of finding mosaicism and to reduce the risk of maternal cell contamination. Our decision tree for the number of metaphase cells to study is similar to the revised recommendations of Hsu and

Benn (31) for a standard, moderate, or extensive workup (Table 31.10).

TABLE 31.10. PRENATAL KARYOTYPE ANALYSIS

Number of cells to study		
Routine colony analysis: 15/15/2 (15 cells from 15 colonies from at least two cultures)		
CVS direct preparation: can serve for no more than four of the 15 cells		
Heart defect by ultrasound: 15/15/2 and FISH for 22q11 deletion		
FISH is not needed for echogenic focus or hypoplastic left heart[a]		
Amniotic fluid study after CVS mosaic diagnosis: 30/30/3		
Nonclonal analysis (suspension harvest or subcultures):		
20 cells with at least two cells from a second culture		
Mixed-colony and nonclonal analysis:		
20 cells including as many different colonies as possible		
Looking for mosaicism:		
If additional analysis is required, study colonies from other cultures		
one colony with one or more cells:	then analyze a total of:	and if normal, interpretation is:
Monosomy	20/20/2	Pseudomosaic
Trisomy	30/30/3	Pseudomosaic
45,X,–Y or 45,X,–X	30/30/3	Pseudomosaic
Balanced rearrangement	20/20/2	Pseudomosaic
Unbalanced structural change	20/20/2	Pseudomosaic
+ marker	20/20/2	Pseudomosaic
Tetraploidy	15/15/2	Artifact
One mixed colony		
46/45 any missing chromosome	15/15/2	Random loss
46/47,+13, +18, +21, ±X, ±Y	20/20/2	Artifact
46/other abnormality	15/15/2	Artifact
46/92	15/15/2	Artifact
multiple colonies in only one culture		
Unbalanced structural change	30/30/3	Pseudomosaic
+ marker	30/30/3	Pseudomosaic
Mixed colonies with 46/92	15/15/2	Artifact
Tetraploidy, pure	30/30/3	May require follow-up
multiple colonies from more than one culture		
46,XY/46,XX	15/15/2 46,XY	Maternal cell admixture, chimerism is very rare
Any other result	30/30/3	True mosaic, may consider UPD studies

These guidelines are used in our laboratory and are provided as an example of protocol development to make the most efficient use of personnel working at the microscope and to assist the staff in the preliminary interpretation of findings that are usually not clinically significant. All abnormal cells are also brought to the attention of the person responsible for clinical interpretation of the karyotype findings.
CVS, chorionic villus sampling; FISH, fluorescent *in situ* hybridization; UPD, uniparental disomy.

[a] Thomas, JA, Graham JM Jr. Chromosome 22qll deletion syndrome: an update and review for the primary pediatrician. *Clin Pediatr* 1997;36:253–266.

For CVS, reliance on direct chromosome preparations alone is not recommended. Cultured cells or pooled analysis from direct preparations and cultures is preferable. If both methods are employed, at least 15 cells should be analyzed, with at least 10 cells coming from the culture harvests.

It is important to confirm abnormal prenatal diagnosis results after birth or pregnancy interruption. The parents very often benefit from certain knowledge of a correct result, and the practice serves as a powerful quality-control device for the cytogenetics laboratory. For a urea termination, fetal tissues are usually nonviable in tissue culture, but an amniotic fluid sample can be obtained before instillation of the urea, and chorionic villi obtained after delivery are also usually viable.

CYTOGENETIC RESULTS

Part of "31 - Prenatal Cytogenetic Diagnosis"

Growth Failure

For amniotic fluid samples, growth failure is rare in an experienced laboratory. Aside from technical catastrophe, sources of growth failure include maternal urine or hygroma fluid submitted as amniotic fluid or heat or cold damage during transport of the specimen to the laboratory. A bloody amniotic fluid sample may grow more slowly, but a result can usually be expected. First-trimester CVS growth failure should be well under 1% for specimens that contain 5 mg or more of chorionic villi and have visible blood vessels (32,33). The amniotic fluid cell culture failure rate should be 0.1% or less.

Maternal Cell Contamination

A few mitotic cells of maternal origin are observed in approximately 0.2% to 0.5% of amniotic fluid cultures. We request at least two aliquots of fluid and find that maternal cell contamination (MCC) is usually confined to cultures established from the first tube collected during amniocentesis. Some laboratories prefer to discard the first 2 mL of amniotic fluid. MCC is usually not an interpretive problem because in almost every recognized case, the majority of colonies in the culture are not maternal, and diagnostic errors (wrong sex or missed chromosome abnormality) are rare. Clusters of MCC can be caused by an inexperienced person performing the amniocentesis procedure.

In CVS direct preparations, the very low rate (less than 0.1%) of MCC is owing to the relative absence of mitotic activity in the decidual fragments. The experience varies among laboratories, but MCC is identified as a 46,XX/46,XY admixture in 0.5% to 2% of CVS cultures. Undoubtedly, there is an equal frequency of unrecognized admixtures of maternal and 46,XX fetal cells. The maternal cells are usually in the minority and have caused few cytogenetic errors in experienced laboratories, although their presence can be problematic for biochemical or DNA assays.

The theoretical concern is that a mixture of 46,XX and 46,XY may represent a true chimera. The incidence of chimeras has been estimated to be one in 60,000 newborns. Weighing this incidence against the rate of 46,XX/46,XY cell admixture in prenatal genetic cell cultures, the likelihood of MCC in any individual case seems very high, even if 46,XX cells are seen in multiple culture vessels. When the 46,XX cells are confined to one or a minority of the cultures, they almost certainly represent MCC. When an XX/XY admixture is identified, some laboratories raise the question of chimerism and some do not. Exceedingly few true chimeras identified in prenatal genetic studies have been described in the literature, and it is likely that a chimera presenting at prenatal diagnosis will not be identified as such. Often, the XX/XY admixture is identified in cultures established from the same tube of amniotic fluid, making MCC the most likely interpretation. Other possible sources of XX/XY admixture include cross-contamination of samples, twins, or a resorbed twin. An ultrasound examination for fetal gender may be of value in selected cases.

Chromosome Rearrangement

If a chromosome rearrangement is identified, parental karyotype studies are indicated because the risk to the fetus depends on whether the rearrangement is balanced and whether it is familial (34,35) (Table 31.5 and Table 31.11). A balanced familial translocation or inversion does not confer a significantly increased risk for birth defects or mental retardation nor does a familial extra marker chromosome. A duplication or deletion involving euchromatin generally results in mental retardation with or without birth defects. A familial robertsonian translocation involving chromosome 14 or 15 appears to carry a small increased risk of uniparental disomy (UPD) (40; LG Shaffer, personal communication, 1999).

TABLE 31.11. ESTIMATE OF THE INCREASED RISK OF MENTAL RETARDATION (WITH OR WITHOUT BIRTH DEFECTS) ASSOCIATED WITH CHROMOSOME REARRANGEMENTS

Type of Rearrangement	Risk Estimate (%)
Unbalanced rearrangement	Nearly 100
Balanced rearrangement	
Familial	0
Robertsonian involving 14 or 15	˜0.5, test for UPD
De novo	
Robertsonian translocation	0–1
Reciprocal translocation	3–15
Inversion	6–20
Extra marker, familial	Near 0
Extra marker, *de novo*	
Not satellited	15–28
Monosatellited	7–20
+i(12p) or +i(18p)	100
Bisatellited	
Monocentric	Probably <5
Dicentric, 1 G-light band	Probably <10
Dicentric, >1 G-light band	High risk

The background risk is estimated to be 3%, so total risk is the sum of the two risk estimates. The total frequencies of each class of abnormality are summarized in Table 31.4. Statistics are from many sources, including Crolla et al. (36), Hsu (37), Hsu and Perlis (38), Warburton (34), Steinbach et al. (39), and Van Dyke, unpublished data.

UPD, uniparental disomy.

Extra Marker Chromosomes

For an extra marker chromosome (also termed extra small accessory chromosome), whether mosaic or not, the karyotype of each parent needs to be studied with specific attention to the presence of a similar marker in a small proportion of their blood cells (35,39). Approximately 60% of such markers identified in our program are familial. If the marker is a small acrocentric, consider the possibility that one parent carries the relatively common t(11;22) translocation associated with cat's-eye syndrome (Table 31.5 and Table 31.11 and Fig. 30.7).

If the marker is not familial, its chromosomal origin should be determined to the extent possible. Approximately 85% of extra markers are satellited, and most of these are derived from chromosome 15 or 22. If the marker is satellited, FISH testing using a beta-satellite probe (for the heterochromatic region of the acrocentric chromosomes) and alpha-satellite probes for each of the acrocentric chromosomes can be performed to determine whether it is mono- or bisatellited. If the marker is bisatellited and either monocentric or dicentric with only a tiny amount of G-band light-staining material between the centromeres, the empirical risk to the fetus is small (Table 31.11). If the marker is

derived from chromosome 15, additional FISH and UPD studies may be required to estimate the risks to the fetus (40,41,42 and 43). If the marker is small and derived from chromosome 14 or 22, it may be important to consider testing for UPD of chromosome 14 and for duplication of proximal 22q (cat eye syndrome, see Chapter 30).

If the marker is not familial and is nonsatellited (approximately 15% of extra markers), FISH techniques such as M-FISH or 23-chromosome centromere kits may be used experimentally to identify its chromosomal origin. At least an X and Y centromere region probe should be considered, as approximately 25% are X or Y derived. A small ring chromosome can easily be distinguished from a rod chromosome using a routinely stained (nonbanded) preparation. A ring chromosome will have ring shape; even a tiny ring chromosome will look like a ring in at least some cells. Many cytogeneticists are using FISH techniques to characterize these small marker chromosomes to improve genetic counseling and to improve our understanding of the associated risks (44,45). Several extra markers have been associated with UPD of the normal homologs, so UPD testing is probably indicated if the chromosomal origin of the marker can be established.

Polyploidy

A diagnosis of tetraploidy or mosaic tetraploidy should be confirmed by an independent specimen and follow-up detailed ultrasound examination, unless defects were observed by ultrasound. Tetraploidy is a common artifact of harvest procedures in monolayer cultures. Triploidy is an occasional finding in prenatal studies, and, when seen in an amniocentesis culture, unambiguous abnormalities are usually evident on ultrasound examination.

Autosomal Monosomy Mosaicism

Rare cases of true mosaicism for monosomy 21 and monosomy 22 have been reported. A diagnosis of monosomy mosaicism should be considered if multiple cultures have colonies with the same monosomy. The diagnosis is sufficiently uncommon that detailed ultrasound and a confirmatory study (amniocentesis or PUBS) can be considered.

Trisomy 20 Mosaicism

Approximately one in 2,500 amniotic fluid cell cultures reveals true trisomy 20 mosaicism. Many cases have been confirmed in fibroblast cultures from various tissue sources from the fetus or newborn, including urine sediment cells (originating from the bladder and kidney) and extrafetal membranes, although almost never from peripheral blood cultures. The clinical significance of this finding is still poorly understood; however, a trisomy 20 syndrome has not been described, but approximately 6% of cases have an abnormal outcome, and among cases with more than 50% of cells with trisomy 20, the risk appears to be closer to 20% (46). Most newborns with this prenatal diagnosis have had a normal phenotype at birth and during early childhood. More long-term follow-up studies are needed to provide better information to parents.

Sex Chromosome Aneuploidy

The phenotypes of the nonmosaic sex chromosome abnormalities 45,X; 47,XXX; 47,XXY; and 47,XYY are described in Chapter 30 and were reviewed by Robinson et al. (47). The incidences

TABLE 31.12. PRENATAL EFFECTS OF CONFINED PLACENTAL MOSAICISM

Chromosome	Risk
2	Moderate for IUGR
3	Moderate for IUGR
7	Low
8	Low
9	Moderate for IUGR
13	Moderate to high for IUGR
15	Moderate for IUGR
16	High for IUGR or IUFD
18	Moderate for IUFD
20	Low
22	Moderate to high for IUGR or IUFD
X	Low
Tetraploid	Low

IUGR, intrauterine growth retardation; IUFD, intrauterine fetal demise.
From Leschot NJ, Schuring-Blom GH, Van Prooijen-Knegt AC, et al. The outcome of pregnancies with confined placental chromosome mosaicism in cytotrophoblast cells. *Prehat Diagn* 1996;16:705–712.

of 47,XXY and 47,XXX increase with advancing maternal age, and together they constitute nearly 10% of the chromosome abnormalities we have detected prenatally.

45,X/46,XY Mosaicism

Determine whether the 45,X cell population is secondary to the presence of a cell population with an unstable Y-chromosome rearrangement such as a ring or dicentric. Such karyotypes represent a distinct and clinically important diagnostic group.

In our experience, simple 45,X/46,XY mosaicism constitutes approximately 2% of all abnormalities found prenatally. In many of the first X/XY cases, the pregnancies were terminated because of the known association with Ullrich-Turner syndrome, mixed gonadal dysgenesis, and ambiguous genitalia. However, because most of the fetuses had normal-appearing male external genitalia, the parents of most of the more recent cases have continued the pregnancies. Although further research is required, the risk of abnormal sexual differentiation in prenatally diagnosed 45,X/46,XY is 5%. Detailed ultrasound examination of the external genitalia can be useful. The mosaicism is occasionally confirmed in fetal or extrafetal tissues. Confirmatory cytogenetic studies and long-term clinical follow-up are encouraged (48).

45,X/46,XX Mosaicism

This prenatal diagnosis requires further study, but in the absence of ultrasound evidence of a cystic hygroma or nonimmune hydrops, there may be little risk of fetal death or an Ullrich-Turner syndrome phenotype. As with 45,X/46,XY, this result appears to be consistent with a normal phenotype in most instances. The risk of isolated gonadal dysgenesis without an Ullrich-Turner phenotype is uncertain in this group. Many have normal puberty and fertility, although there is a suggestion of an increased risk of miscarriage or early menopause, so the 45,X/46,XX mosaic would want to take this into consideration in her own family planning. As with other prenatally diagnosed mosaics, cytogenetic confirmation and long-term follow-up are encouraged.

Prenatal Diagnosis of Mosaicism

Chromosomal mosaicism is one of the most academically interesting areas in clinical cytogenetics. Because of its clinical significance, the diagnosis of mosaicism has been particularly challenging in prenatal cytogenetic studies (38,49,50 and 51). When an abnormality is found in some but not all cells, the clinical significance is not always clear, and several possibilities need to be considered, including MCC, true fetal mosaicism, confined placental mosaicism (CPM), and cell culture artifact.

No cytogenetic techniques can identify all mosaics. Amniotic fluid cell karyotypes exhibit mosaicism in approximately 0.3% of studies, and the majority of these can be confirmed by analysis of other tissues. In practical terms, true mosaicism is considered to be confirmed only if the abnormality is identified in more than one culture vessel or preparation. Mosaicism limited to a single culture vessel is termed *pseudomosaicism* (52). Based on the high frequency of pseudomosaicism (approximately 4% of all amniotic fluid karyotype studies), and the very small number of reported false-negative results, pseudomosaicism probably reflects true fetal mosaicism in fewer than 1% of cases (53,54 and 55). Although it is possible to garner additional data using FISH (56), the additional data are not always easy to interpret, and we are not aware of any laboratories that routinely employ FISH when pseudomosaicism is detected.

Many laboratories use different protocols depending on whether the possible mosaicism is consistent with viability. However, because most rare autosomal aneuploidies have been reported as mosaics in liveborn infants (Table 30.2) (57), we treat almost all cases of possible mosaicism in much the same way (Table 31.10).

Chorionic villus cell karyotype studies exhibit true mosaicism in approximately 2% of studies, but only approximately 10% of these represent true mosaicism in the fetus (58). Most of the remainder reflect mosaicism that is confined to the placental tissues (59). CPM is observed more often in direct preparations of cytotrophoblast cells than in preparations from cultured cells from the villus mesenchyme (50). The rate of true mosaicism interpreted only as pseudomosaicism is probably much lower than for amniotic fluid cell cultures.

CPM can be clinically significant in two principal ways. First, many CPM mosaics identified with CVS probably arise as trisomic conceptions. Fetal viability is salvaged by a compensating mitotic nondisjunction (trisomy rescue) that produces a karyotypically normal cell population, and, as discussed in the previous chapter, trisomy rescue can result in clinically significant UPD.

Second, CPM for some chromosomes is associated with an increased risk of IUGR or intrauterine fetal demise (IUFD) (60). The specific risks are not well understood, but CPM can be confirmed in at least 35% of CVS-identified mosaics that are later associated with IUGR (61). Likewise, detailed cytogenetic studies of term placentas from IUGR pregnancies indicate that at least 5% of isolated IUGR is associated with CPM (62,63). The risks differ depending on the chromosome involved, with CPM trisomy 9, 16, and 22 apparently having the greatest risks of severe IUGR and IUFD. A meiotic origin of the trisomy appears to confer a greater risk than a mitotic origin (64). When CVS-identified CPM is associated with fetal malformations, it seems likely that UPD or true fetal mosaicism is also present. More studies are needed to understand fully the effects of CPM, UPD, and tissue-limited fetal mosaicism on placental insufficiency and fetal demise (Table 31.12).

Among nonmosaic trisomy 13 and 18 liveborn infants, placental karyotype studies indicate that CPM with chromosomally normal placental cells contributes to the viability of these two nearly lethal trisomies (65).

Protocol for Mosaicism

Several protocols have been proposed to distinguish true mosaicism from pseudomosaicism. Because we employ an *in situ* harvest method for both amniotic fluid and chorionic villus cell cultures, our protocol for dealing with possible mosaicism is similar for both cell types (Table 31.10). Stated briefly, true mosaicism is likely only when abnormal colonies are detected in independent preparations.

Mixed Colony

If one colony has a mixture of abnormal and normal cells, and no other colony shows that abnormality, the abnormal cells are usually considered an *in vitro* artifact.

Most amniotic fluid cell colonies represent single clones, but most CVS colonies do not. A colony can arise from several cells attached together in the original inoculum; cells can be dislodged during handling and feeding of cultures; and the cells do have limited mobility, so they can infiltrate nearby colonies or establish new, apparently independent colonies. Furthermore, independent clones can form a single colony as they expand. We have observed mosaicism limited to one side of a large colony, with normal cells on the other side. This might be interpreted as a mixed colony, but examination of the clonal growth patterns under lower magnification can sometimes reveal that there were originally two separate colonies.

Pure Colony

If one colony is abnormal, we usually attempt to score five to 15 additional colonies for that abnormality, particularly colonies in other cultures. In most cases, the additional cells scored do not have the same abnormality, and the abnormal colony is interpreted as a probable artifact of no clinical significance (pseudomosaic).

Single Abnormal Cell

If a colony has only one metaphase, its karyotype is abnormal and the abnormality is not consistent with a viable mosaicism (Table 30.2), it is reasonable to interpret it as an artifact (pseudomosaic) (67). A single hypodiploid cell (even 45,X) almost always represents an artifact of culture or slide preparation. However, if the defect is consistent with a viable numerical or structural mosaicism, we follow the same protocol as for a pure abnormal colony, and score five to 15 additional colonies for that abnormality (Table 31.10).

Multiple Pure Colonies in One Culture

If multiple colonies have the same chromosome abnormality, all from the same culture vessel, we attempt to score 15 or more additional colonies from other culture vessels. In actual practice, the technologists often score all colonies from all available cultures. It seems highly unlikely that a mutation in cell culture could produce more than two or three abnormal colonies. In an amniotic fluid cell culture, as well as in a CVS culture, multiple abnormal colonies may reflect CPM with fragments of cytogenetically abnormal chorionic villi having been introduced into just one culture vessel. If the mosaicism is not confirmed in any other cultures, the finding is usually interpreted as an artifact (pseudomosaic) or confined placental mosaicism.

Although follow-up at birth is encouraged, aggressive prenatal follow-up studies, such as repeat amniocentesis, or amniocentesis after CVS for analysis of another tissue, are generally not done if the mosaicism is limited to a single culture vessel. Such aggressive follow-up would increase the risks to the fetus without ensuring an accurate cytogenetic diagnosis. If all the colonies in one culture have the same abnormal karyotype, but no colonies from other cultures have it, one must consider the possibility of mixed-up samples.

Multiple Mixed Colonies in Multiple Cultures

This situation is very unusual but may raise questions of an unstable chromosome, such as a ring or dicentric, or of a chromosomal instability syndrome, which would need further investigation. Tetraploid and diploid cells appearing together within multiple colonies probably reflect a tissue culture artifact.

Chromosomal fragile sites, other than the fragile site at Xq27.3, are generally considered to be of no clinical significance but can present themselves in this way (Fig. 31.7). Region 1q11–q12 is a C-band positive variable region, also known as 1qh or, in the older literature, the “uncoiler locus.” In CVS preparations, 1qh is true to its original moniker because in some individuals, it uncoils dramatically (Fig. 31.7). This behavior is not observed to the same extent in routine preparations of other cell types and is of no known clinical significance when observed in a CVS preparation.

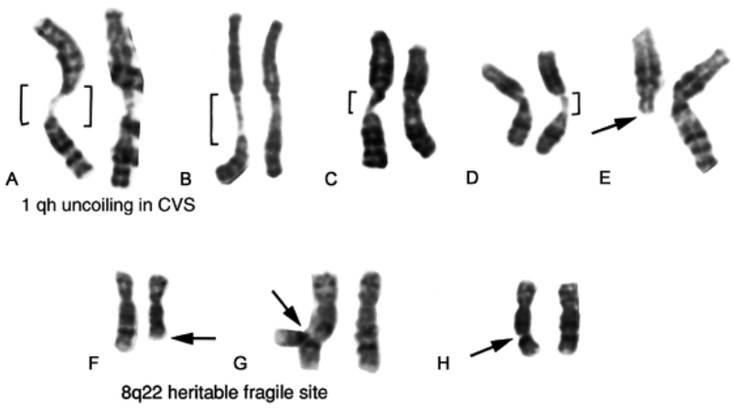


FIGURE 31.7. The uncoiler locus and the fragile site at 8q22. Chromosome 1 uncoiling **(A–D)**, usually not seen in chromosome preparations from other cell types, is a common artifact in CVS preparations. Both homologs (as in **A**) are frequently involved, and in occasional cells **(E)**, an entire arm is lost. A chromosomal fragile site can give the appearance of chromosomal mosaicism by producing multiple colonies with the same abnormality, yet have no clinical significance. The case depicted here had three colonies with one or two cells **(F)** having a del(8)(q22); the remaining cells in each colony had a normal karyotype. One cell **(G)** had a triradial configuration at 8q22. Many other cells had a chromatid break or chromosome break **(H)** at the same site. This finding was interpreted as a heritable fragile site with a high rate of chromosome breakage located at the common fragile site 8q22.

Multiple Culture Vessels

If an abnormality is observed in multiple colonies grown in more than one culture vessel, the interpretation is true mosaicism. Genetic counseling is indicated, and follow-up depends on the cytogenetic abnormality and on whether the mosaicism is identified in CVS or amniotic fluid cell cultures.

Follow-up Studies for Amniotic Fluid Cell Mosaicism

For true mosaicism in an amniotic fluid study, a confirmatory prenatal study is usually not undertaken. A follow-up study and a detailed ultrasound examination of the fetus may be appropriate when (a) the physicians and family agree to confirm mosaicism (or other abnormal result) when the family's decision would be to consider interrupting the pregnancy only if the abnormality

is confirmed, (b) only a single culture is available or an uncertain result is obtained from a very small sample of amniotic fluid or villi, (c) the quality of the cytogenetic preparation is below usual standards and has led to an uncertain result, and (d) a 46,XX and 46,XY admixture in multiple cultures is observed (raising the possibility of chimerism).

A normal result on confirmatory studies cannot rule out mosaicism nor ensure a normal outcome of the pregnancy. Even if mosaicism is confirmed by amniocentesis or CVS, the clinical interpretation is subject to uncertainties about the risk to the fetus and the likelihood of the mosaicism in the fetus. In some circumstances, tissue-limited mosaicism is possible (e.g., trisomy 20 or extra marker chromosome), so a PUBS is not always an appropriate sample for confirming mosaicism.

There are many unknowns in the clinical interpretation of mosaicism, *de novo* balanced rearrangements, and extra marker chromosomes. For this reason, we urge all groups involved in prenatal diagnosis to participate in cytogenetic and clinical follow-up studies, including long-term clinical follow-up, and cytogenetic studies of placenta, amnion, circumcision skin, and other tissues when possible.

In low-level true mosaicism, the clinical risks are not always clear, and follow-up using PUBS and detailed ultrasound examination may be informative (46). UPD studies may be indicated to clarify clinical risks for low-level mosaicism for some abnormalities, such as trisomy 15 (Table 31.13).

TABLE 31.13. UNIPARENTAL DISOMY AND PHENOTYPE

Chromosome	Effect of Maternal UPD	Effect of Paternal UPD
1	None	None
2	IUGR?	No cases
3		No cases
4	None	No cases
5	No cases	None
6	None	Transient neonatal diabetes?
7	IUGR, postnatal GR, Russell-Silver syndrome	None
8	None	None
9	None	No cases
10	None	None
11	No cases	Beckwith-Wiedemann syndrome?
12	No cases	No cases
13	None	None
14	Short stature, precocious, puberty, other features	Severe MR and skeletal abnormalities
15	Prader-Willi syndrome	Angelman syndrome
16	IUGR, birth defects in some?	None
17	No cases	No cases
18	No cases	No cases
19	No cases	No cases
20	None	None
21	None	None
22	None	None
XX	None	Short stature?

The clinical importance of uniparental disomy (UPD) is discussed in greater detail in Chapter 30. Also discussed in Chapter 30 are Russell-Silver, Beckwith-Wiedemann, Prader-Willi, and Angelman syndromes. Listed here are the reported effects of UPD for individual chromosomes (57,62,64,67,68 and 69). The conclusions are often based on very small sample size (sometimes only one case), and in some instances autosomal recessive disease or tissue-limited mosaicism in the fetus cannot be excluded. For most chromosomes, information is based on individual case reports, so the magnitude of the risk of UPD is not known. No studies have been conducted to determine the frequency and clinical effects of UPD in unselected populations, so the risks are uncertain but not negligible (70). Several cases of UPD15 have been diagnosed prenatally after detecting trisomy 15 mosaicism in amniotic fluid or chorionic villus sampling cell cultures (40,71).
None, no clinical effect suspected; no cases, none has been identified as yet; MR, mental retardation.

Follow-up Studies for CVS Mosaicism

Mosaicism found in a CVS requires follow-up. This usually includes amniotic fluid cell karyotype studies. If mosaicism is detected in a direct preparation but not in the cultures, there is approximately a 7% chance that it will be confirmed (58). If the mosaicism is detected in CVS cell cultures, approximately 11% are confirmed, and if both a direct preparation and cultures are mosaic, approximately 21% are confirmed by amniocentesis. Mosaicism for trisomy 13, 18, or 21 is confirmed in approximately 19% of cases, sex chromosome aneuploidy in approximately 15%, extra marker chromosomes in approximately 25%, and all others in fewer than 10% of cases. A normal amniocentesis result indicates that the fetus is most likely not mosaic, although CPM is likely and in some cases this confers a risk of IUGR or IUFD (Table 31.12). If the mosaicism involves a chromosome that is imprinted, UPD studies may be indicated (Table 31.13).

REFERENCES

1. Stranc LC, Evans JA, Hamerton JL. Chorionic villus sampling and amniocentesis for prenatal diagnosis. *Lancet* 1997;349:711–714.

2. Milunsky A, ed. *Genetic disorders and the fetus*, 4th ed. Baltimore: The Johns Hopkins University Press, 1998.

3. Brambati B, Kuliev A, Jackson L, et al. Risk evaluation in chorionic villus sampling. Report of a WHO consultation on first trimester fetal diagnosis, 1 June 1985. *Prenat Diagn* 1986;6:451–456.

4. Lindsten J, Marsk L, Mikkelson M, et al. Role of chorion villi biopsy in prenatal diagnosis of genetic disorders. *Prog Clin Biol Res* 1985; 177:195–212.

5. Rhoads GG, Jackson LG, Schlesselman SE, et al. The safety and efficacy of chorionic villus sampling for early prenatal diagnosis of cytogenetic abnormalities. *N Engl J Med* 1989;320:609–617.

6. Canick JA, Palomaki GE, Osanthanondh R. Prenatal screening for trisomy 18 in the second trimester. *Prenat Diagn* 1990;10:546–548.

7. Hook EB. Cytogenetics: incidence, risks, and recurrence of chromosome abnormalities. In: Brock DJH, Rodeck C, Ferguson-Smith MA, eds. *Prenatal diagnosis and screening*. London: Churchill-Livingstone, 1992:351–392.

8. Eydoux P, Choiset A, Le Porrier N, et al. Chromosomal prenatal diagnosis: study of 936 cases of intrauterine abnormalities after ultrasound assessment. *Prenat Diagn* 1989;9:255–268.

9. Snijders RJM, Nicolaides KH. *Ultrasound markers for fetal chromosomal defects*. London: The Parthenon Publishing Group, 1995.

10. Stoll C, Alembik Y, Roth MP, et al. Study of 224 cases of oligohydramnios and congenital malformations in a series of 225,669 consecutive births. *Commun Genet* 1998;1:71–77.

11. Seeds JW. Ultrasonographic screening for fetal aneuploidy. *N Engl J Med* 1997;337:1689–1690.

12. Thomas JA, Graham JM Jr. Chromosome 22q11 deletion syndrome: an update and review for the primary pediatrician. *Clin Pediatr* 1997;36:253–266.

13. Nyberg DA, Mack LA, Hirsch J, et al. Abnormalities of fetal cranial contour in sonographic detection of spina bifida: evaluation of the “lemon” sign. *Radiology* 1988;167:387–392.

14. Brock DJH, Sutcliffe RG. Alpha-fetoprotein in the antenatal diagnosis of anencephaly and spina bifida. *Lancet* 1972;2:197.

15. Cuckle HS, Wald NJ. Principles of screening. In: Wald NJ, ed. *Antenatal and neonatal screening*. London: Oxford University Press, 1984:1–22.

16. Wald NJ, ed. *Principles of screening. Antenatal and neonatal screening*. London, Oxford University Press, 1984:12–22.

17. Merkatz IR, Nitowsky HM, Macri JN, et al. An association between low maternal serum alpha-fetoprotein and fetal chromosome abnormalities. *Am J Obstet Gynecol* 1984;148:886–892.

18. Canick JA. Screening for Down syndrome using maternal serum alpha-fetoprotein, unconjugated estriol, and HCG. *J Clin Immunol* 1990;13:30–33.

19. Palomaki GE, Haddow JE. Maternal serum alpha fetoprotein, age, and Down syndrome risk. *Am J Obstet Gynecol* 1987;156:460–463.

20. Haddow JE, Palomaki GE, Knight GJ, et al. Second trimester screening for Down's syndrome using maternal serum dimeric inhibin A. *J Med Screen* 1998;5:115–119.

21. Wald NJ, Cuckle HC, Densem JD, et al. Maternal serum screening for Down's syndrome in early pregnancy. *BMJ* 1988;297:883–887.

22. Bogart MH, Pandian MR, Jones OW. Abnormal maternal serum chorionic gonadotropin levels in pregnancies with fetal chromosome abnormalities. *Prenat Diagn* 1987;7:623–630.

23. Canick JA, Knight GJ, Palomaki GE, et al. Low second trimester maternal serum unconjugated oestriol in pregnancies with Down's syndrome. *Br J Obstet Gynaecol* 1988;95:330–333.

24. Cuckle HS, Holding S, Jones R. Maternal serum inhibin levels in second trimester Down's syndrome pregnancies. *Prenat Diagn* 1994;14:387–390.

25. Lambert-Messerlian GM, Canick JA, Palomaki GE, et al. Second trimester levels of maternal serum inhibin A, total inhibin, σ-inhibin precursor, and activin in Down's syndrome pregnancy. *J Med Screen* 1996;3:58–62.

26. Burton BK. Positive amniotic fluid acetylcholinesterase: distinguishing between open defects of the neural tube and ventral wall defects. *Am J Obstet Gynecol* 1986;155:984–986.

27. Wald NJ, Barlow RD, Cuckle HS, et al. Ratio of amniotic fluid acetylcholinesterase as an antenatal diagnostic test for exomphalos and gastroschisis. *Br J Obstet Gynaecol* 1984;91:882.

28. Bryndorf T, Christensen B, Vad M, et al. Prenatal detection of chromosome aneuploidies by fluorescence in situ hybridization: experience with 2000 uncultured amniotic fluid samples in a prospective preclinical trial. *Prenat Diagn* 1997;17:333–341.

29. Lewin P, Kleinfinger P, Bazin A, et al. Defining the efficiency of fluorescence in situ hybridization on uncultured amniocytes on a retrospective cohort of 27407 prenatal diagnoses. *Prenat Diagn* 2000;20:1–6.

30. American College of Medical Genetics. *Standards and guidelines: clinical genetics laboratories*, 2nd ed. Bethesda, MD: American College of Medical Genetics, 1999.

31. Hsu LYF, Benn PA. Revised guidelines for the diagnosis of mosaicism in amniocytes. *Prenat Diagn* 1999;19:1081–1090.

32. Ledbetter DH, Martin AO, Verlinsky Y, et al. Cytogenetic results of chorionic villus sampling: high success rate and diagnostic accuracy in the United States collaborative study. *Am J Obstet Gynecol* 1990;162:495–501.

33. Ledbetter DH, Zachary JM, Simpson JL, et al. Cytogenetic results from the U.S. collaborative study on CVS: high diagnostic accuracy in over 11,000 cases. *Prenat Diagn* 1992;12:317–345.

34. Warburton D. Outcome of cases of de novo structural rearrangements diagnosed at amniocentesis. *Prenat Diagn* 1984;4:69–80.

35. Warburton D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet* 1991;49:995–1013.

36. Crolla J. FISH and molecular studies of autosomal supernumerary marker chromosomes excluding those derived from chromosome 15: II. Review of the literature. *Am J Med Genet* 1998;75:367–381.

37. Hsu LYF. Prenatal diagnosis of chromosome abnormalities through amniocentesis. In: Milunsky A, ed. *Genetic disorders and the fetus*, 4th ed. New York: Plenum Press, 1998:179–248.

38. Hsu LYF, Perlis TE. United States survey on chromosome mosaicism and pseudomosaicism in prenatal diagnosis. *Prenat Diagn* 1984;4: 97–130.

39. Steinbach P, Djalali M, Hansmann I, et al. The genetic significance of accessory bisatellited marker chromosomes. *Hum Genet* 1983;65: 155–164.

40. Robinson WP, Langlois S, Schuffenhauer S, et al. Cytogenetic and age-dependent risk factors associated with UPD 15. *Prenat Diagn* 1996;16:837–844.

41. Cheng SD, Spinner NB, Zackai EH, et al. Cytogenetic and molecular characterization of inverted duplicated chromosomes 15 from 11 patents. *Am J Hum Genet* 1994;55:753–759.

42. Leana-Cox J, Jenkins L, Palmer CG, et al. Molecular cytogenetic analysis of inv dup(15) chromosomes, using probes specific for the Prader-Willi/Angelman syndrome region: clinical implications. *Am J Hum Genet* 1994;54:748–756.

43. Flejter WL, Bennett-Baker PE, Ghaziuddin M, et al. Cytogenetic and molecular analysis of inv dup(15) chromosomes observed in two patients with autistic disorder and mental retardation. *Am J Med Genet* 1996;61:182–187.

44. Hastings RJ, Nisbet DL, Waters K, et al. Prenatal detection of extra structurally abnormal chromosomes (ESACs): new cases and a review of the literature. *Prenat Diagn* 1999;19:436–445.

45. Li MM, Howard-Peebles PN, Killos LD, et al. Characterization and clinical implications of marker chromosomes identified at prenatal diagnosis. *Prenat Diagn* 2000;20:138–143.

46. Wallerstein R, Yu MT, New RL, et al. Common trisomy mosaicism diagnosed in amniocytes involving chromosomes 13, 18, 20, and 21: karyotype-phenotype correlation. *Prenat Diagn* 2000;20:103-122.

47. Robinson A, Linden MG, Bender BG. Prenatal diagnosis of sex chromosome abnormalities. In: Milunsky A, ed. *Genetic disorders and the fetus*, 4th ed. New York: Plenum Press, 1998:249–285.

48. Hsu LYF. Prenatal diagnosis of 45,X/46,XY mosaicism—a review and update. *Prenat Diagn* 1989;9:31–48.

49. Worton RG, Stern R. A Canadian collaborative study of mosaicism in amniotic fluid cell cultures. *Prenat Diagn* 1984;4:131–144.

50. Vejerslev LO, Mikkelsen M. The European collaborative study on mosaicism in chorionic villus sampling: data from 1986 to 1987. *Prenat Diagn* 1989;9:575–588.

51. Ing PS, Van Dyke DL, Caudill SP, et al. Detection of mosaicism in amniotic fluid cultures: A CYTO2000 collaborative study. *Genet Med* 1999;1:94–97.

52. Moertel CA, Stupca PJ, Dewald GW. Pseudomosaicism, true mosaicism, and maternal cell contamination in amniotic fluid processed with in situ culture and robotic harvesting. *Prenat Diagn* 1992;12: 671–683.

53. Vockley J, Inserra J, Breg WR, et al. "Pseudomosaicism" for 4p- in amniotic fluid cell culture proven to be true mosaicism after birth. *Am J Med Genet* 1991;39:81–83.

54. Van Ravenswallij-Arts CMA, Tuerlings JHAM, Van Heyst AFJ, et al. Misinterpretation of trisomy 18 as a pseudomosaicism at third-trimester amniocentesis of a child with a mosaic 46,XY/47,XY,+3/48,XXY,+18 karyotype. *Prenat Diagn* 1997;17: 375–379.

55. Beverstock GC, Hansson K, Helderman-Van Den Enden ATJM, et al. A near false-negative findings of mosaic trisomy 21—a cautionary tale. *Prenat Diagn* 1998;18:742–746

56. Schwartz S, Leana-Cox J. Fluorescent in situ hybridization (FISH): a new application in the delineation of true vs. pseudomosaicism in prenatal diagnosis. *Prenat Diagn* 1993;13:661–670.

57. Hsu LYF, Yu M-T, Neu RL, et al. Rare trisomy mosaicism diagnosed in amniocytes, involving an autosome other than chromosomes 13, 18, 20, and 21: karyotype/phenotype correlations. *Prenat Diagn* 1997;17:201–242.

58. Phillips OP, Tharapel AT, Lerner JL, et al. Risk of fetal mosaicism when placental mosaicism is diagnosed by chorionic villus sampling. *Am J Obstet Gynecol* 1996;174:850–855.

59. Henderson KG, Shaw TE, Barrett IJ, et al. Distribution of mosaicism in human placentae. *Hum Genet* 1996;97:650–654.

60. Jackson A, Wapner RJ, Davis GH, et al. Mosaicism in chorionic villus sampling: an association with poor perinatal outcome. *Obstet Gynecol* 1990;75:573–577.

61. Kalousek DK, Howard PP, Olson SB, et al. Confirmation of CVS mosaicism in term placentae and high frequency of intrauterine growth retardation in association with confined placental mosaicism. *Prenat Diagn* 1991;11:743–750.

62. Wolstenholme J, Rooney DE, Davison EV. Confined placental mosaicism, IUGR, and adverse pregnancy outcome: a controlled retrospective U.K. collaborative study. *Prenat Diagn* 1994;14:345–361.

63. Artan S, Basaran N, Hassa H, et al. Confined placental mosaicism in term placentae: analysis of 125 cases. *Prenat Diagn* 1995;15:1135–1142.

64. Robinson WP, Barrett IJ, Bernard L, et al. Meiotic origin of trisomy in confined placental mosaicism is correlated with presence of fetal uniparental disomy, high levels of trisomy in trophoblast, and increased risk of fetal intrauterine growth restriction. *Am J Hum Genet* 1997;60: 917–927.

65. Harrison KJ, Barett IJ, Lomax BL, et al. Detection of confined placental mosaicism in trisomy 18 conceptions using interphase cytogenetic analysis. *Hum Genet* 1993;92:353–358.

66. Leschot NJ, Schuring-Blom GH, Van Prooijen-Knegt AC, et al. The outcome of pregnancies with confined placental chromosome mosaicism in cytotrophoblast cells. *Prenat Diagn* 1996;16:705–712.

67. Bui TH, Iselius L, Lindsten J. European collaborative study on prenatal diagnosis: mosaicism, pseudomosaicism and single abnormal cells in amniotic fluid cell cultures. *Prenat Diagn* 1984;4:145–162.

68. Ledbetter DH, Engel E. Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. *Hum Mol Genet* 1995;4:1757–1764.

69. Kotzot D. Abnormal phenotypes in uniparental disomy (UPD): Fundamental aspects and a critical review with bibliography of UPD other than 15. *Am J Med Genet* 1999;82:265–274.

70. EUCHROMIC (European Collaborative Research on Mosaicism in CVS). Trisomy 15 CPM: probable origins, pregnancy outcome and risk of fetal UPD. *Prenat Diagn* 1999;19:29–35.

71. Christian SL, Smith ACM, Macha M, et al., Prenatal diagnosis of uniparental disomy 15 following trisomy 15 mosaicism. *Prenat Diagn* 1996;16:323–332.

32

Cytogenetic Studies in Neoplastic Hematologic Disorders

Gordon W. Dewald

Rhett P. Ketterling

William A. Wyatt

Peggy J. Stupca

Conventional cytogenetic studies help classify malignant hematologic disorders that are associated with specific chromosome abnormalities. The observation of cytogenetic subclones provides clues to disease progression. Cytogenetic studies are useful to assess the effectiveness of treatment and to monitor remission. Moreover, fluorescence *in situ* hybridization (FISH) with chromosome-specific DNA probes enhances testing for patients with malignant hematologic disorders. FISH readily permits analysis of large numbers of proliferating (metaphase) and nonproliferating (interphase nuclei) cells and is useful in establishing the percentage of neoplastic cells before and after therapy. Consequently, cytogenetic studies are widely used by hematologists in the diagnosis and management of their patients. In 1995, the 190 United States' cytogenetics laboratories collectively reported doing 82,855 chromosome studies on patients with neoplastic hematologic disorders (1).

This chapter presents a model for the role of chromosome abnormalities in the origin and progression of malignant hematologic disorders. Information is provided about the types of specimens to collect, procedures for transporting specimens, common laboratory procedures, and the interpretation of results for conventional cytogenetic studies. The application of FISH in the workup and management of patients with hematologic disorders is emphasized. An overview of genetic testing to assess the effectiveness of treatment and discriminate between congenital and acquired neoplastic cytogenetic abnormalities is presented. Finally, the kinds of chromosome abnormalities that occur and their association with current classification schemes for hematologic disorders are presented.

CHROMOSOMAL BASIS OF MALIGNANCY

Part of "32 - Cytogenetic Studies in Neoplastic Hematologic Disorders"

Types of Chromosome Abnormalities

Chromosome abnormalities generally are classified as either numerical or structural (Fig. 32.1). The numerical abnormalities are subclassified into polyploid or aneuploid. The term polyploid refers to chromosome complements that are multiples of 23, which is the haploid number of chromosomes; for example, diploidy refers to 46 chromosomes, triploidy to 69 chromosomes, and tetraploidy to 92 chromosomes. In neoplastic disorders, most polyploid clones associated with advanced stages of disease are derived either from the fusion of neoplastic cells or from endoreduplication.

Aneuploid refers to chromosome complements that involve irregular multiples of the haploid number. For example, a cell that has trisomy 8 is characterized by 47 chromosomes, including three number 8 chromosomes. A cell that is monosomy 7 contains 45 chromosomes and is lacking a chromosome 7. Aneuploid abnormalities usually occur as a consequence of mitotic malfunction, such as chromosome nondisjunction.

Structural abnormalities are classified as translocations, deletions, inversions, duplications, or isochromosomes. Reciprocal translocations involve the interchange of parts of different chromosomes and are the most common type of translocation in hematologic disorders. Deletions involve the loss of part of a chromosome and are either terminal or interstitial. Inversions produce a reversal in the direction of an interstitial part of a chromosome and are either pericentric or paracentric. Pericentric inversions involve both arms while paracentric inversions occur on one arm. Duplications produce two or more copies of a particular DNA segment on the same chromosome. Isochromosomes produce a mirror-image band pattern with respect to the center of the chromosome and arise from a break and fusion of sister chromatids or a translocation between homologous chromosomes.

Origin of Abnormalities

Most structural abnormalities of chromosomes originate during replication or repair of DNA, when the DNA is particularly vulnerable to breakage and fusion. Structural abnormalities of chromosomes occur in interphase when chromosomes are uncoiled and DNA of different chromosomes is overlapped or in close association. DNA breakage and refusion can involve multiple loci and occasionally DNA from several chromosomes. These events may be from recombination involving repetitive or homologous DNA sequences. The subsequent refusion or repair of broken DNA can result in the formation of variant translocations with

or without DNA loss. This observation explains why certain variants of classical structural abnormalities are associated with specific hematologic disorders.

The Ph-chromosome is an example of variation in DNA breakage and fusion in the origin of a specific translocation. Conventional cytogenetic studies indicate several forms of Ph-chromosomes, including classical t(9;22)(q34;q11.2), complex (three or more chromosomes) or masked (submicroscopic insertion). Molecular genetic studies provide evidence of further genetic variation among Ph-chromosomes. In some patients, small deletions and duplications of DNA sequences have been detected within the BCR and ABL gene. Moreover, the DNA break and fusion point with the BCR may occur within the major or minor BCR region and sometimes elsewhere within the BCR region.

FISH studies suggest even further genetic variation among Ph-chromosomes (2). The typical Ph-chromosome involves a break in the BCR region on chromosome 22 and ABL region on chromosome 9. Usually, a reciprocal translocation forms so that a BCR/ABL fusion occurs on both the abnormal chromosomes 9 and 22. Approximately 20% of these patients have an atypical FISH pattern at the fusion site on the abnormal chromosome 9, which indicates loss of portions of the BCR and/or ABL hybridization sites and some adjacent DNA sequences. This loss of DNA sequences most likely occurs during the formation of translocations and does not appear to affect the oncogenic process of chronic myeloid leukemia.

Chromosome abnormalities occur sporadically in cells of every person on a regular basis. In the authors' cytogenetic practice, structural chromosome abnormalities are observed in normal bone marrow specimens at a frequency of 0.6%. These sporadic cells occur in all tissues and represent new mutations rather than clinically significant clones. If more than one metaphase in 30 has nonclonal structural abnormalities, most cytogeneticists become concerned that the patient has a breakage syndrome or was exposed to toxic substances such as chemotherapy or an environmental carcinogen. Several genetic mechanisms to repair DNA damage exist. Most damaged DNA is repaired without producing chromosome abnormalities.

Any mutation affecting the genes that control repair processes can significantly affect the proportion of cells with chromosome abnormalities. This explains the considerable incidence of cells with structurally abnormal chromosomes in patients with chromosome breakage syndromes. The incidence of malignant disorders is higher among people with these genetic disorders than among people with normal DNA repair mechanisms.

Patients with Fanconi's anemia have chromosomes that are particularly subject to breakage by bifunctional alkylating agents such as mitomycin C. Culturing cells of Fanconi anemia patients in the presence of alkylating agents like mitomycin C or

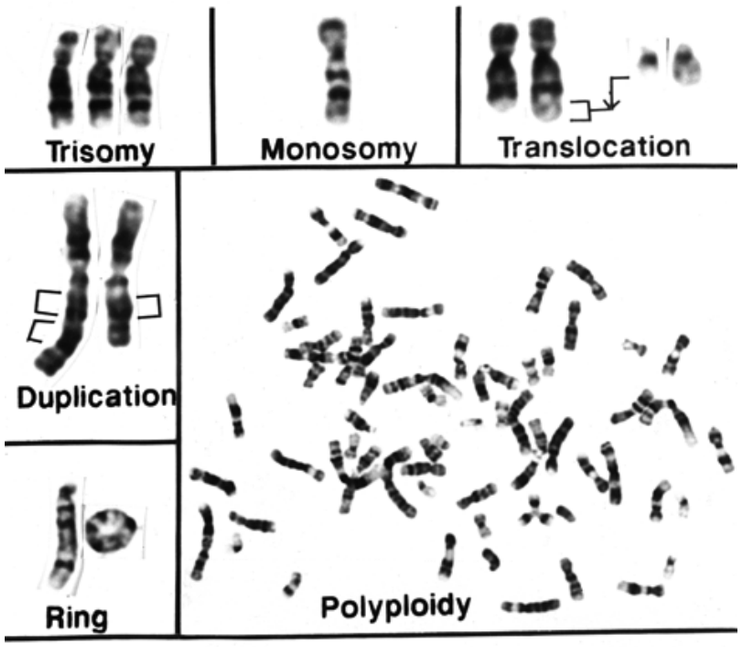


FIGURE 32.1. Examples of common chromosome abnormalities. Trisomy is the presence of three copies of any chromosome. Monosomy denotes the presence of a single copy of any chromosome. Translocations involve the exchange of chromatin between any two or more nonhomologous chromosomes. Duplications produce structural abnormalities with multiple copies of genes. Deletions result in the loss of a portion of a chromosome; a ring chromosome is an example of a deletion. Polyploid cells involve 69 or 92 chromosomes or some exact multiple of the haploid chromosome number.

diepoxybutane, produce many metaphases with radial formations (3). This test is particularly important in the workup of children with aplastic anemia, which may be undiagnosed Fanconi anemia. A chromosome study on peripheral blood must be performed in these patients because they are hypersensitive to alkylating agent chemotherapy, which is used to ablate the bone marrow prior to transplantation. If the patient's chromosomes are hypersensitive, a lower dose of ablative chemicals is needed to avoid causing excess genetic damage in nonneoplastic cells.

Formation of Clones

The clonal pattern of malignancies has been demonstrated with cytogenetic studies, molecular genetic methods, immunocytochemical techniques, enzyme markers, and many other procedures. Cytogenetic studies suggest that nearly all malignant disorders are associated with chromosomally abnormal clones. Numerous cytogenetic reports have been published in recent years recording the presence of chromosomally abnormal clones in a wide variety of hematologic disorders and solid tumors (4).

The formation of each neoplastic clone probably begins with a chromosome abnormality in a single cell, which then proliferates (Fig. 32.2). Observation of identical chromosome abnormalities in different cells of the same tumor is evidence of clonality. The formation of structural abnormalities of chromosomes may be relatively random with respect to the site of breakpoints. Theoretically, a break could affect any site on any chromosome. Considering that the human diploid genome has approximately 6 billion base pairs distributed among 46 chromosomes, about 36×10^{18} different chromosome rearrangements are possible. Certain chromosome abnormalities may produce biologically significant genetic imbalances that affect critical cellular pathways and may result in cell death prior to proliferating. Other chromosome abnormalities may be balanced, which may not affect the cell phenotype or may produce clones with malignant potential. This effect has been observed in patients with ataxia-telangiectasia.

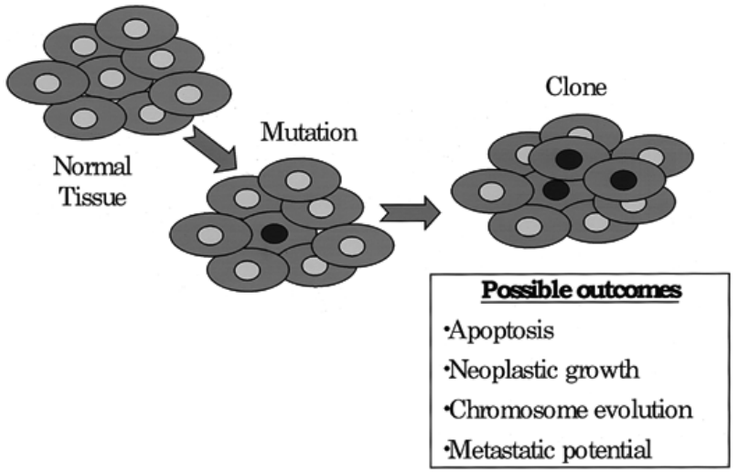


FIGURE 32.2. Formation of a clone. In the early stages, the cells of a chromosomally abnormal clone may die as a consequence of genetic imbalance or may be destroyed by the immune system. If not, the clone may proliferate and form a malignant tumor, and subclones may form as a result of chromosome evolution.

Some chromosome changes are not lethal, but rather enhance the proliferative potential of cells and promote malignancy. These abnormalities may activate oncogenes or disrupt the action of tumor-suppressor genes, which control critical cellular pathways. This explains the correlation of these chromosome changes to specific malignant neoplastic disorders.

Breakage of chromosomes is most likely not random. The likelihood of developing a chromosome abnormality with oncogenic potential is affected by the presence of fragile sites, the association of homologous DNA sequences, carcinogens, and cell survival (Fig. 32.2). In fact, because the number of possible chromosome rearrangements is 36×10^{18}, the probability of developing a common chromosome abnormality by accident seems so remote that chance alone would not explain the relatively high frequency of hematologic disorders.

Chromosome Evolution and Tumor Progression

As neoplastic cells proliferate in most tumors, additional chromosome abnormalities appear in sporadic malignant cells by “chromosome evolution.” These cells divide to produce subclones that contain the initial chromosome abnormality as well as one or more secondary chromosome abnormalities. For example, in chronic myeloid leukemia a t(9;22)(q34;q11.2) is considered the primary chromosome anomaly. The observation of t(9;22)(q34;q11.2) and trisomy 8 in a subset of cells is evidence of a subclone.

Chromosome evolution is responsible for the complex karyotypes observed in many hematologic disorders. The mechanisms of chromosome evolution include nondisjunction, cell fusion, and a wide variety of structural abnormalities. The role of chromosome evolution in tumor progression is not entirely understood. A general correlation does exist between the aggressiveness of hematologic disorders and the appearance of subclones. The number of chromosome abnormalities in an abnormal clone may provide a crude measurement of tumor progression. Indeed, many high-grade lymphomas have very complex karyotypes, while low-grade lymphomas often have only one or two chromosome abnormalities. Chronic myeloid leukemia and the 5q- syndrome are two well-studied disorders where chromosome evolution correlates with more aggressive stages of the disease. Evidence is rapidly accumulating that suggests a similar correlation exists between karyotype complexity and disease progression in other hematologic disorders, especially among the myeloproliferative and myelodysplastic syndromes.

Chromosome evolution may affect tumor response to therapy. Treating a tumor with multiple subclones may be akin to dealing with a patient who has several concurrent neoplastic disorders. For these patients, it may be necessary to develop more than one form of treatment to combat the genotype of each subclone. Evidence suggests that patients with acute myeloid leukemia evolving from a myelodysplastic syndrome with complex karyotypes do not respond well to standard therapies. By comparison, patients with classical chromosome abnormalities associated with overt acute myeloid leukemia such as t(8;21) and t(15;17), experience better response to standard therapies.

Oncogenes

Cell division is initiated and regulated by numerous genes. Under normal circumstances, the genes that govern cell division serve important functions during embryogenesis, growth, and cell repair to ensure the production of new cells. In mature cells in which mitosis is no longer necessary, these genes are inactive. Through mutation, one or more of the genes that control cell division may become reactivated. These genes are inappropriately expressed and, because of lack of regulation, can result in uncontrolled cell proliferation. These aberrantly expressed developmental genes have become known as oncogenes.

The oncogene copy number is increased by trisomy and decreased by monosomy, but the degree of genetic imbalance seldom is more than one gene copy. Recent evidence has emerged to suggest that in some patients, neoplastic cells with trisomy 11 have a submicroscopic duplication of a portion of the MLL oncogene. It is postulated that this duplication is the primary event and trisomy 11 is a result of chromosome evolution. Similar mechanisms may be important for trisomies of other chromosomes.

Structural changes may be more significant than numerical abnormalities in the activation of oncogenes. Evidence indicates that one or more oncogenes are located near sites of chromosome breakage for several of the specific chromosome abnormalities associated with neoplastic disorders (5). These structural changes alter gene expression either by position effect or gene mutation leading to transcription of the oncogene.

Activation of the c-MYC oncogene by the t(8;14)(q24;q32) in Burkitt's lymphoma is an example of a chromosome abnormality, which deregulates cell division. In this translocation, the breakpoints on chromosome 14 are at 14q32, the site of the heavy-chain immunoglobulin gene and on chromosome 8 at 8q24, the site of the c-MYC oncogene. As a consequence of the 8;14 translocation, the c-MYC oncogene is juxtaposed with the heavy-chain immunoglobulin locus on chromosome 14. Transcription of the c-MYC gene then is controlled by the heavy-chain immunoglobulin gene.

Two other translocations associated with Burkitt's lymphoma provide further evidence for the activation of c-MYC in lymphoproliferative disorders. Some patients with Burkitt's lymphoma have a t(8;22)(q24;q11). This translocation brings the lambda immunoglobulin light-chain gene into juxtaposition with the c-MYC oncogene on chromosome 8. Other patients with Burkitt's lymphoma have a t(2;8)(p12;q24), which brings the kappa immunoglobulin light-chain gene into juxtaposition with the c-MYC oncogene. Each of these translocations is known to activate transcription of the c-MYC oncogene.

This mechanism offers an explanation as to why specific chromosome abnormalities are associated with particular neoplastic disorders. B-cells with a t(8;14)(q24;q32) are neoplastic because the heavy-chain immunoglobulin locus induces a hybrid ribonucleic acid (RNA) molecule that is part c-MYC and part heavy-chain immunoglobulin. This translocation has not been observed in other cell types and thus does not appear to have neoplastic potential in cells in which the heavy-chain immunoglobulin locus is not normally transcribed.

Tumor-Suppressor Genes

Tumor-suppressor genes provide another mechanism for tumorigenesis (5). In normal cells, tumor-suppressor genes regulate cellular proliferation. Germline mutations in tumor suppressor genes predispose individuals to developing malignancies. The malignant process is initiated by a mutation or loss of the normal gene, resulting in absence of gene expression. The loss of the normal gene happens by chromosome nondisjunction or by structural rearrangements such as deletions and unbalanced translocations. In a similar fashion, somatic mutations in both copies of a tumor suppressor gene will result in loss of regulation and tumorigenic potential.

Inactivation of tumor-suppressor genes has been implicated in many hematologic disorders. For example, monosomy 7, del(5)(q13q33), del(6)(q14q32), and other losses of chromosomes are candidates for this mechanism although the specific tumor suppressor genes have not be defined.

CONVENTIONAL CYTOGENETIC ANALYSIS

Part of "32 - Cytogenetic Studies in Neoplastic Hematologic Disorders"

Choice of Tissue

Hematologists who use cytogenetic studies need to provide specimens for chromosome analysis that includes malignant cells. Lymph node biopsy specimens are most suitable for lymphomas. Bone marrow specimens are best for patients with myeloproliferative

disorders, myelodysplastic syndromes, chronic lymphocytic leukemias, and acute leukemias. Chromosome studies often can be accomplished on bone marrow biopsy specimens even if the bone marrow aspirate is paucicellular. If the disease has circulating cells, a peripheral blood specimen can be collected, but for myeloid disorders, this should be done only if bone marrow is not available. Only about 60% of blood specimens produce adequate metaphases as compared to more than 95% of bone marrow specimens. Chromosome analysis of peripheral blood is informative mainly in advanced disorders, chronic lymphocytic leukemia, and some cutaneous T-cell lymphomas.

In certain clinical situations, cytogenetics is required on various body fluids for diagnosis or to rule out metastasis or relapse. These fluids include pleural or ascitic effusions, spinal fluid, and rarely other specimens. In pleural fluids, lymphomas are often more readily diagnosed by cytogenetic techniques than by standard cytologic examination (6). Cytogenetics can be successful even on low volume specimens that appear to lack cells.

Most therapeutic agents interfere with attempts to produce metaphases from specimens because these agents are intended to prevent cell division. Successful chromosome studies usually are possible 2 to 3 weeks after treatment. Although the agents used to treat malignancies may induce sporadic abnormalities in chromosomes, they should not interfere with the identification of chromosomally abnormal clones.

Specimen Collection and Transportation

Routine chromosome studies are performed on mitotically active cells. Bone marrow specimens should be <1 mL to reduce the chance of clotting. Successful chromosome studies routinely are accomplished on 0.25 to 0.50 mL of bone marrow. The specimen should be collected with sodium heparin using sterile techniques. If multiple tests are requested for a bone marrow specimen, cytogenetic studies should receive an early tap to improve the likelihood of obtaining dividing cells. This can be done without interfering with the collection of fresh specimens for standard hematologic examination by redirecting the needle during bone marrow collection.

When transporting bone marrow specimens to a cytogenetic laboratory, particular attention should be paid to transportation methods. Bone marrow specimens up to 1.0 mL can be transported in culture medium or in sodium heparin vacutainers. When fresh bone marrow is transported, it is important to send the sample as soon as possible. Alternatively, preprocessing bone marrow specimens with colcemid solution, hypotonic solution, and fresh methanol-glacial acetic acid may be performed (7). Because the cells are transported in fixative, transportation time and environmental conditions do not affect the laboratories' success in obtaining analyzable metaphases. If cells can be processed locally, both a direct technique and short-term culture methods can be used in concert. This method produces successful chromosome studies approximately 96% of the time.

Chromosome analysis of peripheral blood requires 7.0 to 10.0 mL collected in a sterile syringe containing sodium heparin. The specimen should be transferred immediately to a sodium heparin vacutainer to prevent clotting. Sodium heparin is highly recommended as the anticoagulant of choice because other anticoagulants may interfere with cell viability. A cytogenetic laboratory should be consulted if sodium heparin is unavailable to determine a viable second choice.

Chromosome studies can be done on lymph-node biopsy specimens of 1.0 to 3.0 cm^2. The specimen should be transported in a screw-capped, sterile container with sufficient sterile tissue medium with serum, sterile Hank's balanced salt solution, or a sterile isosaline solution to prevent tissue dehydration. The specimen can be refrigerated or sent to the cytogenetic laboratory at ambient temperature.

Processing Specimens for Chromosome Analysis

Various methods exist to process specimens for chromosome analysis. Specimens can be processed immediately after collection using a direct method. Cells are treated with Colcemid for 1 hour, warm hypotonic solution (0.075 M KCl) for 15 minutes and fixed with 3:1 methanol-glacial acetic acid. This method has a success rate of more than 92% and allows a diagnosis within 24 hours.

Short-term cultures of 24 to 48 hours are used by many laboratories. These cultures use a standard medium such as RPMI 1640 with fetal bovine serum, and antibiotics. Cells are incubated in one to three 25-mL T-flasks at 37°C. It is desirable to use a low-oxygen environment (5% oxygen, 5% carbon dioxide, and 90% nitrogen) with these cultures. Cultures can be incubated overnight (14 to 16 hours) with Colcemid (Roche Pharmaceuticals) and FUdR to improve metaphase quality. Cells are harvested for chromosome analysis using the standard procedures with Colcemid, hypotonic solution, and fixative. These methods are successful for more than 90% of bone marrow specimens.

The *in situ* method involves culturing the cells on six to eight chamber slides or coverslips. The cells are cultured in low oxygen using a culture medium such as RPMI 1640 and harvested. This method avoids loss of cells associated with the T-flask technique, is amenable to robotic harvesting (Tecan, Inc.), and allows cultures to be processed easily with different methods (e.g., with or without mitogens) (8). Because few cells are lost during the harvest process, this method is particularly successful on specimens with a low mitotic index.

Most laboratories process hematologic specimens using both a direct preparation and a short-term culture. A comparison of the direct technique versus 24-hour unstimulated cultures in a blind study of 100 unselected bone marrow specimens revealed no significant differences with respect to chromosome morphology, mitotic index, or the detection of an abnormal clone (9). Nevertheless, because one method occasionally fails, it is useful to routinely use at least two methods with each specimen.

Mitogens are used to stimulate division of neoplastic cells in chromosome studies of patients with lymphoproliferative disorders. Mitogens such as phytohemagglutinin, 12-*O*-tetradecanoylphorbol-13-acetate (TPA), Interleukin-2, pokeweed, and inactivated Epstein-Barr virus have been used successfully. Unfortunately, mitogens also may stimulate normal B-cells and T-cells to divide which can interfere with the analysis. Furthermore, not all cells of lymphoproliferative disorders respond to

mitogens. Some laboratories have success by adding hematopoietic growth factors such as PHA-stimulated leukocyte-conditioned medium.

Several techniques have been used to optimize the band resolution of chromosomes. Longer chromosomes are produced by incorporating DNA intercalaters (e.g., actinomycin D or ethidium bromide), methotrexate, or low concentrations of Colcemid into the culture method.

Staining

The most common chromosome staining methods are G-banding and Q-banding. G-bands are easily induced by a brief pretreatment with an enzyme such as trypsin, followed by staining with Leishman's, Giemsa, or Wright's. G-bands produce permanent preparations that easily permit the identification of each chromosome and the detection of subtle structural abnormalities.

Q-banding involves staining the chromosomes with quinacrine mustard and using a fluorescence microscope system. Because this stain usually fades in a few minutes, n-propyl gallate is used to slow the rate of fluorescence loss. This method permits the identification of all chromosomes and detection of subtle structural abnormalities. Q-banding is the staining technique of choice when there are only a few metaphases or when chromosome morphology or metaphase spreading is poor.

Slide Preparation

Several techniques are used to optimize metaphase spreading. These techniques include the use of cold fixative, dropping cell suspensions from increased heights, flame drying the slides, and various methods to modify the rate at which slides dry. These techniques attempt to influence the characteristics of chromosome spreading by modifying the rate at which metaphases dry and chromosomes spread on the coverslip or microscope slide (10). An environmental chamber by Thermatron can be used to control temperature and humidity, which helps to optimize chromosome spreading.

Analysis of Metaphases

At least 20 metaphases are analyzed in order to statistically exclude a clone that accounts for more than 14% of cells with 95% confidence. The analysis of 30 cells excludes a clone that accounts for more than 10% of cells with 95% confidence. In either case, this assumes that metaphases are selected randomly for analysis. Because not all metaphases provide equal information, many technologists do not randomly select metaphases, and usually examine more metaphases than they actually report. During analysis, the number and band pattern of each chromosome is established. Interfering factors for effective analysis include overlapping chromosomes, inconsistent staining, and variable morphology.

Metaphase and Karyotype Documentation

Several metaphases should be documented by photographs or computer images. Many laboratories use computer-assisted karyotype systems to eliminate darkroom procedures, capture FISH images, and integrate data into the laboratory computer system. For each specimen, formal karyotypes are prepared from representative metaphases to objectively compare band patterns between chromosomes. Most laboratory-certifying agencies require preparation of at least two karyotypes from the primary clone and one karyotype from each subclone.

Definition of a Clone

The standard cytogenetic definition of a clone requires at least two metaphases with the same extra chromosome or the same structural abnormality, or the presence of three or more metaphases missing the same chromosome. The stricter definition for chromosome loss helps exclude technical artifacts.

The definition of a clone should include a maximal number of metaphases to be analyzed because nonclonal chromosome abnormalities are relatively common. We prefer to analyze no more than 30 metaphases in each study to avoid false-positive findings and to use laboratory time most efficiently. Nevertheless, the observation of even one cell with a classic chromosome abnormality could be significant. If this abnormality is reported, it should be stated that this discovery does not meet the minimal criteria for a clone.

FLUORESCENCE IN SITU HYBRIDIZATION (FISH)

Part of "32 - Cytogenetic Studies in Neoplastic Hematologic Disorders"

The advent of fluorescent-labeled DNA probes with *in situ* hybridization greatly enhanced the usefulness of cytogenetics to study hematologic disorders. Many fluorescent-labeled DNA probes can be purchased from commercial companies. Other probes are readily made as “home-brew” products using cosmids, plasmid artificial chromosomes, yeast artificial chromosomes, and bacterial artificial chromosomes. Numerous FISH strategies are available to detect abnormalities associated with neoplastic processes. Moreover, FISH can be performed on both proliferating (metaphase) and nonproliferating (interphase) cells. FISH bridges the testing gap between cytogenetic and molecular genetic studies, is less expensive than conventional cytogenetics, can be performed overnight (sometimes in less than 6 hours), and can produce quantitative results with considerable precision.

FISH is used to accurately define the disease-state of many hematologic disorders at diagnosis, during therapy, and remission. It is beyond the scope of this chapter to discuss the entire field of FISH in hematologic disorders. Enough information about FISH is provided to use this technology in clinical practice.

FISH in Clinical Practice

The utility and accuracy of FISH to detect chromosome abnormalities has been established in hematologic disorders. However, the Food and Drug Administration (FDA) has approved only a few commercial FISH probes. Recent data from the College of American Pathologists and American College of Medical Genetics indicate that FISH is widely used by cytogenetic laboratories (11). Nomenclature for both metaphase and interphase FISH

has been created and adopted by cytogeneticists (12). Accrediting agencies have published criteria for FISH in routine practice and at least one accrediting agency provides an on-going proficiency testing program for FISH(11, 13).

Validation of FISH Tests

Validation of new tests is necessary to ensure safe and effective testing for the specific intended clinical purpose. Validation of FISH tests is a multistep process designed to assess probe performance in normal and abnormal specimens, and to establish the ability of laboratory personnel to perform the test consistently.

The procedures to validate FISH tests vary significantly depending upon the specific clinical application and FISH methodology (i.e., whole chromosome-specific paints, centromere-specific probes, locus-specific or telomere-specific probes). Moreover, the extent of validation among laboratories varies depending on probe approval by the FDA and the extent of published work regarding the test. Procedures to validate FISH tests have been reported (13, 14), but because this is a new field, it is anticipated that this subject will be dealt with in more detail by accrediting agencies in the future.

Validation and scoring criteria should consider several factors. Certain biological and technical factors introduce false signal patterns and affect the accuracy of FISH to detect and quantify neoplastic cells. Split signals can result from G_2 separation and asynchronous replication of DNA sequences. In addition, cells in normal individuals occasionally demonstrate aneuploidy. Probe signals sometimes appear to split because of incomplete saturation of the hybridization site, coincidental overlap of signals, and occasional cross-hybridization.

Advantages of FISH

FISH tests can be performed on uncultured interphase nuclei using overnight procedures. Thus, for hematologic disorders where a quick diagnosis is important, FISH is particularly useful. Acute promyelocytic leukemia (AML-M3) is an example of a hematologic disorder where a rapid diagnosis is possible by FISH and these patients can die quickly if left without proper treatment.

Sometimes hematologic specimens are not available for conventional cytogenetic studies. In these cases, FISH studies of interphase nuclei can be performed on a variety of specimens, including fixed cells from peripheral blood or bone marrow, smears made from peripheral blood or bone marrow, touch preparations, and paraffin embedded specimens.

FISH is useful to detect abnormalities when results of conventional cytogenetic studies are inconclusive because of the absence of metaphases, or normal because of cryptic or masked translocations. For example, 5% of patients with chronic myeloid leukemia have a Ph-chromosome from a submicroscopic insertion translocation, which results in BCR/ABL fusion. This translocation is undetectable (masked) in conventional cytogenetic studies, but is readily detected by FISH studies. Similarly, the 12;21 translocation associated with 20% of children with B-cell acute lymphocytic leukemia is undetectable by conventional cytogenetic studies. A FISH test is available to detect the TEL/AML1 fusion, which results from this 12;21 translocation. Thus, when conventional cytogenetics are normal and B-cell acute lymphocytic leukemia is suspected, it is useful to perform FISH studies.

Types of FISH Probes

FISH probes can be classified into four categories: (i) centromere-specific, (ii) telomere-specific, (iii) locus-specific, and (iv) chromosome-specific paints. Probes in these categories can be used alone or in combination and can be labeled with different colored fluorochromes.

Centromere-specific probes are used to detect numeric abnormalities. These probes hybridize to highly repetitive alpha satellite DNA associated with the centromere and are available from commercial sources for most human chromosomes. The gain or loss of specific chromosomes is visualized as an increase or decrease in the number of centromeric FISH signals within a nucleus.

Telomere-specific probes are commercially available and are useful to detect subtle chromosome abnormalities that involve the ends of chromosomes.

Locus-specific probes are used to detect structural abnormalities. The gain of locus-specific probe signals within a nucleus corresponds to duplications or amplification. The loss of locus-specific probe signals indicates a deletion. Locus-specific probes also can be used to detect reciprocal translocations and inversions by using probes that span the breakpoint cluster regions.

Chromosome-specific paints are composed of unique sequence DNA probes that hybridize to loci along the length of each chromosome. The application of these probes is restricted to analysis of metaphases where the morphology of individual chromosomes is discernable. This method is useful to identify marker chromosomes and to detect cryptic translocations. Recent modifications of chromosome-specific paints permit the simultaneous identification of each human chromosome by their color in a single metaphase (Fig. 32.3A). Several different methods have emerged that can be collectively referred to as multicolored FISH.

Probe Strategies and Analytical Sensitivities

Probes used individually or in combination with different colored fluorochromes allow numerous FISH testing strategies. Some FISH strategies employ a probe of one color to detect a genomic target of clinical significance and a different colored probe for a control locus. The addition of a control probe assures that the FISH procedure is working correctly, improves the accuracy of the scoring criteria, and verifies that reportable ranges are maintained. A normal interphase signal pattern is shown in Figure 32.3B.

Strategies that use both a centromere-specific probe and locus-specific probe for the same chromosome permit simultaneous detection of several kinds of chromosome abnormalities with different testing sensitivities. In our experience analyzing 200 cells, the upper boundary of the normal range for detecting a monosomy signal pattern (Fig. 32.3C) is 3% to 5%; for a trisomy signal pattern (Fig. 32.3D), 1% to 3%; for duplications (Fig. 32.3E), 8% to 10%; for deletions (Fig. 32.3F), 8% to 10%; and for gene amplification (Fig. 32.3G), less than 1%.

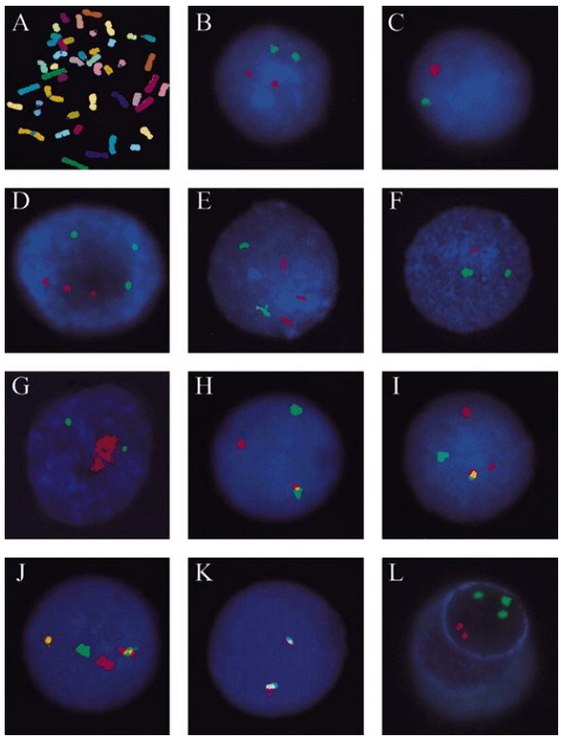


FIGURE 32.3. Example FISH strategies for detection of chromosome abnormalities in hematologic disorders. **(a)** Multicolored FISH chromosome-specific painting system for metaphase analysis. **(B–G)** are examples of interphase FISH patterns using red target and green control loci on the same chromosome: **(B)** normal signal pattern, **(C)** monosomy, **(d)** trisomy, **(E)** duplication, **(F)** deletion, and **(G)** amplification. **(H–K)** are examples of two color interphase FISH fusion strategies: **(H)** S-FISH, **(I)** ES-FISH, **(J)** D-FISH, and **(K)** break-apart. **(L)** Combination immunostain/FISH strategy, using plasma cell cytoplasmic immunoglobulin stain and FISH centromere probes.

FISH strategies to detect translocations and inversions in interphase nuclei have been developed. These strategies have different sensitivities as demonstrated by those employed for chronic myeloid leukemia. The t(9;22)(q34;q11.2) is associated with chronic myeloid leukemia and involves break and fusion points within the BCR region of chromosome 22q11.2 and the ABL oncogene of chromosome 9q34.

One FISH strategy uses a probe for the centromeric side of the BCR region and a different colored probe for the distal side of the ABL oncogene (Fig. 32.3H). As a consequence of the 9;22 translocation, the ABL signal is fused with the BCR signal on chromosome 22, creating a single-fusion signal. This single-fusion strategy is called S-FISH (15). The coincidental overlap of BCR and ABL signals produces false-positive fusion signals in 3% to 5% of normal nuclei, which results in an upper boundary of the normal range of 10%.

Another FISH strategy to detect t(9;22)(q34;q11.2) uses the same BCR probe, but a larger ABL probe to span the entire breakpoint region of ABL. With this method, abnormal nuclei show a BCR/ABL fusion signal on chromosome 22 and a small ABL signal on the abnormal chromosome 9 (Fig. 32.3I). This extra ABL signal strategy is called ES-FISH. This method reduces the problem of scoring cells with false-positive BCR/ABL fusion signals. Nevertheless, false-positive signal patterns result from splitting of ABL signals with coincidental overlap of BCR signals. In our experience, the upper boundary of normal range for this method is 3%.

The most sensitive translocation strategy for t(9;22) (q34;q11.2) involves probes that span both the BCR and ABL breakpoint regions. With this method, abnormal nuclei show two BCR/ABL fusion signals; one on the abnormal chromosome 9 and the other on the abnormal chromosome 22 (Fig. 32.3J). This double-fusion strategy is called D-FISH (16). Nearly all technical and biological artifacts are avoided with D-FISH and this results in an upper boundary of the normal range of <1%.

A break-apart FISH strategy has been developed to detect specific structural rearrangements. This strategy has been used to detect breakpoints within the CBFß gene on chromosome 16q22, the MLL gene on chromosome 11q23, and the ALK gene on chromosome 2p23. This FISH strategy uses different colored probes that hybridize to sequences flanking both sides of the common breakpoint region. In normal nuclei, this signal pattern would appear as two fusion signals (Fig. 32.3K). Any structural abnormality that alters the breakpoint region separates these signals changing both the number and color of the signals in the nucleus (Fig. 32.3H). In our experience, the upper boundary of the normal range for this method is 1% to 3%.

FISH combined with immunostaining methods can be used to establish which hematopoietic cell line is neoplastic. For example, in multiple myeloma, an immunostain for cytoplasmic immunoglobulin to detect plasma cells can be combined with a FISH strategy to detect certain chromosome abnormalities (Fig. 32.3L)(17).

Several FISH strategies can be applied to a single slide to permit “panel FISH testing.” For example, using a set of probes for chromosomes 5, 7, 8, 11, 13 and 20 will allow detection of the most common chromosome abnormalities in myelodysplasia. Probes to detect t(11;14)(q13;q32), +12, del(6)(q13q23), del(11)(q23), del(13)(q12q22), and del(17)(p11) would be a useful FISH panel to study patients with lymphoproliferative disorders.

GENETIC TESTING IN CLINICAL PRACTICE

Part of "32 - Cytogenetic Studies in Neoplastic Hematologic Disorders"

Assessing Effectiveness of Treatment

Numerous treatment protocols for hematologic disorders require quantitative procedures to assess responsiveness to therapy. This can be done by using various genetic tests including quantitative cytogenetic studies, FISH, Southern blot analysis, and reverse transcriptase polymerase chain reaction (RT-PCR). No single genetic testing procedure fulfills all the needs of clinical care for patients with hematologic disorders. Thus, it has become important to use combinations of testing methods that are both accurate and cost-effective for each clinical situation. Testing strategies change over time as methodologies improve or change and new clinical challenges arise. Thus, the review process to develop better genetic testing methods should be an ongoing activity.

The application of genetic testing strategies for chronic myeloid leukemia may be a good model for most hematologic disorders. At diagnosis, it is useful to perform quantitative cytogenetic analysis of 25 or more metaphases from bone marrow. To monitor patients during therapy, quantitative cytogenetic studies on 25 or more metaphases from bone marrow should be done at 3- to 6-month intervals. In lieu of this approach, FISH can be used to study 500 interphase nuclei from blood or bone marrow before and after treatment to track changes in the percentage of cells with BCR/ABL fusion (16). When the results of FISH on interphase nuclei are within the normal range (usually 1% if D-FISH is used), either RT-PCR analysis can be performed or D-FISH analysis of 6,000 or more nuclei could be studied to look for residual disease (2, 16). Remission and relapse can be defined by the test used to detect the abnormal clone. For example, a patient can be in morphological, cytogenetic, FISH, or RT-PCR remission or relapse.

Bone Marrow Transplantation Programs

Chromosome studies are useful before transplantation to establish the karyotype of the neoplastic process. This information helps identify patients who benefit from this procedure and determine the best form of bone marrow transplantation (autologous versus allogenic, purging B-cells, etc.). Genetic testing also is used to establish remission and to determine the best time for bone marrow transplantation.

FISH studies are useful to assess the success of bone marrow transplantation. For opposite sex bone marrow transplantation, DNA probes for the X and Y chromosomes can establish the presence or absence of hematopoietic chimerism. This FISH test can detect donor or recipient cells at a frequency of <1%. FISH with DNA probes specific for a neoplasm is useful to quantify the tumor burden, but the sensitivity of this method varies according to the FISH strategy employed. In some cases, it is useful to use molecular genetic techniques such as RT-PCR to assess the presence of neoplastic cells.

Chromosome studies can determine when relapse is related to the original clone or if a new neoplasm or subclone has emerged. Chromosomal polymorphisms help determine whether the new leukemia involves the donor's cells or the recipient's cells.

Hematopoietic chimerism is common in the posttransplantation period. Molecular genetic methods often detect residual recipient cells as long as 3 years after bone marrow transplantation. Some investigators believe chimerism plays an important role in graft versus host disease.

Constitutional Chromosome Abnormalities

Congenital chromosome abnormalities are part of the constitutional karyotype and, except in cases of mosaicism, occur in all tissues of the patient. Approximately one in 500 adults carries a balanced structural chromosome abnormality and usually is phenotypically normal. However, they have a high risk of having children with birth defects and an elevated rate of miscarriage. The detection of congenital structural chromosome abnormalities is expected to occur during the course of cytogenetic studies for neoplastic disorders. These findings have significant clinical implications for the patients and their blood relatives and they should receive genetic consultation.

In some bone marrow analyses, no clinical or cytogenetic clue is available to clearly distinguish between a constitutional or acquired chromosome abnormality. In these cases, chromosome analysis on peripheral blood cultured in medium containing phytohemagglutinin to stimulate T-cell mitosis is helpful to compare the karyotype of normal T-cells with neoplastic cells.

Certain aneuploid abnormalities are congenital, such as Down, Turner, and Klinefelter syndromes. In these cases, the diagnosis usually is evident by physical examination. Some of these congenital syndromes are associated with a predisposition to develop hematologic disorders. However, these same aneuploid abnormalities are not uncommon in hematologic disorders (18). Fortunately, in hematologic disorders, these aneuploid conditions usually are associated with additional chromosome abnormalities. In these cases, the chromosomally abnormal cells are unbalanced and are associated with neoplastic disease.

CHROMOSOME ABNORMALITIES AND THEIR FREQUENCY OF OCCURRENCE

Part of "32 - Cytogenetic Studies in Neoplastic Hematologic Disorders"

Frequency of Abnormal Clones

All patients with malignancies have one or more clones of neoplastic cells that contain abnormal genes, many of which are detectable by cytogenetic studies. In clinical practice, the frequency with which chromosomally abnormal clones are found in patients with specific hematologic disorders varies considerably. In the authors' cytogenetic and clinical practice, abnormal clones are found in approximately 70% of patients with acute myeloid leukemia, 80% with acute lymphocytic leukemia, 100% with chronic myeloid leukemia, 55% with myelodysplastic syndromes, 39% with polycythemia vera, and 79% with non-Hodgkin's lymphoma. These observations are based on cytogenetic data collected at the time of initial diagnosis. Higher frequencies of chromosomally abnormal clones are observed among patients in more advanced stages of hematologic disorders and lower frequencies are found among patients that undergo successful treatment for their hematologic disorder.

Absence of Abnormal Clones

Several factors contribute to the variations among laboratories with respect to the frequency of detecting chromosomally abnormal clones. The referral patterns of physicians who use cytogenetic studies differ with regard to the type of disease, treatment, and stage of illness. Many cytogenetic studies are done as part of the initial examination of a patient to help establish a diagnosis. This analysis often occurs relatively early in the disease process, when normal cells predominate. The typical cytogenetic study is done on a single bone marrow specimen. Frequently, the most cellular part of the specimen is used for other laboratory tests, such as standard hematologic studies. In such cases, the cytogenetic study is relegated to the “bloody” part of the specimen, often lacking mitotic neoplastic cells. In some hematologic disorders, the neoplastic cells rarely divide, which interferes with the chances of finding a chromosomally abnormal clone. In other disorders, the oncogenic process involves cryptic chromosome abnormalities or submicroscopic mutations which are not detected by standard cytogenetic studies.

In routine practice, usually no more than 4 or 5 hours can be allotted to the analysis of metaphases for each case and investing excessive hours would significantly affect the cost. This work limitation reduces the total number of metaphases that can be analyzed practically in routine practice and increases the chances of missing neoplastic metaphases by sampling error. Moreover, technical experience of the laboratory staff and the processing procedure affects the chances of finding chromosomally abnormal clones. Cytogenetic results often are needed within a few days so that the physician can use the results to manage the patient.

Frequency of Specific Chromosome Abnormalities

Prior to 1985, 43 specific chromosome abnormalities were associated with hematologic disorders. In 1985, the frequency for each of these abnormalities was calculated from 748 predominantly adult patients with an abnormal clone seen at one institution (19). Collectively, these 748 patients had 1,352 chromosome abnormalities. The 43 recognized chromosome abnormalities were seen in 61% of patients. The remaining 39% of patients had other chromosome abnormalities that were not yet associated with any specific hematologic disorder.

Translocations were the most common abnormalities, followed by trisomies, deletions, and monosomies (Table 32.1 and Table 32.2, and Fig. 32.4). The most common specific chromosome abnormality was the Ph-chromosome, followed by +8, -Y, 5q-, -7, 20q-, and +21. The most common trisomy was trisomy 8, followed by trisomy of chromosomes 21, 19, 9, 11, and 7. Loss of the Y chromosome was the most common monosomy, followed by monosomy of chromosomes 7, 5, and X. The most common deletion was 5q-, followed by 20q-, 7q-, 6q-, 13q-, 11q-, 9q-, and 16q-. The most common translocation

was t(9;22), followed by t(8;21), t(1;7), t(8;14), t(9;11), t(4;11), and t(14;18). Isochromosomes, inversions, duplications, haploidy, and polyploidy were infrequent.

TABLE 32.1. COMMON STRUCTURAL CHROMOSOME ABERRATIONS IN HEMATOLOGIC DISORDERS

				Frequency in 748 pt[c]		
Abnormality[a]	Example	Associated Genes	Associated Disorders[b]	Total pt	Pt with one abn	Pt with multiple abn
Translocations						
t(1;2)(q21;p23)		?, ALK	ALCL	0	0	0
t(1;3)(p36;q21)			AML-M1, AML-M4, MDS	0	0	0
t(1;6)(q23-25;p21-25)	Fig. 32.5w		MPD	0	0	0
der(1;7)(q10;p10)	Fig. 32.5a		AML-M1, AML-M2, AML-M4, MDS, t-MDS	13	7	6
t(1;11)(p32;q23)	Fig. 32.5b	TAL1, MLL	ALL, AML	0	0	0
t(1;11)(q21;q23)		AFIρ, MLL	AML-M4, AML-M5	0	0	0
t(1;14)(p32;q11)	Fig. 32.5x	TAL1, TCRαδ	T-ALL	0	0	0
t(1;17)(p36;q21)		?, RARα	AML-M3	0	0	0
t(1;19)(q23;p13)	Fig. 32.5y	PBX1, E2A	Pre-B-ALL	1	1	0
t(1;22)(p13;q13)			AML-M7	0	0	0
t(2;3)(p12;q27)		Igκ, BCL6	DLCL, FL	0	0	0
t(2;3)(p23;q21)		ALK, ?	ALCL	0	0	0
t(2;5)(p23;q35)	Fig. 32.5aa	ALK, NPM	ALCL	0	0	0
t(2;8)(p12;q24)	Fig. 32.5bb	Igκ, c-MYC	ALL-L3, BL, NHL	0	0	0
t(2;18)(p12;q21)			NHL	0	0	0
t(2;11)(p21;q23)	Fig. 32.5c		MDS	0	0	0
t(2;14)(p13;q32)	Fig. 32.5d		B-CLL	0	0	0
t(3;3)(q21;q26)	Fig. 32.5cc	Ribophorin I, EVII	AML, MDS	0	0	0
ins(3;3)(q26;q21q26)	Fig. 32.5e		AML, MDS	2	1	1
t(3;5)(q21;q35)	Fig. 32.5f		AML-M6	0	0	0
t(3;5)(q25;q35)		MLF1, NPM	AML-M6	0	0	0
t(3;14)(p21;q32)	Fig. 32.5dd	?, IgH	NHL	0	0	0
t(3;14)(q27;q32)		BCL6, IgH	DLCL, FL	0	0	0
t(3;21)(q26;q22)	Fig. 32.5g	EAP, AML1	AML, CML Ph+, MDS	0	0	0
t(3;22)(q27;q11)		BCL6, Igλ	DLCL, FL	0	0	0
t(4;11)(q21;q23)	Fig. 32.5h	AF4, MLL	ALL, AML	5	3	2
t(4;14)(p16.3;q32)		FGFR3, IgH	MM	0	0	0
t(5;12)(q33;p13)		PDGFRB, TEL	CMML, MDS, MPD	0	0	0
t(5;14)(q31;q32)		IL3, IgH	ALL	0	0	0
t(5;17)(q35;q21)		NPM, RARα	AML-M3	0	0	0
t(6;9)(p23;q34)	Fig. 32.5i	DEK, CAN	AML-M1, AML-M2, AML-M4, MDS	1	1	0
t(6;11)(q27;q23)	Fig. 32.5j	AF6, MLL	AML-M4, AML-M5	0	0	0
t(6;12)(q15;p13)			CLD	0	0	0
t(6;14)(p25.3;q32)		IRF4, IgH	MM	0	0	0

t(7;11)(p15;p15)	Fig. 32.5ee	HOXA9, NUP98	AML, AML-M2	0	0	0
t(8;13)(p11;q12)		FGFR1, ZNF198	MPD	0	0	0
t(8;14)(q11;q32)			ALL	0	0	0
t(8;14)(q24;q32)	Fig. 32.5l	c-MYC, IgH	ALL-L3, BL, MM, NHL	8	4	4
t(8;14)(q24;q11)	Fig. 32.5k	c-MYC, TCRαδ	T-ALL	0	0	0
t(8;16)(p11;p13)	Fig. 32.5ff		AML-M4, AML-M5	0	0	0
t(8;21)(q22;q22)	Fig. 32.5m	ETO, AML1	AML-M2, AML-M4, MDS	15	6	9
t(8;22)(q24;q11)	Fig. 32.5gg	c-MYC, Igλ	ALL-L3, BL	0	0	0
t(9;11)(p22;q23)	Fig. 32.5n	AF9, MLL	ALL, AML-M5, MDS, t-AML	6	3	3
dic(9;12)(p13;p11)			ALL	0	0	0
t(9;14)(p13;q32)		PAX5, IgH	B-NHL, LPL	0	0	0
t(9;22)(q34;q11.2)	Fig. 32.5o	ABL, BCR	ALL, AML-M1, AML-M2, CML, MPD	378	292	86
t(10;11)(p13;q23)	Fig. 32.5p	AFI0, MLL	AML-M4, AML-M5	0	0	0
t(10;14)(q24;q11)		HOX11, TCRαδ	T-ALL	0	0	0
t(11;14)(p13;q11)	Fig. 32.5q	Rhom2, TCRαδ	T-ALL, NHL	3	1	2
t(11;14)(q13;q32)	Fig. 32.5r	CCND1, IgH	B-PLL, CLD, MM, MCL, MGUS, NHL	5	0	5
t(11;17)(q23;q21)		PLZF, RARα	AML-M3	0	0	0
t(11;17)(q23;q21)		MLL, AFI7	AML-M4, AML-M5	0	0	0
t(11;18)(q21;q21)		API2, MLT	MZL, NHL	0	0	0
t(11;19)(q23;p13.1)		MLL, ELL	ALL, AML-M4, AML-M5, t-AML			
t(11;19)(q23;p13.3)	Fig. 32.5hh	MLL, ENL	ALL, AML-M4, AML-M5, t-AML	0	0	0
t(12;21)(p13;q22)		TEL, AML1	ALL	0	0	0
t(14;16)(q32;q23)		IgH, c-MAF	MM	0	0	0
t(14;18)(q32;q21)	Fig. 32.5s	IgH, BCL2	CLD, FL, DLCL, MM, NHL	5	0	5
t(14;19)(q32;q13)	Fig. 32.5t	IgH, BCL3	CLD, CLL, NHL	0	0	0
t(14;22)(q32;q11)	Fig. 32.5u	IgH, BCR	ALL	0	0	0
t(15;17)(q22;q21)	Fig. 31.5v	PML, RARα	AML-M3, CML Ph+	1	1	0
t(16;16)(p13;q22)		MYH11, CBFß	AML-M4Eo, MDS	0	0	0
t(16;21)(q24;q22)		MTG16, AML1	AML-M2, AML-M4, Childhood AML	0	0	0
t(16;21)(q24;q22)		FUS, ERG	AML-M1, AML-M2	0	0	0
t(17;19)(q21-22;p13)		HLF, E2A	ALL	0	0	0
t(X;11)(q13;q13)		AFIP, MLL	T-ALL	0	0	0
der(Y)(Y;1)(q12;q21)			MDS	0	0	0
Isochromosomes						
i(1)(q10)	Fig.31.6b		CLD, NHL	0	0	0
i(3)(q10)			NHL	0	0	0
i(6)(p10)	Fig. 31.6g		ALL, CLD, NHL	0	0	0
i(7)(q10)	Fig. 31.6ii		ALL, AML, HSγδTL	0	0	0
i(8)(q10)	Fig. 31.6m		T-CLL, T-PLL	0	0	0
i(9)(q10)	Fig. 31.6jj		ALL, NHL	0	0	0
i(10)(q10)			ALL, AML, CLD	0	0	0
i(11)(q10)			AML	0	0	0
i(17)(q10)	Fig. 31.6w		AML, CML Ph+, MDS, MPD	33	1	32
i(21)(q10)	Fig. 31.6z		AML, ALL, MDS, NHL	0	0	0

idic(X)(q13)	Fig. 31.6s		MDS, RARS	5	3	2
Inversions						
inv(3)(q21q26)	Fig. 31.6d	Ribophorin I, EVII	AML-M4, AML-M6, CML Ph+, MDS	3	3	0
inv(14)(q11q32)	Fig. 31.6u	TCRαδ, IgH	T-CLD, T-CLL, T-NHL, T-PLL	0	0	0
inv(16)(p13q22)	Fig. 31.6v	MYH11, CBFß	AML-M4Eo	3	0	3
Duplications						
dup(1)(q12q32)	Fig. 31.6cc		ALL, NHL	0	0	0
dup(1)(q21q32)	Fig. 31.6c		ALL, CLD, NHL	12	5	7
dup(1)(q25q44)			NHL	0	0	0
dup(11)(q13q25)	Fig. 31.6ll		NHL	0	0	0
dup(12)(q13q21)	Fig. 31.6mm		NHL	0	0	0
Deletions						
del(1)(p32p36)	Fig. 31.6bb		CLD, NHL	0	0	0
del(1)(p22)	Fig. 31.6a		AML, ALL, MDS, NHL	0	0	0
del(1)(q21)	Fig. 31.6dd		NHL	0	0	0
del(1)(q32)			ALL, NHL	0	0	0
del(1)(q42)	Fig. 31.6ee		NHL	0	0	0
del(2)(p23)			AML	0	0	0
del(2)(p21)			CLD, NHL	0	0	0
del(2)(q31)	Fig. 31.6ff		AML, CLD	0	0	0
del(3)(p21)	Fig. 31.6gg		ALL, MDS	0	0	0
del(3)(q21)			ALL, AML, NHL	0	0	0
del(4)(p14)	Fig. 31.6e		NHL	0	0	0
del(4)(q21)			NHL	0	0	0
del(5)(p13)	Fig. 31.6hh		NHL	0	0	0
del(5)(q13q33)	Fig. 31.6f		AML, MDS, MPD, 5q-syn	101	32	69
del(6)(p21)	Fig. 31.6h		ALL, AML	0	0	0
del(6)(q13-15q23)	Fig. 31.6i		ALL, CLL, FL, MCL, SLL, T-NHL	21	1	20
del(6)(q21)			ALL, AML, CLL, NHL	0	0	0
del(7)(p13)	Fig. 31.6j		AML, CLD, NHL	0	0	0
del(7)(q11)			ALL, AML, MDS	0	0	0
del(7)(q22q34)	Fig. 31.6k		AML, CLD, MPD, MDS, NHL	29	5	24
del(7)(q32)	Fig. 31.6l		AML, CLL, MDS, NHL	0	0	0
del(8)(p21)			ALL, CLD, NHL	0	0	0
del(8)(q22)			AML, NHL	0	0	0
del(9)(p13)	Fig. 31.6n		ALL, NHL	0	0	0
del(9)(p21)	Fig. 31.6o		ALL, AML, CLD, NHL	0	0	0
del(9)(q11q13-q22)	Fig. 31.6p		AML, MDS	10	2	8
del(10)(p13)			NHL, CLD	0	0	0
del(10)(p12)			AML	0	0	0
del(10)(q22)	Fig. 31.6kk		NHL			
del(10)(q24)			ALL, CLD, NHL	0	0	0
del(11)(p11)			CLD, NHL	0	0	0
del(11)(q13q14-q23)			AML, CLD, CLL, MDS, NHL	0	0	0
del(11)(q23)	Fig. 31.6q		AML, ALL, CLD, CLL, MDS, NHL	19	8	11

del(12)(p12)	Fig. 31.6r		ALL, AML-M2, AML-M4, MDS, NHL	0	0	0
del(12)(p11p12-p13)			AML, MDS	0	0	0
del(12)(q22)	Fig. 31.6nn		NHL	0	0	0
del(13)(q12-q22)	Fig. 31.6t	RB1, D13S25	AML, AMM, CLD, CLL, MM, MDS, NHL	20	9	11
del(13)(q12-q14)			AML, AMM, CLD, CLL, MDS, NHL	0	0	0
del(14)(q24)	Fig. 31.6oo		CLD, NHL	0	0	0
del(15)(q21q22)			NHL	0	0	0
del(16)(q22)	Fig. 31.6pp		AML, AML-M4Eo, NHL	3	2	1
del(17)(p11)		P53	ALL, AML, CLD, MDS, NHL	0	0	0
del(17)(q23)			NHL	0	0	0
del(18)(p11)			AML	0	0	0
del(18)(q21)			AML, CLD, NHL	0	0	0
del(19)(q13)			NHL	0	0	0
del(20)(q11)	Fig. 31.6x		AML, MDS, MPD, PV	65	3	34
del(20)(q11q13)	Fig. 31.6y		AML, MDS, MPD, PV	0	0	0
del(22)(q11)	Fig. 31.6aa		AML, ALL, CLD, MDS, NHL	0	0	0
del(X)(q24)			NHL	0	0	0

[a] The terminology is from the International System for Human Cytogenetic Nomenclature. The following standard notations are used for chromosome anomalies: del, deletion; der, derived; dup, duplication; i, isochromosome; idic, isodicentric; ins, insertion translocation; inv, inversion; p, short arm; q, long arm; t, translocation. The numbers within the first set of parentheses adjacent to the type of anomaly identify the chromosome or chromosomes involved. The numbers within the second set of parentheses identify the breakpoints. Thus, the expression t(9;22)(q34;q11.2) indicates a reciprocal translocation between parts of chromosomes 9 and 22 with breakpoints at 9q34 and 22q11.2.

[b] ALL, acute lymphocytic leukemia (subtype L3 where indicated); AML, acute myeloid leukemia (subtype M1-M7 where indicated); AMM, agnogenic myeloid metaplasia; ALCL, anaplastic large cell lymphoma; ATL, adult T-cell lymphoma; B-CLL, B-cell chronic lymphocytic leukemia; BL, Burkitt's lymphoma; B-NHL, B-cell non-Hodgkin's lymphoma; B-PLL, B-cell prolymphocytic leukemia; CLD, chronic lymphoproliferative disorder; CLL, chronic lymphocytic leukemia; CML Ph+, chronic myeloid leukemia with Philadelphia chromosome; CMML, chronic myelomonocytic leukemia; DLCL, diffuse large cell lymphoma; FL, follicular lymphoma; HSγδTL, hepatosplenic gamma delta T-cell lymphoma; HCL, hairy-cell leukemia; LPL, lymphoplasmacytic lymphoma; MCL, mantle-cell lymphoma; MDS, myelodysplastic syndrome; MM, multiple myeloma; MPD, myeloproliferative disorder; MGUS, monoclonal gammopathy of undetermined significance; MZL, mantle zone lymphoma; NHL, Non-Hodgkin's lymphoma; pre-B-ALL, precursor B-cell acute lymphocytic leukemia; PV, polycythemia vera; RARS, refractory anemia with ringed sideroblasts; SLVL, splenic lymphoma with villous lymphocytes; T-ALL, T-cell acute lymphocytic leukemia; T-CLL, T-cell chronic lymphocytic leukemia; T-CLD, T-cell chronic lymphoproliferative disorder; T-NHL, T-cell non-Hodgkin's lymphoma; T-PLL, T-cell prolymphocytic leukemia; t-AML, therapy-related acute myeloid leukemia; t-MDS, therapy-related myelodysplastic syndrome; 5q- syn, 5q- syndrome.

[c] Frequency of each chromosome abnormality seen in bone marrow specimens among 748 patients with an abnormal clone seen at Mayo Clinic. Because it was not part of our cytogenetic practice to study lymph nodes or patients with most lymphoproliferative disorders during the time period of this study, the data are more representative of acute leukemia, myelodysplastic syndromes, and myeloproliferative disorders. Moreover, physicians at Mayo Clinic primarily see adult patients with hematologic disorders. Data from Dewald GW, Noel P, Dahl RJ, et al., Chromosome abnormalities in malignant hematologic disorders. *Mayo Clin Proc* 1985;60:675–689.

Since 1985, many other chromosome abnormalities associated with hematologic disorders have been reported (4). Many of the newly discovered chromosome abnormalities were either relatively rare or cryptic in nature. Eventually, the clinical disorder(s) associated with each specific chromosome abnormality will be correlated. Currently, physicians and cytogeneticists who interpret chromosome abnormalities need to correlate the cytogenetic results with the clinical scenario of the patient.

Several important international attempts have been made to review the cytogenetic literature in order to correlate chromosome abnormalities with specific hematologic disorders (20, 21 and 22). Most recently, breakpoints for 1,803 structural abnormalities including 215 balanced and 1,588 unbalanced anomalies associated with hematologic disorders and solid tumors were summarized (22).

A list of the most common structural chromosome aberrations in hematologic disorders is shown in Table 32.1 and Table 32.2. Examples of many of these abnormalities are shown in Figure 32.5 and Figure 32.6, and are cross-indexed with the chromosome abnormalities listed in Table 32.1.

SPECIFIC CYTOGENETIC ABNORMALITIES IN HEMATOLOGIC DISORDERS

TABLE 32.2. COMMON AnEUPLOID ABNORMALITIES IN HEMATOLOGIC DISORDERS

Type of Anomaly	No. of Specimens			
	Total	Only Abnormality	Plus Other Abnormalities	Reported Associated Disorders[a]
Near haploid	1	1	0	ALL
Monosomy				
–5	25	0	25	MDS, AML, t-AML
–7	97	21	76	CML in blast crisis, MDS, AML, t-AML
–Y	118	66	52	Age-associated loss and CML, MDS, AML
–X	20	4	16	AML
Trisomy				
+3	5	0	5	ATL, SLVL
+7	10	1	9	ATL, NHL
+8	161	55	106	CML in blast crisis, MDS, AML, t-AML
+9	20	4	16	MM, t-AML, PV, Various disorders
+11	20	2	18	MM
+12	5	0	5	CLL, HCL, NHL
+15	9	3	6	MM
+18	8	0	8	NHL
+19	31	4	27	CML in blast crisis, AML
+21	53	8	45	CML in blast crisis, MDS, AML

[a] For explanation of abbrevations, see footnote b in Table 32.1.

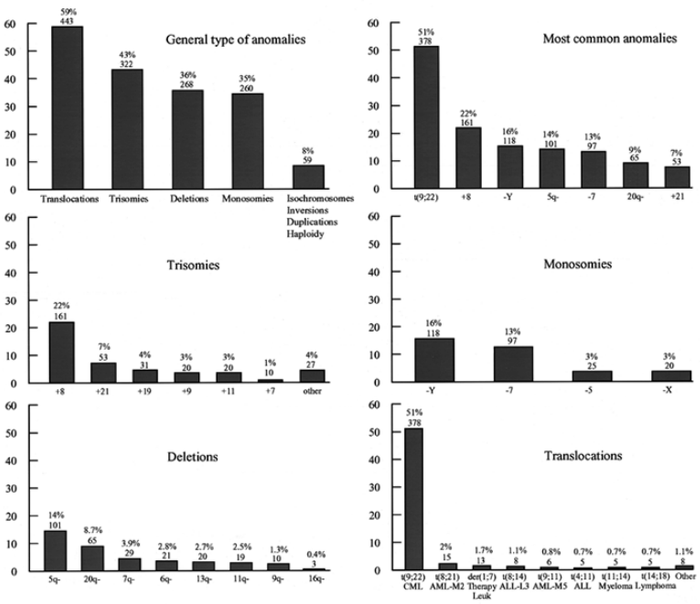


FIGURE 32.4. Frequency of different kinds of chromosome abnormalities in hematologic disorders. ALL, acute lymphocytic leukemia (L3 subtype); AML, acute myeloid leukemia (M2 and M5, subtypes); CML, chronic myeloid leukemia; Therapy Leuk, therapy-related leukemia.

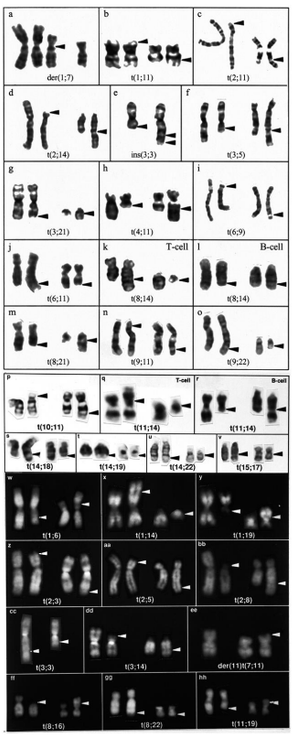


FIGURE 32.5. Examples of common chromosome translocations in hematologic disorders. See Table 32.1 for more information about the breakpoints, associated hematologic disorders, associated genes, and frequency among 748 patients with a chromosomally abnormal clone.

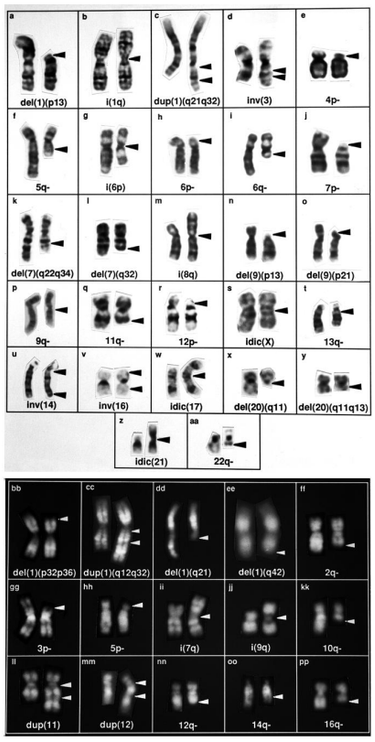


FIGURE 32.6. Examples of common isochromosomes, deletions, inversions, and duplications in hematologic disorders. See Table 32.1 for more information about the breakpoints, associated hematologic disorders, associated genes, and frequency among 748 patients with a chromosomally abnormal clone.

Neoplastic hematologic disorders have been identified in nearly every cell type that comprise the normal bone marrow and lymphoid systems. Not surprisingly, classification of hematologic disorders has evolved according to these primary cell types. Briefly, neoplastic hematologic disorders are grouped into myeloid and lymphoid cell types. Neoplastic myeloid disorders include acute myelogenous leukemias or AML (≥30% myeloblasts in the bone marrow), myelodysplastic syndromes, and chronic myeloproliferative disorders. Each of these general categories is further subdivided into several clinical and/or morphological neoplasms, which are best summarized by the French-American-British (FAB) classification system (Table 32.3, Table 32.4, and Table 32.5)(23, 24).

TABLE 32.3. CLASSIFICATION SYSTEM FOR ACUTE LEUKEMIAS (24)

Acute myeloid leukemia[a]			
FAB Subtype	Frequency in Children	Frequency in Adults	
M0	<5%	2%–3%	Agranular myeloblasts; <3% blasts reactive to myeloid antigens
M1	11%	20%	Lack of maturation beyond the blast stage, ≥3% blasts reactive to myeloid antigens
M2	32%	25%–30%	Maturation to the progranulocyte state or beyond
M3	8%	8%–15%	Abnormal (hypogranular or hypergranular) progranulocytes that display bundles of Auer rods
M4	24%	20%–25%	Blasts with both monocytic and granulocytic characteristics. Monocytes exceed 20% of the nucleated cell population
M5a	11%	5%	Monoblastic cells >50% monocytic differentiation
M5b			Monocytic cells >50% monocytic differentiation
M6	5%	5%	Erythrocytic cells predominately erythroid precursors
M7	9%	1%–2%	Megakaryocytic cells predominately megakaryocytes precursors
Acute lymphocytic leukemia[b]			
FAB Subtype	Frequency in Children	Frequency in Adults	
L1	80%	30%	Small blasts, regular nuclear outline, few or no visible nucleoli, scant cytoplasm
L2	15%–20%	65%–70%	Larger blasts (x2 small lymphocyte), irregular or cleft nucleoli, one or more prominent nucleoli, variable to abundant cytoplasm
L3	3%	3%	“Burkitt's” type, large cells, dense chromatin, round or oval nucleus, moderately abundant cytoplasm, deeply basophilic cytoplasm with prominent vacuoles

[a] Acute myeloid leukemia is classified into eight categories on the basis of cell morphology.
[b] Acute lymphocytic leukemia is classified into three categories on the basis of cell morphology.

Neoplastic lymphoid disorders include acute lymphocytic leukemias (ALL) (≥30% lymphoblasts in the bone marrow), lymphomas, and chronic lymphoproliferative disorders (Table 32.3, Table 32.6). The main division of lymphomas and chronic lymphoproliferative disorders results in two main categories: Hodgkin's disease and non-Hodgkin's lymphoma.

Although non-Hodgkin's lymphoma can be broadly divided into B-cell or T-cell disorders, these subclassifications have had many discrepant and conflicting classification schemes including the Rappaport, Kiel, Lukes-Collins, and Working Formulation (25, 26, 27 and 28). The current classification scheme utilized by the majority of hematologists is the Revised European-American Lymphoma (REAL) classification (Table 32.6)(29). However, a new lymphoma classification scheme by the World Health Organization (WHO) is now available (30).

Because the determination of cell morphology often is difficult, most physicians use supplemental methods such as cytogenetics, cytochemistry, immunologic markers, biochemical markers, and B-cell or T-cell rearrangement studies to help classify hematologic disorders. Much effort has been expended by cytogeneticists, hematologists, and pathologists attempting to correlate the findings of cytogenetic studies with the hematologic classification systems.

Since 1978, six International Workshops on Chromosomes in Leukemia and Lymphoma have been convened for this purpose (31, 32, 33, 34, 35 and 36). These workshops have demonstrated a strong correlation between specific chromosome abnormalities and specific clinical disorders that have helped document the biological validity of the classification schemes. However, these workshops identified many other chromosome abnormalities that do not

TABLE 32.4. FAB CLASSIFICATION SYSTEM FOR MYELODYSPLASTIC SYNDROMES (23, 24)

Subtype	Peripheral Blood	Bone Marrow
Refractory anemia (RA)	<1% blasts	Erythroid hyperplasia
		Dyserythropoiesis
		<5% blasts
		No or few ringed sideroblasts
Refractory anemia with ringed sideroblasts (RARS)	<1% blasts	Erythroid hyperplasia
	Occasional dysplastic neutrophils and platelets	
		>15% ringed sideroblasts
Refractory anemia with excess blasts (RAEB)	Cytopenias	Panmyeloid hyperplasia
		5%–20% blasts
	<5% blasts	
Refractory anemia with excess blast in transition (RAEBIT)	Cytopenias	Panmyeloid hyperplasia
		20%–30% blasts
		Auer rods
	5–29% blasts	
Chronic myelomonocytic leukemia (CMML)	>1 × 10^9/L monocytes	<20% blasts
		Increase in monocytes and precursors
	<5% blasts	
Myelodysplastic syndrome unclassified[a]	Cytopenias	<5% blasts
(MDS-U)	<1% blasts	Multilineage dysplasia

[a] Proposed WHO classification “refractory cytopenia with multilineage dysplasia” or RCMD.

TABLE 32.5. CLASSIFICATION FOR CHRONIC MYELOPROLIFERATIVE DISORDERS

Predominant Abnormal Cell Line	Myeloproliferative Disorder
Erythroid	Polycythemia vera (PV)
Granulocytic	Chronic myeloid leukemia (CML)
	Chronic myelomonocytic leukemia (CMML)
Megakaryocytic	Essential thrombocytemia (ET)
Fibroblast	Agnogenic myeloid metaplasia (AMM)

correlate well with the current classification schemes. In the future, classification schemes are expected to become more based on cytogenetic and/or molecular genetic abnormalities.

Herein, an outline of the current cytogenetics literature is provided for each of the main hematologic disorders including acute leukemia, therapy-related AML and therapy-related myelodysplasia, myelodysplastic syndromes, chronic myeloproliferative disorders, lymphomas and chronic lymphoproliferative disorders, and Hodgkin's disease. For each specific cytogenetic anomaly described, an attempt was made to present a concise summary of main disease associations, unique clinical features, median age at diagnosis, cytogenetic variants, associated cytogenetic anomalies, FISH strategies, gene(s) involved, and molecular testing strategies.

Acute Leukemia (Table 32.3)(37, 38, 39, 40, 41, 42, 43, 44, 45, 46 and 47)

- Among children with acute leukemia, 85% have ALL and 15% have AML
- Among adults with acute leukemia, 20% have ALL and 80% have AML
- Yearly incidence of AML: 1.2/100,000 (children), 3.5/100,000 (age 50), 15/100,000 (age 70), and 35/100,000 (age 90)

TABLE 32.6. REVISED EUROPEAN-AMERICAN LYMPHOMA (REAL) CLASSIFICATION SYSTEM (29)

B-cell neoplasms
Precursor B-cell neoplasm
Precursor B-lymphoblastic leukemia and lymphoma
Peripheral B-cell neoplasms
B-cell chronic lymphocytic leukemia prolymphocytic leukemia, small lymphocytic lymphoma
Lymphoplasmacytoid lymphoma, immunocytoma
Mantle cell lymphoma
Follicle center lymphoma, follicular lymphoma
Provisional cytologic grades: I (small cell), II (mixed small and large cells), III (large cells)
Provisional subtype: diffuse, predominantly small cell type
Marginal zone lymphoma
Extranodal (mucosa-associated lymphoid tissue type with or without monocytoid B-cells)
Provisional subtype: nodal (with or without monocytoid B-cells)
Provisional entity: splenic marginal zone lymphoma (with or without villous lymphocytes)
Hairy cell leukemia
Plasmacytoma, plasma cell myeloma
Diffuse large B-cell lymphoma[a]
Subtype: primary mediastinal (thymic) B-cell lymphoma
Burkitt's lymphoma
Provisional entity: high-grade B-cell lymphoma, Burkitt's like[a]
T-cell and putative natural killer cell neoplasms
Precursor T-cell neoplasm
Precursor T-lymphoblastic leukemia or lymphoma
Peripheral T-cell and natural killer cell neoplasms
T-cell chronic lymphocytic leukemia, prolymphocytic leukemia
Large granular lymphocyte leukemia
T-cell type
Natural killer cell type
Mycosis fungoides, Sézary syndrome
Peripheral T-cell lymphomas, unspecified[a]
Provisional cytologic categories: medium-sized cell, mixed medium and large cell, large cell, lymphoepithelioid
Provisional subtype: hepatosplenic γδ T-cell lymphoma
Provisional subtype: subcutaneous panniculitis T-cell lymphoma
Angioimmunoblastic T-cell lymphoma
Angiocentric lymphoma
Intestinal T-cell lymphoma (may be enteropathy associated)
Adult T-cell lymphoma or leukemia
Anaplastic large-cell lymphoma (CD30+, T-, and null-cell types)
Provisional entity: anaplastic large-cell lymphoma, Hodgkin's-like
Hodgkin's disease
Type I. Lymphocyte predominance
Type II. Nodular sclerosis
Type III. Mixed cellularity
Type IV. Lymphocyte depletion
Type V. Provisional entity: lymphocyte-rich classic Hodgkin's disease

[a] These categories are thought likely to include more than one disease entity.

- Yearly incidence of ALL: 9.8/100,000 (children), 12.8/100,000 (adults)
- AML (> 55 years of age) often have chromosome abnormalities associated with progression of myelodysplastic syndromes
- AML (< 55 years of age) often have chromosome abnormalities linked with de novo AML

t(1;14)(p32;q11) (Fig. 32.5X)

- Specific for T-ALL in 10% to 30% of patients
- Often young males with mediastinal mass
- Positive cerebro-spinal fluid and high white count are common
- 1p32 breakpoint in TAL1; 14q11 breakpoint in TCR α or δ (T-cell receptor genes)
- Commercial FISH strategy is not available, but possible to detect TAL1/TCRαδ fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(1;19)(q23;p13) (Fig. 32.5Y)

- 30% of children with Pre-B-cell ALL
- Low leukocyte count
- Cytogenetic variant: t(17;19)(q21-22;p13)
- Poor to intermediate prognostic significance
- 1q23 breakpoint in PBX1; 19p13 breakpoint in E2A
- Commercial FISH strategy is not available, but “home-brew” methods have been reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(1;22)(p13;q13)

- Specific for AML-M7 in 25% to 35% of patients
- Found predominantly in infants
- Associated with organomegaly from leukemic infiltration
- Poor prognostic significance
- Commercial FISH strategy is not available
- Gene breakpoints are undefined so molecular tests are not available

t(3;5)(q25.1;q35) (Fig. 32.5F)

- Associated with AML-M6
- Young age
- 3q25.1 breakpoint in MLF1; 5q34 breakpoint in NPM
- Commercial FISH strategy is not available, but possible to detect MLF1/NPM fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(3;21)(q26;q22) (Fig. 32.5G)

- AML, myelodysplastic syndromes, and secondary anomaly in chronic myeloid leukemia
- 3q26 breakpoint in EAP; 21q22 breakpoint in AML1
- Commercial FISH strategy is available to detect break apart of AML1 locus
- Gene breakpoints are undefined so molecular tests are not available

t(6;9)(p23;q34) (Fig. 32.5I)

- 1% to 2% of patients with AML (AML-M2, M4 and rarely M1); refractory anemia with excess blasts
- May be associated with bone marrow basophilia
- Disease onset 20 to 40 years old
- Poor to intermediate prognostic significance
- 6p23 breakpoint in DEK; 9q34 breakpoint in CAN
- Commercial FISH strategy is not available, but possible to detect DEK/CAN fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(8;14)(q24;q32) (Fig. 32.5L)

- 100% of patients with ALL-L3 and Burkitt's lymphoma (different presentations of same disease)
- 5% of patients with adult ALL
- For more information see section on non-Hodgkin's lymphoma

t(8;16)(p11;p13) (Fig. 32.5FF)

- Associated with AML-M4 and M5b
- Found mainly in infants and children
- Associated with disseminated intravascular coagulation (DIC) and erythrophagocytosis
- Poor prognostic significance
- Commercial FISH strategy is not available
- Gene breakpoints are undefined so molecular tests are not available

t(8;21)(q22;q22) (Fig. 32.5M)

- 5% to 10% of AML; 10% to 30% of AML-M2; 7% of AML-M4; 20% of childhood AML
- Disease onset is 25 to 30 years
- Generally good prognostic factor, although long-term survival in children is poor
- 70% of males have associated loss of the Y chromosome (likely no prognostic importance)
- 60% of females have associated loss of the X chromosome (likely no prognostic importance)
- t(16;21)(q24;q22) may be a rare variant
- 8q22 breakpoint in ETO; 21q22 breakpoint in AML1
- Commercial FISH strategy available to detect fusion of ETO and AML1
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(9;22)(q34;q11.2) (Fig. 32.5O)

- “Philadelphia chromosome,” (Ph)
- 5% of children with ALL; 25% to 35% of adults with ALL (usually L2); 1% of AML-M1
- Poor prognostic significance
- 50% of adult ALL and nearly 100% of childhood ALL have fusion gene e1a2
- For more information see section on chronic myeloproliferative disorders

t(10;14)(q24;q11)

- 1% to 3% of patients with T-cell ALL
- 10q24 breakpoint in HOX11; 14q11 breakpoint in TCRαδ (T-cell receptor αδ)
- Commercial FISH strategy is not available, but possible to detect HOX11/TCRαδ fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(11;var)(q23;var)

- Group of translocations with breakpoint at 11q23 and various partner chromosomes
- Mostly children and young adults
- 70% of children <1 year old with acute leukemia; 3% of patients with therapy related AML
- Prognostic significance is poor in ALL and intermediate in AML
- t(4;11)(q21;q23) (Fig. 32.5H) occurs in 5% to 8% of children with ALL who are <1 year of age
- t(9;11)(p22;q23) (Fig. 32.5N) or t(11;19)(q23;p13) (Fig. 32.5HH) occurs in 35% of patients with AML-M5 and 50% of patients with AML-M5a
- t(10;11)(p13;q23) (Fig. 32.5P) occurs in 30% of patients with AML-M4 or M5
- t(X;11)(q13;q23) in T-cell ALL
- Over 30 different chromosome partners described
- 11q23 breakpoint in MLL (mixed lineage leukemia gene)
- 1p32 breakpoint in AF1P; 4q21 in AF4; 9p22 in AF9; 10p13 in AF10; 19p13.1 in ELL; 19p13.3 in ENL, Xq13 in AFX1
- Tandem duplications of exons 2 to 6 in MLL account for 10% of cytogenetically normal AML
- Commercial FISH strategy is available to detect break apart of MLL gene
- Breakpoints in MLL cluster in an 8.3-kb region detected by Southern Blot or RT-PCR analysis

t(12;21)(p13;q22)

- 25% of children with ALL; 3% of adults with ALL
- Good prognostic significance
- Difficult to detect by conventional cytogenetic studies
- Often an unbalanced translocation with loss of abnormal chromosome 12
- 12p13 breakpoint in TEL; 21q22 breakpoint in AML1
- Commercial FISH strategy is available to detect fusion of TEL and AML1
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(15;17)(q22;q21) (Fig. 32.5V)

- Specific for AML-M3 (also known as acute promyelocytic leukemia, APL)
- Accounts for up to 30% to 40% of AML in certain ethnic groups
- 20% patients have a hypogranular or microgranular variants (M3v)
- Median age of onset is 38 years
- Patients frequently present with disseminated intravascular coagulation (DIC)
- Good prognostic significance; respond to All- *trans* -retinoic acid treatment
- 1% of patients have cryptic translocations or variants
- Most common variants include t(11;17)(q23;q21), t(5;17)(q35;q21), t(1;17)(p36;q21)
- 15q22 breakpoint in PML (aka MYL); 17q21 breakpoint in RARα (retinoic acid receptor)
- Commercial FISH strategy is available to detect PML/RARα fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

inv(3)(q21q26) (Fig. 32.6D)

- AML-M2, M4, and M6; 2% of patients with AML
- Frequently associated with thrombocytosis (rare in other forms of AML)
- Poor prognostic significance
- Cytogenetic variants include t(3;3)(q21;q26) (Fig. 32.5CC) and ins(3;3)(q26;q21q26) (Fig. 32.5E)
- 3q21 breakpoint in Ribophorin; 3q26 breakpoint in EVI1
- Commercial FISH strategy is not available, but “home-brew” methods have been reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

inv(16)(p13q22) (Fig. 32.6V)

- This anomaly or a variant occurs in 100% of patients with AML-M4Eo
- 25% of patients with AML-M4; 6% of patients with AML
- Clinically associated with bone marrow eosinophilia, lymphadenopathy, and hepatomegaly
- Good prognostic significance
- Most common cytogenetic variant is t(16;16)(p13;q22)
- del(16)(q22) (Fig. 32.6pp) linked with AML-M4Eo, but clinical phenotype may differ
- 16p13 breakpoint in MYH11; 16q22 breakpoint in CBFß (aka PEBP2B)
- Commercial FISH strategy is available to detect break apart of CBFß locus

- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

del(5)(q13q33) (Fig. 32.6F)

- 10% of patients with AML; usually previous myelodysplastic syndromes
- Poor prognostic significance
- See section on myelodysplastic syndromes for more information

del(12)(p12) (Fig. 32.6R)

- 2% of patients with AML, most have M2 or M4
- Associated with basophilia
- Poor prognostic significance
- Frequently associated with monosomy 7
- Gene at 12p13.1 breakpoint is unknown, but TEL or KIP1 (p27 protein) frequently lost
- Commercial FISH strategy available to detect loss of TEL1
- Gene breakpoints are undefined so molecular tests are not available

Hyperdiploidy without apparent structural anomalies

- 25% to 30% of children with ALL, often non-T-cell, non-B-cell
- Neoplastic clones usually have 50 to 60 chromosomes
- Most common trisomies include +4, +6, +10, +14, +17, +18, +20, +21, +X
- Trisomy 21 or tetrasomy 21 occurs in nearly 100% of patients
- Good prognostic significance, especially when +4 and +10 are involved
- Commercial FISH strategy is available using centromere-specific probes to detect aneuploidy
- Molecular strategy is not applicable

Near Haploidy

- 1% of children with ALL
- Neoplastic clones with 26 to 36 chromosomes
- Diploid copies of chromosomes 10, 14, 18, 21 and X or Y are common
- Poor prognostic significance
- Commercial FISH strategy is available using centromere-specific probes to detect aneuploidy
- Molecular strategy is not applicable

Therapy-related leukemia (t-AML) and therapy-related myelodysplasia (t-MDS)(38, 41, 48, 49 and 50)

Three cytogenetic categories: hypodiploid with abnormalities of chromosomes 5 and/or 7, der(1;7) translocations, and t(11;var)(q23;var)

- Among 3,976 patients with hematologic disorders, 121 were hypodiploid with abnormalities of chromosomes 5 and/or 7, 39 had der(1;7) translocations and 29 had t(11;var)(q23;var)
- Up to 20% of patients with aggressive chemotherapy or radiation treatment therapies
- Same chromosome abnormalities associated with exposure to environmental carcinogens
- 10% to 20% of AML
- Chromosomally abnormal clones identified in 90% of patients with t-MDS or t-AML

Hypodiploid and Abnormalities of Chromosomes 5 and 7

- Modal number of chromosomes is <46
- Associated with prior treatment with alkylating chemotherapeutic agents or radiation
- Time from treatment to bone marrow dysfunction is 2 to 5 years
- Patients usually present with myelodysplastic syndromes
- 50% of patients with t-MDS progress to AML in 6 months
- Poor prognostic significance
- Critical region is most likely in 5q31, 7q11-q21 and/or 7q31-q36
- FISH strategy available to detect structural and numeric anomalies of chromosomes 5 and 7
- Molecular strategy is not applicable

der(1;7)(q10;p10) (Fig. 32.5A)

- Unbalanced translocation with loss of the chromosome 7 q-arm and three copies of the chromosome 1 q-arm
- 70% of patients have a therapy-related hematologic disorder
- Associated with prior treatment with alkylating chemotherapeutic agents or radiation
- Time from treatment to bone marrow dysfunction is 2 to 5 years
- Patients usually present with myelodysplastic syndromes and rarely progress to AML
- Pathogenesis may be related to loss of tumor suppressor genes on the chromosome 7 q-arm
- Commercial FISH strategy is available to detect loss of chromosome 7 q-arm and fusion of the centromere of chromosome 1 and 7
- Molecular strategy is not applicable

t(11;var)(q23;var)

- Translocations with breakpoints at 11q23 and various partner chromosomes
- Due to Topoisomerase II inhibitor chemotherapy such as Doxorubicin and Etoposide
- No myelodysplastic syndrome prodrome
- AML develops 1 to 2 years following treatment
- Favorable response to subsequent chemotherapy
- Most common translocations are t(9;11)(p22;q23) (Fig. 32.5N) and t(10;11)(p13;q23) (Fig. 32.5P)
- 11q23 breakpoint in MLL (mixed lineage leukemia gene)
- For more information see section on acute leukemia

Myelodysplastic Syndromes (Table 32.4) (38, 40, 41, 48, 51, 52, 53 and 54)

- "5q- syndrome" included in this section, although not officially recognized as an FAB subgroup
- Overall, 40% of patients with myelodysplastic syndromes have a chromosomal anomaly
- Chromosomally abnormal clones are found in 31% of patients with refractory anemia, 29% with refractory anemia with ringed sideroblasts, 46% with refractory anemia with excess blasts, 50% with refractory anemia with excess blasts in transition, 45% with chronic myelomonocytic leukemia, and 100% with 5q- syndrome
- Many patients with myelodysplastic syndromes progress to acute leukemia

del(5)(q13q33) (Fig. 32.6F)

- "5q- syndrome" accounts for one-third of patients with a 5q- chromosome
  - Subset of young women with 5q- as sole abnormality
  - Small megakaryocytes and abnormal erythropoiesis
  - Good prognosis
  - Chromosome evolution and progression to AML is uncommon
- Two-thirds patients with 5q- have various myelodysplastic syndromes subtypes
  - 50% patients have additional chromosome abnormalities with associated worse prognosis
  - Common variants include del(5)(q31q35) and del(5)(q22q33)
  - Breakpoint at 5q33 occurs in 70% of patients
  - Most patients lack 5q31; candidate genes include EGR1, CSF1R, and IRF1
  - Commercial FISH strategy is available to detect loss of EGR1 and CSF1R
  - Molecular strategy is not applicable

idic(X)(q13) (Fig. 32.6S)

- 0.7% of patients with adult myeloid disorders; specifically associated with refractory anemia with ringed sideroblasts
- Affects only women, mostly >60 years old
- Rarely progresses to AML
- Variants include translocations with breakpoints at Xq13
- Subclones with two copies of the idic(X) are frequent
- ABC7 gene associated with familial refractory anemia with ringed sideroblasts is at Xq13
- Commercial FISH strategy is not available, but probes to detect loss of ABC7 gene is possible
- Molecular strategy is not applicable

i(17q) (Fig. 32.6W)

- Most common isochromosome in hematologic disorders
- Usually associated with complex karyotypes
- Primary anomaly in myelodysplastic syndromes and chronic myeloproliferative disorders
- Isolated i(17q) has poor prognostic significance and progression to AML is frequent
- Pathogenesis most likely related to loss of a tumor suppressor gene such as p53 at 17p13
- Commercial FISH strategy is available to detect the chromosome 17 centromere and q-arm
- Molecular strategy is not applicable

Chronic Myeloproliferative Disorders (Table 32.5) (37, 38, 41, 55)

t(5;12)(q33;p13)

- 2% to 5% of patients with chronic myelomonocytic leukemia
- Associated with myelodysplastic syndromes or chronic myeloproliferative disorder
- Variant translocation partners of 12p13 include 3q26 and 10q24
- 5q33 breakpoint in PDGFRB; 12p13 breakpoint in TEL
- Commercial FISH strategy is available to detect abnormalities of TEL
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(9;22)(q34;q11.2) (Fig. 32.5O)

- "Philadelphia Chromosome," (Ph)
- Ph-chromosome or a variant occurs in 100% of patients with chronic myeloid leukemia
- "Ph-negative CML" does not exist
- Chronic myeloid leukemia also known as chronic granulocytic leukemia
- Variant translocations in 5% to 8% of patients can be complex (involving two or more chromosomes) or cryptic (submicroscopic insertion translocations)
- More than 90% of patients with chronic myeloid leukemia have t(9;22) in all metaphases
- Chronic phase has no other cytogenetic changes
- Myeloid blastic crisis has >50% with additional cytogenetic changes such as multiple copies of Ph-chromosome, +8, i(17q), and +19
- Lymphoid blastic crisis often has anomalies of chromosome 7
- 9q34 breakpoint in ABL; 22q11.2 breakpoint in BCR (breakpoint cluster region)
- Fusion gene usually is b2a2 or b3a2
- Commercial FISH strategies are available to identify BCR/ABL fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis
- For more information on Ph-chromosome in ALL and AML see section on acute leukemia

Lymphomas and Lymphoproliferative Disorders (Table 32.6)

- Classified into non-Hodgkin's lymphoma (T-cell or B-cell) and Hodgkin's disease
- T-cell receptor genes on chromosomes 7 and 14 usually linked with T-cell disorders (Fig. 32.7)

- Light-chain immunoglobulin genes on chromosomes 2 and 22, and heavy-chain immunoglobulin gene on chromosome 14 usually associated with B-cell disorders (Fig. 32.7)

Non-Hodgkin's Lymphoma (Table 32.6) (37, 39, 40 and 41, 56, 57, 58, 59, 60, 61, 62 and 63)

- 80% of patients have detectable chromosome abnormalities

t(2:5)(p23;q35) (Fig. 32.5AA)

- Most likely occurs in 100% of patients with anaplastic large cell lymphoma (Ki-1/CD30+)
- Diverse clinical presentation
- Tetraploidy is associated with poor prognosis
- High frequency in childhood non-Hodgkin's lymphoma
- Variants: t(1;2)(q21;p23), t(1;2)(q25;p23), t(2;3)(p23;q21), t(2;13)(p23;q34), inv(2)(p23q35)
- 2p23 breakpoint in ALK; 5q35 breakpoint in NPM
- Commercial FISH strategy is available to detect break apart of ALK gene
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(3;14)(q27;q32)

- Associated with diffuse large cell lymphoma and 5% to 10% of patients with follicular lymphoma
- Variants include t(3;22)(q27;q11) and t(2;3)(p12;q27)
- 3q27 breakpoint in BCL6; 14q32 breakpoint in IgH
- Commercial FISH strategy is not available, but “home-brew” methods have been reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(8;14)(q24;q32) (Fig. 32.5L)

- 100% of patients with ALL-L3 (Burkitt's leukemia) and Burkitt's lymphoma (different presentations of same disease); occasionally, diffuse large cell lymphoma
- May be positive or negative for Epstein Barr virus

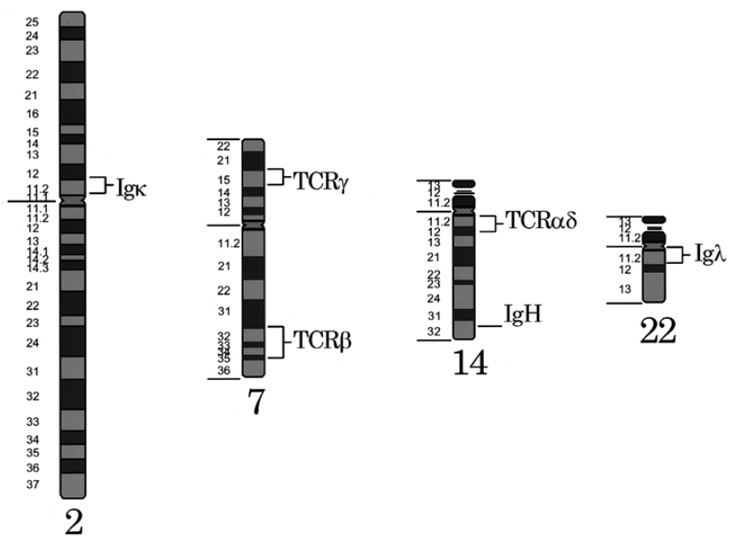


FIGURE 32.7. The immunoglobulin loci on chromosomes 2, 14, and 22; chromosome abnormalities involving any of these sites usually are B-cell disorders. The T-cell receptor (TCR) loci on chromosomes 7 and 14; chromosome abnormalities involving these sites usually are T-cell disorders.

- Variants include t(2;8)(p12;q24) (Fig. 32.5BB) and t(8;22)(q24;q11) (Fig. 32.5GG)
- 8q24 breakpoint in c-MYC; 14q32 in IgH; 2p12 in Igκ; 22q11 in Igλ
- With endemic Burkitt's, the breakpoint is 5′ to c-MYC gene
- Sporadic or AIDS associated Burkitt's, the breakpoint is within c-MYC gene
- Commercial FISH strategy unavailable, but “home-brew” methods to detect c-MYC/IgH fusion reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(9;14)(p13;q32)

- Low-grade NHL with plasmacytoid differentiation
- 9p13 breakpoint in PAX5; 14q32 breakpoint in IgH locus
- Commercial FISH strategy is not available, but possible to detect PAX5/IgH fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(11;14)(q13;q32) (Fig. 32.5R)

- 5% to 10% of patients with non-Hodgkin's lymphoma
- Variants include t(11;22)(q13;q11) and t(2;11)(q13;p13)
- Mantle-cell lymphoma
  - 100% of patients with mantle-cell lymphoma
  - 11q13: 80% of breakpoints in major translocation cluster (MTC), 5′ of Cyclin D1 gene (CCND1)
  - 14q32 breakpoint adjacent to IgH enhancer region
- Multiple myeloma
  - Often prodrome of monoclonal gammopathy of undetermined significance (MGUS)
  - 30% to 50% of patients with multiple myeloma (using FISH)
  - 14q32 breakpoint in IgH usually in switch region
  - 70% of patients with multiple myeloma have a breakpoint at 14q32
  - Variants include t(4;14)(p16.3;q32) and t(14;16) (q32;q23)
  - Other translocation partners include chromosomes 6, 8, and 11
  - 11q13 breakpoint in Cyclin D1 gene is variable
  - 1q abnormalities are common secondary anomalies
  - Evolution to plasma cell leukemias is associated with complex cytogenetic abnormalities
- B-cell prolymphocytic leukemia
  - 20% of patients with prolymphocytic leukemia
  - Cases may represent erroneously classified mantle cell leukemia
- B-cell chronic lymphocytic leukemia is not associated with t(11;14)(q13;q32), and when observed in bone marrow or blood, mantle-cell leukemia should be suspected
- 11q13 breakpoint in Cyclin D1 (CCND1, PRAD1, BCL1); 14q32 breakpoint in IgH
- Commercial FISH strategies available to detect fusion of IgH and Cyclin D1
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(11;18)(q21;q21)

- 25% to 50% marginal zone lymphomas (low grade mucosa-associated lymphoid tissue, MALT)
- Good prognostic significance
- 11q21 breakpoint in API2; 18q21 breakpoint in MLT
- Commercial FISH strategy not available, but “home brew” methods have been reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(14;18)(q32;q21) (Fig. 32.5S)

- 80% of patients with follicular small-cleaved cell; 39% of patients with diffuse large cell
- Good prognostic significance
- Together with +17, +12, or del(1)(p32p36) denotes poor prognostic significance
- 14q32 breakpoint in IgH; 18q21 breakpoint in BCL2
- Two breakpoints in BCL2 gene: 50% to 60% in MBR (major breakpoint region) and 25% to 30% in MCR (minor cluster region)
- Commercial FISH not available, but “home-brew” methods for BCL2/IgH fusion reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

t(14;19)(q32;q13) (Fig. 32.5T)

- B-cell chronic lymphocytic leukemia and small lymphocytic lymphoma (different presentations of same disease)
- Poor prognostic significance
- Most patients less than 40 years old
- Associated with trisomy 12 in 50% of patients
- 14q32 breakpoint in IgH; 19q13 breakpoint in BCL3
- Commercial FISH strategy not available, but possible to detect IgH/BCL3 fusion
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

del(6)(q14q23) (Fig. 32.6I)

- 80% of patients with large cell lymphoma; common in chronic lymphocytic leukemia and T-cell lymphoma
- Good prognostic significance in chronic lymphocytic leukemia
- Candidate gene is MYB
- Commercial FISH strategy not available, but “home-brew” methods for loss of MYB reported
- Molecular strategy is not applicable

inv(14)(q11q32) (Fig. 32.6U)

- 70% of patients with T-cell prolymphocytic leukemia
- Variant: t(14;14)(q11;q32)

- Associated with i(8q)
- 14q11 breakpoint in TCR αδ gene; 14q32 breakpoint in IgH locus
- Commercial FISH strategy in development and “home brew” methods reported
- Defined gene breakpoints are amenable to Southern Blot or RT-PCR analysis

Hodgkin's Disease (Table 32.6) (64, 65)

- Classic Hodgkin's and Reed-Sternberg cells are commonly identified
- 40% of patients with Hodgkin's disease have chromosome anomalies
- No specific chromosome abnormalities described, but recurrent breakpoints include 1p13, 1p36, 2p16-21, 4q25-28, 6q15-21, 7q32-34, 12q22q23, and 19p13

Chromosome Abnormalities That Cross FAB Classification Categories(40, 41, 44, 48, 51, 52 and 53, 55, 66)

del(7)(q22q36) (Fig. 32.6K)

- 9% of patients with AML; myelodysplastic syndromes or t-AML
- Poor prognostic significance
- Commercial FISH strategy is available to detect loss of D7S486 (Elastin gene)
- Molecular strategy is not applicable

Monosomy 7

- Associated with myelodysplastic syndromes and juvenile chronic myelomonocytic leukemia
- Isolated -7 has variable prognostic significance
- Commercial FISH strategy is available to detect loss of chromosome 7 centromere
- Molecular strategy is not applicable

Trisomy 8

- Isolated anomaly in 13% of AML (M2, M4 or M5); 10% of myelodysplastic syndromes; chronic myeloproliferative disorders
- Frequent secondary chromosome anomaly in Ph-positive chronic myeloid leukemia
- Intermediate prognostic significance in AML
- More aggressive as secondary chromosome abnormality
- Commercial FISH strategy is available to detect the chromosome 8 centromere
- Molecular strategy is not applicable

del(13)(q12q22) (Fig. 32.6T)

- 6% of patients with agnogenic myeloid metaplasia
- Associated with myelofibrosis in other disorders such as polycythemia vera and chronic myeloid leukemia
- Secondary anomaly in multiple myeloma, chronic lymphocytic leukemia, and non-Hodgkin's lymphoma
- Critical gene loss may be D13S319 at 13q13
- Commercial FISH strategy available to detect loss of D13S319
- Molecular strategy is not applicable

del(20)(q11q13) (Fig. 32.6Y)

- 2% to 5% of AML; usually prior myelodysplastic syndrome or myeloproliferative disorder
- 6% of patients with polycythemia vera; 1% of patients with myelodysplastic syndromes
- Abnormal erythropoiesis
- Commercial FISH strategy is available to detect loss of D20S108 at 20q12
- Molecular strategy is not applicable

REFERENCES

1. Knutsen T. *AGT International Laboratory Directory*. Lenexa, KS: Association of Genetic Technologists, 1997

2. Dewald G, Wyatt W, Silver R. Atypical BCR and ABL D-FISH patterns in chronic myeloid leukemia and their possible role in therapy. *Leuk Lymphoma* 1999; 34:481–491.

3. Kuffel D, Lindor N, Litzow M, et al. Mitomycin C chromosome stress test to identify hypersensitivity to bifunctional alkylating agents in patients with Fanconi anemia or aplastic anemia. *Mayo Clin Proc* 1997;72:579.

4. Mitelman F. Catalog of chromosome aberrations in cancer. New York: Alan R. Liss, 1991.

5. Friend S, Dryja T. Oncogenes and tumor-suppressing genes. *N Engl J Med* 1988;318:618.

6. Dewald G, Hicks G, Dines D, et al. Usefulness of culture methods to aid cytogenetic analyses for the diagnosis of malignant pleural effusions. *Mayo Clin Proc* 1982;57:488.

7. Dewald G, Allen J, Strutzenberg D, et al. A cytogenetic method for mailed-in bone marrow specimens for the study of hematologic disorders. *Lab Med* 1982;13:225.

8. Spurbeck J, Carlson R, Allen J, et al. Culturing and robotic harvesting of bone marrow, lymph nodes, peripheral blood, fibroblasts, and solid tumors with in situ techniques. *Cancer Genet Cytogenet* 1988;32:59.

9. Dewald G, Broderick D, Tom W, et al. The efficacy of direct, 24-hour culture, and mitotic synchronization methods for cytogenetic analysis of bone marrow in neoplastic hematologic disorders. *Cancer Genet Cytogenet* 1985;18:1.

10. Spurbeck J, Zinsmeister A, Meyer K, et al. Dynamics of chromosome spreading. *Am J Med Genet* 1996;61:387.

11. Dewald G, Brothman A, Butler M, et al. Pilot studies for proficiency testing using fluorescence in situ hybridization with chromosome-specific DNA probes. A College of American Pathologists/American College of Medical Genetics Program. *Arch Path* 1997;121:359.

12. Mitelman F, ed. *An International System for Human Cytogenetic Nomenclature*. Basel:S. Karger, 1995.

13. American College of Medical Genetics Laboratory Practice Committee. Metaphase fluorescence in situ hybridization (FISH). In: *Standards and Guidelines: Clinical Genetics Laboratories*, 1996:1.

14. Schad C, Dewald G. Building a new clinical test for fluorescence *in situ* hybridization. *Appl Cytogenet* 1995;21:1.

15. Dewald G, Juneau A, Schad C, et al. Cytogenetic and molecular genetic methods for diagnosis and treatment response in chronic granulocytic leukemia. *Cancer Genet Cytogenet* 1997;94:59.

16. Buno I, Wyatt W, Zinsmeister A, et al. A special fluorescent in situ hybridization technique to study peripheral blood and assess the effectiveness of interferon therapy in chronic myeloid leukemia. *Blood* 1998;92:2315.

17. Ahmann G, Jalal S, Juneau A, et al. A novel three-color, clone-specific fluorescence in situ hybridization procedure for monoclonal gammopathies. *Cancer Genet Cytogenet* 1998;101:7.

18. Dewald G, Diez-Martin J, Steffen S, et al. Hematologic disorders in 13 patients with acquired trisomy 21 and 13 individuals with Down syndrome. *Am J Med Genet Suppl* 1990;7:247.

19. Dewald G, Noel P, Dahl R, et al. Chromosome abnormalities in malignant hematologic disorders. *Mayo Clin Proc* 1985;60:675.

20. Mitelman F, Kaneko Y, Trent J. Report of the committee on chromosome changes in neoplasia. *Cytogenet Cell Genet* 1991;58:1053.

21. Mitelman F, Kaneko Y, Berger R. Report of the committee on chromosome changes in neoplasia. In: Cuticchia A, Chipperfield M, Foster P, eds. *Human Gene Mapping 1995: A Compendium*. Baltimore: Johns Hopkins University Press, 1996:1332.

22. Mitelman F, Mertens F, Johansson B. A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nat Genet* 1997;(special issue):417.

23. Bennett J, Catovsky D, Daniel M, et al. Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol* 1982;51:189.

24. Bennett J, Catovsky D, Daniel M, et al. Proposed revised criteria for the classification of acute, myeloid leukemia. A report of the French-American-British Cooperative Group. *Ann Intern Med* 1985;103:620.

25. Mathe G, Rappaport H, O'Conor G, et al. Histological and cytological typing of neoplastic diseases of haematopoietic and lymphoid tissues. In: *International Histological Classification of Tumours*. Geneva: World Health Organization, 1976:45.

26. Non-Hodgkin's lymphoma pathologic classification project. National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a working formulation for clinical usage. *Cancer* 1982;49:2112.

27. Lukes R, Collins R. Immunologic characterization of human malignant lymphomas. *Cancer* 1974;34:1488.

28. Stansfeld A, Diebold J, Kapanci Y, et al. Updated Kiel classification for lymphomas. *Lancet* 1988;i:292.

29. Pileri S, Leoncini L, Falini B. Revised European-American lymphoma classification. *Curr Opin Oncol* 1995;7:401.

30. Harris N, Jaffe E, Diebold J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: Reports of the clinical advisory committee meeting—Airlie House, Virginia, November 1997. *J Clin Oncol* 1999.

31. First International Workshop on Chromosomes in Leukaemia. Chromosomes in Ph1-positive chronic granulocytic leukaemia. *Br J Haematol* 1978;39:305.

32. Second International Workshop on Chromosomes in Leukemia, 1979. Morphological analysis of acute promyelocytic leukemia (M3) and t(8;21); cytogenetic, morphologic, and clinical correlations in acute nonlymphocytic leukemia with t(8q;21q+); chromosomes in acute promyelocytic leukemia; chromosomes in preleukemia. *Cancer Genet Cytogenet* 1980;2:89.

33. Third International Workshop on Chromosomes in Leukemia, 1980. Chromosomal abnormalities in acute lymphoblastic leukemia: structural and numerical changes in 234 cases; clinical significance of chromosomal abnormalities in acute lymphoblastic leukemia; report on essential thrombocythemia. *Cancer Genet Cytogenet* 1981;4:95.

34. Fourth International Workshop on Chromosomes in Leukemia, 1982. A prospective study of acute nonlymphocytic leukemia. *Cancer Genet Cytogenet* 1984;11:249.

35. Fifth International Workshop on Chromosomes in Leukemia-Lymphoma. Correlation of chromosome abnormalities with histologic and immunologic characteristics in non-Hodgkin's lymphoma and adult T cell leukemia-lymphoma. *Blood* 1987;70:1554.

36. Pierre R, Catovsky D, Mufti G, et al. Clinical-cytogenetic correlations in myelodysplasia (preleukemia). (Sixth International Workshop on Chromosomes on Leukemia, 1987). *Cancer Genet Cytogenet* 1989;40:149.

37. Drexler H, Borkhardt A, Janssen J. Detection of chromosomal translocations in leukemia-lymphoma cells by polymerase chain reaction. *Leuk Lymphoma* 1995;19:359.

38. Brunning R, McKenna R. Tumors of the bone marrow. In: Atlas of Tumor Pathology. Washington, D.C.: Armed Forces Institute of Pathology, 1994.

39. McCarthy K. Molecular diagnosis of lymphomas and associated diseases. *Cancer Metastasis Rev* 1997;16:109.

40. Block A. Cancer Cytogenetics. In: Keagle M, Gersen S, eds. *The Principles of Clinical Cytogenetics*. Totowa, NJ: Humana Press Inc., 1999; 345.

41. Roulston D, Le Beau M. Cytogenetic analysis of hematologic malignant diseases. In: Barch MJ, Knutsen T, Spurbeck J, eds. *The AGT Cytogenetics Laboratory Manual*, 3rd ed. Philadelphia: Lippincott-Raven Publishers, 1997:325.

42. Chessells J, Swansbury G, Reeves B, et al. Cytogenetics and prognosis in childhood lymphoblastic leukaemia: results of MRC UKALL X. *Br J Haematol* 1997;99:93.

43. Lu G, Altman A, Benn P. Review of the cytogenetic changes in acute megakaryoblastic leukemia: one disease of several? *Cancer Genet Cytogenet* 1993;67:81.

44. Barnard D, Kalousek D, Wiersma S, et al. Morphologic, immunologic, and cytogenetic classification of acute myeloid leukemia and myelodysplastic syndrome in childhood: a report from the Childrens Cancer Group. *Leukemia* 1996;10:5.

45. Harrison C, Secker-Walker L. The importance of cytogenetics and associated molecular techniques in the management of patients with leukaemia. *Clin Oncol* 1998;10:255.

46. Gilliland DG. Molecular genetics of human leukemia. *Leukemia* 1998;12:S7.

47. Johansson B, Mertens F, Mitelman F. Secondary chromosomal abnormalities in acute leukemias. *Leukemia* 1994;8:953.

48. Vallespi T, Imbert M, Mecucci C, et al. Diagnosis, classification, and cytogenetics of myelodysplastic syndromes. *Haematologica* 1998;83:258.

49. Giles F, Koeffler H. Secondary myelodysplastic syndromes and leukemias. *Current Opinion in Hematology* 1994;1:256.

50. Thirman M, Larson J. Therapy-related myeloid leukemia (Review). *Hematology-Oncology Clin North Am* 1996;10:293.

51. Fenaux P, Morel P, Lai J. Cytogenetics of Myelodysplastic syndromes. *Sem Hematol* 1996;33:127.

52. Jotterand M, Parlier V. Diagnostic and prognostic significance of cytogenetics in adult primary myelodysplastic syndromes. *Leuk Lymphoma* 1996;23:253.

53. Mecucci C. Molecular features of primary MDS with cytogenetic changes. *Leuk Res* 1998;22:293.

54. McClure R, Dewald G, Hoyer J, et al. Isolated isochromosome 17q: a distinct type of mixed myeloproliferative disorder/myelodysplastic syndrome with an aggressive clinical course. *Br J Haematol* 1999;106:445.

55. Dewald G, Wright P. Chromosome abnormalities in the myeloproliferative disorders. *Sem Oncol* 1995;22:341.

56. Fonseca R, Coignet L, Dewald G. Cytogenetic abnormalities in multiple myeloma. In: Kyle R, Gertz M, eds. *Hematology/Oncology Clinics of North America*. Philadelphia: Saunders, 1999:1169–1179.

57. Ong S, LeBeau M. Chromosomal abnormalities and molecular genetics of non-Hodgkin's lymphoma. *Sem Oncol* 1998;25:447.

58. Juliusson G, Gahrton G. Cytogenetics in CLL and related disorders. *Baillieres Clin Haematol* 1993;6:821.

59. Maljaei S, Brito-Babapulle V, Hiorns L, et al. Abnormalities of chromosomes 8, 11, 14, and X in T-prolymphocytic leukemia studied by fluorescence in situ hybridication. *Cancer Genet Cytogenet* 1998; 103:110.

60. Ott G, Katzenberger T, Greiner A, et al. The t(11;18)(q21;q21) chromosome translocation is a frequent and specific aberration in low-grade but not high-grade malignant non-Hodgkin's lymphomas of the mucosa-associated lymphoid tissue (MALT-) type. *Cancer Res* 1997;37:3944.

61. Offit K, Parsa N, Filippa D, et al. t(9;14)(p13;q32) Denotes a subset of low-grade non-Hodgkin's lymphoma with plasmacytoid differentiation. *Blood* 1992;80:2594.

62. Rosenwald A, Ott G, Pulford K, et al. t(1;2)(q21;p23) and t(2;3)(p23;q21): Two novel variant translocations of the t(2;5)(p23;q35) in anaplastic large cell lymphoma. *Blood* 1999;94:362.

63. Michaux L, Dierlamm J, Wlodarska I, et al. t(14;19)/BCL3 rearrangements in lymphoproliferative disorders: a review of 23 cases. *Cancer Genet Cytogenet* 1997;94:36.

64. Atkin N. Cytogenetics of Hodgkin's disease. *Cytogenet Cell Genet* 1998;80:23.

65. Deerberg-Wittram J, Weber-Matthiesen K, Schlegelberger B. Cytogenetics and molecular cytogenetics in Hodgkin's disease. *Annals Oncol* 1996;4(suppl 7):S49.

66. Kurtin P, Dewald G, Shields D, et al. Hematologic disorders associated with deletions of chromosome 20q: A clinicopathologic study of 107 patients. *Am J Clin Pathol* 1996;106:680.

33

The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics

James H. Ray

James German

The so-called chromosome-breakage syndromes are rare, genetically unrelated, and clinically diverse disorders that sometimes are grouped because they share two important features, excessive genomic instability and an increased risk of developing cancer. Most regularly included in the group are Bloom's syndrome (BS), Fanconi's anemia (FA), ataxia-telangiectasia (AT), the Nijmegen breakage syndrome (NBS), xeroderma pigmentosum (XP), and Werner's syndrome (WS).

Each disorder is inherited as an autosomal-recessive trait. In all the disorders except XP, the excessive genomic instability presents itself in cultured but otherwise untreated cells in the form of increased numbers of chromatid and isochromatid gaps and breaks and chromosomal rearrangements. In XP, excessive chromatid breakage occurs only after exposure of cells to ultraviolet (UV) irradiation and UV-mimetic agents. Cells from all the disorders are hyperresponsive to certain chemical or physical agents that damage DNA. Genomic instability has been demonstrated *in vivo* in at least four of the disorders. Although individuals with all the disorders have an increased risk of developing cancer, the distribution of types and sites of cancer is different in each disorder.

Remarkable progress has been made in the last decade in elucidating the molecular and biochemical bases for each of the chromosome-breakage syndromes. The chromosomal location of at least one gene responsible for each disorder has been mapped, the genes have been isolated, and the protein products have been identified. Today, the roles that the proteins play in cellular metabolism are being elucidated, in several cases bringing to light cellular mechanisms not previously known to exist in man.

Here, the clinical features, the cytogenetics, and the molecular genetics of each of the chromosome-breakage syndromes are summarized.

BLOOM'S SYNDROME

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

BS is the rarest of the chromosome-breakage syndromes. Fewer than 200 affected individuals are known to have been diagnosed since 1954 when the disorder originally was described (1). The incidence of BS in the United States is estimated to be 1/6,331,000. In Ashkenazi Jews it is more common, the incidence having been projected to be 1/60,000; 1 in 107 Ashkenazi Jews in New York City is heterozygous for a BS-associated mutation. Extensive clinical and laboratory records on individuals with BS are maintained in the Bloom's Syndrome Registry, which has made possible an accurate definition of the clinical disorder as well as its complications (2).

Bloom initially described three patients with stunted growth and sun-sensitive telangiectatic erythema of the face (1). Additional features are listed in Table 33.1. The BS fetus is abnormally small, the mean birth weight being 1,906 g for males and 1,810 g for females. Although the postnatal growth rate is normal, both height and weight remain below the third percentile throughout childhood. The vast majority of adults are abnormally short. Men with BS attain an average height of 150 cm; women average 145 cm. Body proportions are normal except for mild microcephaly, dolicocephaly, and a characteristic facies, the result of malar hypoplasia, a relatively prominent nose and ears, and a relatively small mandible. The voice characteristically is high-pitched and squeaky.

Facial erythema typically appears during the first or second summer of life following exposure to the sun. The cheeks and nose, and occasionally the dorsa of the hands and forearms, are affected. The severity of the facial lesion varies greatly and generally decreases with age. Well-circumscribed regions of hypo- and hyperpigmentation — café-au-lait spots — ranging in size from a few millimeters to many centimeters also almost always are observed, and these, in contrast to the limitation of the erythematous and sun-sensitive lesion on the face, are found on the trunk and extremities.

Individuals with BS are immunologically compromised. Most have recurring respiratory-tract infections, specifically otitis media and pneumonia. Abnormally low concentrations of one or more circulating immunoglobulins are observed in 87% of patients. Approximately 12% of the 169 persons followed in the Registry have developed noninsulin-dependent diabetes mellitus. Vomiting and gastrointestinal infections commonly occur during infancy, the former being a result of, in at least some

cases, severe gastroesophageal reflux. Menstruation begins at the usual age in girls with BS but ceases prematurely. Although infertility is the rule, three women have become pregnant and have had normally developed children. Men with BS invariably fail to produce spermatozoa. The average age of individuals with BS at the time of death is 23.6 years, with cancer being the primary cause and chronic lung disease being the cause in a few.

Over 100 cancers have been diagnosed in the 169 individuals under surveillance in the Registry (3). The mean age at the time of cancer diagnosis is 24.4 years. Twenty-nine individuals have developed multiple primary cancers. The cancers in BS are of similar cellular types and arise at similar sites as in the general population; however, in addition to being more common, they arise decades earlier than normal. Included among the cancers that have been diagnosed in BS are acute leukemias, lymphomas, and carcinomas, the latter having arisen most often in the gastrointestinal tract but also in the uterus, breast, larynx, liver, and skin. The rare tumors of childhood and germ-cell tumors also have been observed. Leukemias occur predominately during childhood. Carcinomas appear in late adolescence and predominate thereafter. Cancer has been the cause of death of fully one third of persons with BS.

Cytogenetics

That increased spontaneous chromosome breakage is a feature of BS cells was reported in the mid-1960s (4, 5). Chromosome breakage has been observed in blood lymphocytes and bone marrow cells in short-term culture and in dermal fibroblasts and Epstein-Barr virus-transformed lymphoblastoid cell lines (LCLs) in long-term culture. Chromosome instability manifests itself in cultured but otherwise untreated BS cells in the form of increased numbers of chromatid and isochromatid gaps and breaks, acentric fragments, dicentric and otherwise structurally rearranged chromosomes, telomere associations, and quadriradial configurations (Qrs).

The most characteristic cytogenetic abnormality in untreated cells is the Qr (Fig. 33.1). Qrs in BS typically are symmetrical, consisting of homologous chromosomes associated in a four-armed configuration in which the centromeres are positioned in opposite arms and the opposite arms are equal in length. Qrs result

TABLE 33.1. FEATURES OF BLOOM'S SYNDROME

Clinical
Pre- and postnatal growth deficiency, with relatively normal body proportions
Telangiectatic erythema of face, sun-sensitive
Microcephaly with dolichocephaly
Malar and mandibular hypoplasia
Relative prominence of nose and ears
High-pitched squeaky voice
Hypo- and hyperpigmented areas of skin
Immunodeficiency
Increased otitis media and pneumonia
Bronchiectasia and pulmonary failure
Male infertility; female subfertility and premature menopause
Diabetes mellitus, usually adult onset
Increased incidence of neoplasia of many types and sites
Cytogenetic
Excessive spontaneous chromosomal breakage
Homologous quadriradial configurations
Elevated sister-chromatid exchange frequency
Telomere association

from the exchange of chromatid segments between homologous chromosomes at what appear to be homologous sites; therefore, they are cytological evidence of somatic recombination — somatic crossing-over. The discovery of Qrs in BS was of general biological significance because the exchange of segments between homologous chromosomes may, depending on the segregation pattern of the recombinant chromosomes in subsequent mitoses, lead to the loss of heterozygosity thereby unmasking deleterious recessive genes. Thus, crossing-over is a mechanism by which expression of an altered cellular phenotype can arise in somatic cells.

Increased somatic recombination also has been demonstrated *in vivo* in BS. Multiple chromosome satellite distribution patterns were observed on the group D and G chromosomes in T lymphocytes from an individual who had satellite polymorphisms. A significant percentage of the lymphocytes were homozygous for satellite polymorphisms that constitutionally were present in the heterozygous state in the individual. Also indicative of excessive somatic crossing-over *in vivo* is the finding of an increased incidence of MM and NN erythrocytes in the circulating blood of persons who constitutionally are blood type MN.

In the mid-1970s, a technique was developed that permitted intrachromosomal recombination to be observed (in contrast to the interchromosomal recombination signified by the Qr just discussed). Cells are harvested after growth for two cell cycles in the presence of the thymidine analogue bromodeoxyuridine (BrdU). Slide preparations made from the cultures are exposed to ultraviolet (UV) light and then stained with quinacrine mustard or Giemsa depending on whether fluorescence or light microscopy, respectively, is to be used for viewing the chromosomes. Following treatment with the aforementioned protocol, the chromatids of some chromosomes are differentially stained uniformly along their lengths, one chromatid exhibiting dull fluorescence or pale staining and the other chromatid exhibiting bright fluorescence or dark staining depending on the method of microscopy employed. The sister chromatids of other chromosomes exhibit alternating pale and dark stained regions along their lengths, with the staining patterns of the sister chromatids being mirror images of one another. The points on these latter chromatids at which the staining patterns change represent the points at which exchanges of DNA segments occurred between the sister chromatids of the chromosomes during DNA replication.

When blood lymphocytes from normal (nonBS) individuals are examined using the sister-chromatid exchange (SCE) technique just described, fewer than 10 SCEs/cell are observed (Fig. 33.2A). BS blood lymphocytes, on the other hand, exhibit a 10- to 15-fold increase in the average number of SCEs/cell (Fig. 33.2B) (7). BS is the only human genetic disorder in which BrdU-substituted, but otherwise untreated cells, are known to exhibit an increased frequency of SCE.

As blood lymphocytes from increasing numbers of persons with BS were examined using the SCE technique, an unexpected finding was made, one that was to lead to the discovery of two genetic mechanisms previously not known to be employed in mammalian cells. Some individuals with BS exhibited two populations of cells, one with the high-SCE frequency that is uniquely characteristic of BS and the other with the low-SCE frequency that is observed in normal cells (8). A number of hypotheses were proposed over the years to account for the enigmatic presence of cells with a nonBS (normal) phenotype in persons with a recessive disorder. It was not until the molecular basis for the disorder was identified (see below) that two different genetic mechanisms capable of producing low-SCE somatic cells in persons with BS were discovered: (i) recombination between the two chromosomes 15 within the BLM gene in individuals who were compound heterozygotes (i.e., individuals having mutations at different sites in *BLM*) and (ii) back mutation at *BLM*.

High-SCE BS cells are hyperresponsive with regard to survival and SCE induction following exposure to certain DNA-damaging agents. Chemicals that have been studied include mitomycin C (MMC), monofunctional mitomycin C, 4-nitroquinoline-1-oxide (4NQO), and ethyl methanesulfonate (EMS). BS cells generally only are able to survive exposure to concentrations of chemicals that are an order of magnitude less than normal cells. The number of SCEs induced by a given concentration of a chemical is significantly greater in the already high-SCE BS cells than in normal cells. Survival and SCE induction in the exceptional low-SCE cells in some persons with BS, mentioned above, are comparable to that in normal cells.

BS obligate heterozygotes, who are normally developed and healthy, exhibit normal levels of chromosome breakage and SCE. Some heterozygous men do produce abnormally great numbers of sperm with multiple chromosomal breaks and rearrangements.

Genomic instability in BS also has been demonstrated in several other ways. BS cells are hypermutable: an increased incidence of diphtheria-toxin resistant and hypoxanthine guaninephosphoribosyl transferase (HPRT)-deficient cells have been detected in cultured BS cells. The incidence of HPRT-deficient lymphocytes also is elevated in the circulating blood of persons with BS. In addition, the frequency of erythrocytes mutant at the glycophorin A locus on chromosome 4 is dramatically elevated above normal in BS; increased numbers of M/null and N/null cells and, as mentioned above, increased numbers of MM and NN cells are observed in persons who constitutionally are MN, demonstrating that crossing-over occurs excessively in erythrocyte stem-cell precursors. Increased somatic recombination also has been documented for the HLA-A locus. Increased somatic recombination *in vitro* has been demonstrated by study of polymorphic microsatellite markers in subclones of BS LCLs originally derived from a single EBV-transformed B lymphocyte (i.e., subclones of a clone). Although the microsatellite markers were identical in all cells of a given LCL when it was established by the nature of its origin, the markers that originally existed in the heterozygous state were converted to homozygosity over time as the result of somatic recombination between homologous chromosomes. The frequency of cells with micronuclei also is significantly increased in exfoliated oral mucosal and bladder cells, demonstrating that excessive chromosome breakage occurs *in vivo* as well as *in vitro*.

Genetic Heterogeneity and Complementation

Somatic cell hybridization utilizing cells lines derived from persons with BS representing four distinctive ethnic groups were used to search for evidence of genetic heterogeneity in BS (9).

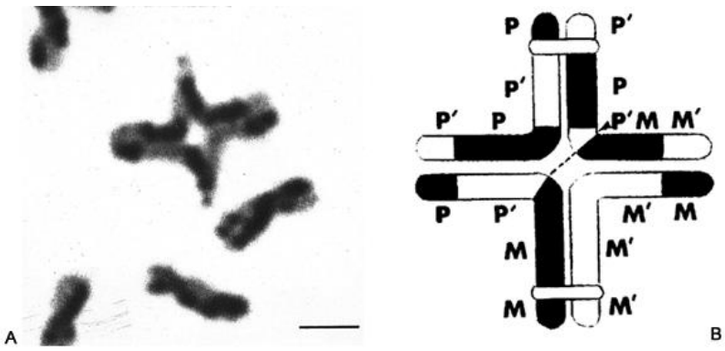


FIGURE 33.1. A: Qr configuration formed between two apparently homologous chromosomes in a BS T lymphocyte. Chromosomes are differentially stained as result of growth for two cell cycles in BrdU-containing medium. **B:** Diagramatic representation of event that led to Qr formation in (**A**). Chromatids arbitrarily have been designated P and P′ for the paternally derived chromosome and M and M′ for the maternally derived chromosome. The presence of P′ segments on P chromatids and P segments on P′ chromatids indicates that exchanges have taken place between the sister chromatids of the paternal chromosome. This also is true for M′ and M chromatids of maternal origin. The presence of M/M′ segments on P/P′ chromatids and P/P′ segments on M/M′ chromatids signifies the exchange of DNA segments between nonsister chromatids of the homologous chromosomes that led to the formation of the Qr. The point at which the nonsister chromatid exchange occurred is designated by the dashed line. Bar represents 2 µm. (From Ray JH, German J. The chromosome changes in Bloom's syndrome, ataxia-telangiectasia, and Fanconi's anemia. In: Arrighi FE, Rao PN, Stubblefield E, eds. *Genes, chromosomes, and neoplasia*. New York, Raven Press, 1981:351–378.)

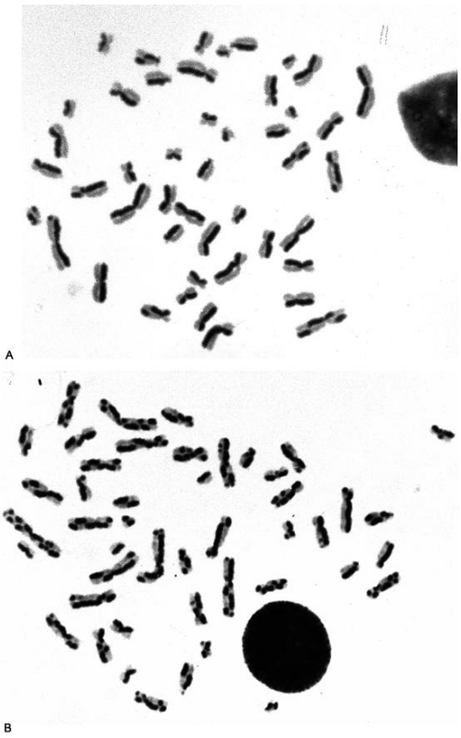


FIGURE 33.2. Differentially stained chromosomes in (A) normal and (B) BS PHA-stimulated lymphocytes demonstrating the dramatically increased SCE frequency of BS cells. Bar represents 10 µm. (From Ray JH, German J. The chromosome changes in Bloom's syndrome, ataxia-telangiectasia, and Fanconi's anemia. In: Arrighi FE, Rao PN, Stubblefield E, eds. Genes, chromosomes, and neoplasia. New York, Raven Press, 1981:351–378.)

The SCE frequency remained elevated in all hybrid cell lines, indicating that mutation at a single locus was responsible for the elevated SCE frequency in all the parental cell lines that were studied. Thus far, molecular genetic studies support the existence of a single locus for BS.

Molecular Genetics

Results of microcell-mediated chromosome transfer experiments, in which normal human chromosomes tagged with the dominant selectable marker for neomycin resistance were introduced into BS cells, provided the first indication that the BS locus was on chromosome 15 (10). Use of structurally rearranged chromosomes further localized the BS locus to the distal part of 15q. That the BS locus was present on chromosome 15 was confirmed by study of an individual with BS who presented with features of both the Prader-Willi syndrome and BS and had maternal uniparental disomy for chromosome 15. The mother was heterozygous for mutation at the BS locus, and her child was homozygous by virtue of having inherited two copies of the segment of her chromosome 15 that carried the *BLM* mutation.

Subsequent to the assignment of *BLM* to chromosome 15, the gene's location was sublocalized by homozygosity mapping. Genotyping of polymorphic DNA markers in persons with BS born to consanguineous parents, taking advantage of the likelihood that the offspring of such unions would have identical mutations at the two BS loci and that DNA sequences closely linked to BLM also would be identical as result of the descent of both chromosomes from a single ancestor, demonstrated tight linkage of *BLM* to the protooncogene *FES*, which was known to map to band 15q26.1 (11). The precise map position of BLM was determined by taking advantage of the low-SCE populations of cells that exist in persons with BS who are the products of nonconsanguineous unions. Previous analysis of polymorphic DNA markers in low-SCE cells from multiple individuals with BS had shown that markers distal to *BLM* were homozygous whereas markers proximal to *BLM* were heterozygous, demonstrating that these individuals were compound heterozygotes, having different mutations at the *BLM* alleles. Intragenic recombination in such individuals can give rise to a wild-type BLM allele, effectively making the cells heterozygous for BS (Fig. 33.3). By comparing polymorphic markers in low-SCE and high-SCE cells from the same individual and determining the precise point at which recombination had occurred giving rise to the low-SCE cells, as indicated by the fact that polymorphic markers changed from exhibiting heterozygosity to exhibiting homozygosity, the map position of *BLM* was identified, and a candidate cDNA was isolated and examined for the presence of mutations (13).

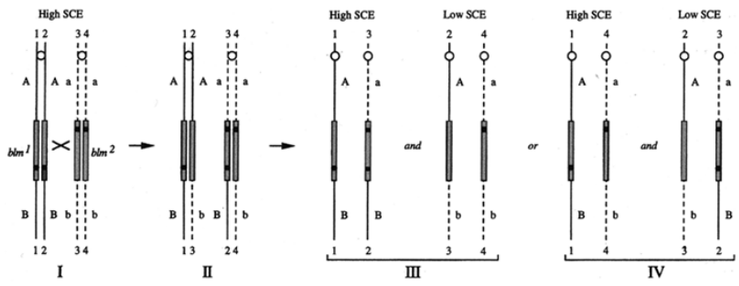


FIGURE 33.3. Model demonstrating how a wild-type BLM allele and low-SCE frequency cells can be generated in an individual with BS by intragenic recombination. **I.** The two chromosomes 15 of a G_2 somatic cell in a BS compound heterozygote are depicted. The mutations in *BLM*, blm^1 and blm^2, are designated by the black dots at different sites in the hatched rectangle. The sister chromatids of the chromosome with the blm^1 mutation are designated 1/1 and 2/2; those of the chromosome with the blm^2 mutation are designated 3/3 and 4/4. The individual is heterozygous for flanking markers designated A/a and B/b. **II**. The G_2 somatic cell in **I** after exchange of segments between nonsister chromatids 2/2 and 3/3. The site of exchange was within *BLM* at a point between blm^1 and blm^2. The resulting BLM locus on chromatid 2/3 (left chromosome 15) is wild type whereas that on chromatid 3/2 (right chromosome 15) has two mutations (blm^1 and blm^2). Flanking markers distal to the point of interchange also are different, being b for chromatid 2/3 and B for chromatid 3/2. Proximal markers remain the same. **III** and **IV**. Alternative chromatid segregation patterns in the subsequent mitosis are depicted. In **III** chromatids 1/1 and 3/2 co-segregate to one daughter cell and 2/3 and 4/4 cosegregate to the other. The daughter cells are homozygous (B/B or b/b) for markers distal to *BLM*. In **IV** chromatids 1/1 and 4/4 cosegregate as do chromatids 2/3 and 3/2. The daughter cells remain heterozygous (B/b) for markers distal to *BLM*. Daughter cells to which chromatids 1/1 and 3/2 or 1/1 and 4/4 cosegregate will maintain the high-SCE frequency characteristic of BS whereas those to which chromatids 2/3 and 4/4 or 2/3 and 3/2 cosegregate will exhibit a low-SCE frequency by virtue of the fact that the BS allele on one chromosome 15 is wild-type. (From Ellis NA, Lennon DJ, Proytcheva M, et al. Somatic intragenic recombination within the muted locus *BLM* can correct the high sister-chromatid exchange phenotype of Bloom syndrome cells. *Am J Hum Genet* 1995:57:1019–1027.)

BLM encodes a 1417 amnio-acid polypeptide with a predicted molecular mass of 159 kDa. The BLM protein exhibits homology to the RecQ helicases, a subfamily of DexH box-containing DNA/RNA helicases. The RecQ helicases are part of a larger group of proteins that contain seven motifs that constitute a highly conserved helicase domain. BLM functions as all 3′ -> 5′ DNA helicases *in vitro*.

Multiple unique BLM mutations have been identified in different individuals with BS. Individuals of Ashkenazi Jewish ancestry share a single mutation, a 6 bp deletion and 7 bp insertion at nucleotide 2,281 BLM, proving that founder effect is responsible for the elevated frequency of BS in that population. More than 60 mutations have been observed in BS individuals of non-Ashkenazi ancestry. Most of the mutations lead to the production of truncated proteins although missense mutations also have been identified.

In spermatogenesis, the BLM protein moves onto the meiotic bivalent during zygotene and pachytene, pointing to a role also in germline recombination.

Prenatal Diagnosis

Determination of the SCE frequency of amniocytes or chorionic villus cells can be offered to couples when both partners are known to be obligate heterozygotes for a BS-producing mutation by virtue of having had an affected child. Use of this technique has permitted the diagnosis BS to be ruled out in a number of pregnancies. It is surprising that only two BS fetuses have been identified among the approximately 20 that have been studied. With the molecular identification of *BLM* and the finding that a single mutation is responsible for BS in individuals of Ashkenazi Jewish ancestry, it has become possible to offer carrier testing and prenatal diagnosis using molecular genetic methods to couples when both partners belong to that ethnic group; such testing is commercially available. Molecular methods also can be employed for couples of non-Ashkenazi ancestry to whom a child with BS has been born when the mutation(s) in *BLM* that they carry has been defined at the molecular level.

FANCONI'S ANEMIA

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

Fanconi's anemia (FA) occurs with a frequency of 1–3/1,000,000 individuals; approximately 1 in 200 – 300 is heterozygous for mutation at one of the FA loci. Clinical and laboratory records on affected individuals are maintained in the International Fanconi Anemia Registry (14).

The disorder first was described in three sibs with pancytopenia and multiple anatomical defects (15, 16). The phenotype is highly variable, with both intrafamilial and interfamilial differences commonly being observed as well as differences between monozygotic twins. Some individuals exhibit severe anatomical malformations whereas others have none (Table 33.2).

TABLE 33.2. FEATURES OF FANCONI'S ANEMIA

Clinical
Moderate pre- and postnatal growth deficiency
Hyperpigmentation of the skin
Anatomical defects
Absence/hypoplasia of radii and thumbs
Kidney malformation
Strabismus
Micropathalmia
Hypogonadism
Bone marrow hypoplasia
Pancytopenia
Increased neoplasia
Benign and malignant liver tumors
Acute nonlymphocytic leukemia
Cytogenetic
Increased spontaneous chromosomal breakage
Triradial and nonhomologous quadriradial configurations
Hyperresponsiveness to DNA-crosslinking agents

The individual with FA exhibits moderate prenatal and postnatal growth deficiency. Height, weight, and head circumference consistently fall below the fifth percentile. Bruises begin to appear at 5 to 10 years of age, and anemia is diagnosed. In the absence of treatment, the pancytopenia progresses during subsequent years to bone marrow failure. This limits the life expectancy of persons, with 81% dying by 40 years of age.

Hyperpigmentation of various parts of the skin commonly are observed. Approximately 50% of affected individuals have skeletal abnormalities affecting the radial side of the forearms, typically unilateral or bilateral absence or hypoplasia of the radii or thumbs. Kidneys may be missing, misplaced, or horseshoe-shaped, and anomalies of the intestinal tract, central nervous system, and heart sometimes are present. Eye abnormalities include strabismus and micropthalmia. A significant percentage of affected individuals exhibit hypogonadism. Cryptorchidism is observed in approximately 30% of males; spermatogenesis is abnormal, and fertility is significantly reduced. Females, while fertile, may experience menstrual irregularities, secondary amenorrhea, and premature menopause. Hypothyroidism, requiring hormone treatment, is present in some individuals with FA.

Strict adherence to the diagnostic criteria established by Fanconi has led to serious underdiagnosis of FA. As many as one third of affected individuals exhibit only minor phenotypic abnormalities, including height, weight, or head circumference below the fifth percentile, skin pigmentation abnormalities, and/or microphthalmia; therefore, individuals in whom these minor abnormalities are identified should receive appropriate clastogen testing (see below) to determine the appropriateness of FA as a clinical diagnosis.

Treatment of FA consists mainly of blood transfusions and maintenance of bone marrow function by administration of anabolic steroids. While such treatment often is beneficial, bone marrow transplantation now has become the therapy of choice for prolonging the lives of individuals with FA. Marrow from a histocompatible sib and reasonably well-matched placental blood from an unrelated person have served as suitable sources of hematopoietic stem cells.

Early experience with bone marrow transplantation demonstrated that individuals with FA were hypersensitive to the high-dose cyclophosphamide pretreatment regimen and were susceptible

to post-transplant graft-versus-host disease. Modification of the pretreatment regimen to include lower doses of cyclophosphamide followed by thoraco-abdominal irradiation with a moderate dose of 1 – 5 Gy has resulted in improvement of long-term posttransplant survival. Cyclosporin A also now is administered to prevent or lessen the severity of the graft-versus-host disease.

A complication of FA is cancer, which arises in at least 20% of affected individuals. Curiously, the association of FA with predisposition to cancer was not recognized until the mid-1960s. One possible explanation for the delay in recognizing FA as a cancer-associated disease is the fact that during the first several decades after the condition was described affected individuals died early in life from bone marrow failure. In the late 1950s long-term treatment with anabolic steroids began to prolong life. The treatment regimen and/or the aging of the FA population may have contributed to the appearance of cancer. The cancers that do arise, predominantly acute nonlymphocytic leukemia and hepatocellular neoplasms, clearly are completely different in type and distribution from that in the general population.

Obligate FA heterozygotes are not known to be cancer-prone.

Cytogenetics

Chromosome instability was reported to be a feature of FA in the mid-1960s. Schroeder et al. (17) reported that phytohemagglutinin (PHA)-stimulated lymphocyte cultures from affected brothers contained an increased frequency of cells with chromosomal abnormalities. Dermal fibroblast cultures and LCLs also exhibit increased spontaneous chromosome breakage.

The chromosomal aberration frequency in short-term lymphocyte cultures from a given individual with FA may vary over time; the number of cells with chromosomal aberrrations varied from 5% to 42% in a 3-year study of one individual and from 11% to 52% in a 6-year study of another. The frequency of chromosomal aberrations is affected by the concentration of oxygen in which the cells are cultured. Reduction of the oxygen concentration from 20% to 5% significantly reduces the level of chromosome breakage. Further suggestion of a role for active oxygen species in the cellular defect in FA is the finding that antioxidants and the iron chelator desferrioxamine reduce the frequency of spontaneous chromosomal breaks.

The chromosomal abnormalities most commonly observed include chromatid and isochromatid gaps and breaks, acentric fragments, rearranged monocentric chromosomes, and Qrs. Qrs in FA differ from those in BS in that nonhomologous rather than homologous chromosomes typically are involved (Fig. 33.4); therefore, the interchanges in FA would be expected to give rise to cells with partial monosomy, partial trisomy, or balanced translocations depending on the segregation patterns of chromatids in subsequent cell cycles. An increased frequency of cells with endoreduplicated chromosomes also has been observed in cultured cells.

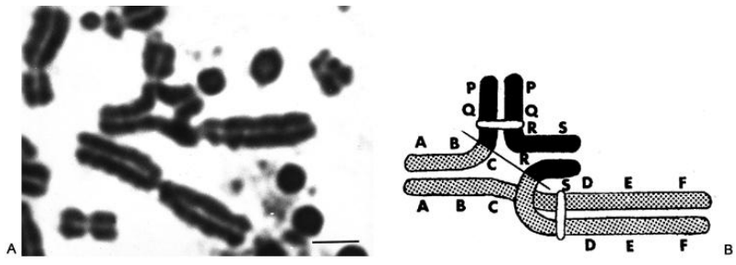


FIGURE 33.4. A: Nonhomologous Qr in FA PHA-stimulated lymphocyte. **B:** Diagramatic representation of event that led to Qr formation in (**A**). Segments on one chromosome are designated ABCDEF; those on the other are designated PQRS. The nonsister chromatid exchange occurred between segments CD on one chromosome and segments QR on the other. The exact point of exchange is denoted by the solid line. Depending on the chromatid segregation pattern in the subsequent mitosis the karyotype of the resulting cells can be either normal, balanced, or unbalanced. For example, if the chromatids containing segments PQCBA and SRDEF cosegregate to one daughter cell and the chromatids containing segments PQRS and ABCDEF cosegregate to the other, the karyotype of the former cell will be balanced and the latter will be normal; however, daughter cells with unbalanced karyotypes will result if the chromatids containing segments PQCAB and ABCDEF cosegregate to one cell and chromatids containing segments PQRS and SRDEF to the other. (From Ray JH, German J. The chromosome changes in Bloom's syndrome, ataxia-telangiectasia, and Fanconi's anemia. In: Arrighi FE, Rao PN, Stubblefield E, eds. *Genes, chromosomes, and neoplasia*. New York, Raven Press, 1981:351–378.)

In addition to displaying an increased frequency of spontaneous chromosomal abnormalities, FA cells exhibit hyperresponsiveness to DNA-crosslinking agents. Exposure to concentrations of DNA-crosslinking agents that have little or no effect on normal cells drastically increases the chromosomal aberration frequency. Agents to which FA cells exhibit hyperresponsiveness include nitrogen mustard (HN2), MMC, long wave UV light in

the presence of 8-methoxypsoralen, cyclophosphamide, cis-diamminedichloroplatinum (II) and (IV), and diepoxybutane (DEB). FA cells respond normally to monofunctional alkylating agents.

The demonstrated hyperresponsiveness to the DNA-crosslinking agents MMC and DEB led to the development of diagnostic tests for individuals with suggestive clinical phenotypes. Soon thereafter, their utility for the prenatal diagnosis of FA in amniocytes and chorionic villus cells was demonstrated (18).

Approximately 15% of persons with FA apparently have two populations of lymphocytes — one that is characteristically hyperresponsive following exposure to DEB and one that responds normally. The finding of two populations of cells with regard to DEB sensitivity is reminiscent of the high-SCE and low-SCE populations of cells that coexist in some individuals with BS who are compound heterozygotes. At least one mechanism that has been shown to account for the existence of cells exhibiting a “wild-type” phenotype in FA is the occurrence of a second mutation in one of the mutant FA alleles that negates the effect of the original mutation on truncation of the encoded protein. The resulting protein, while not being normal in the strictest sense, is able to correct the hyperresponsiveness of FA cells to DNA-crosslinking agents.

In support of *in vivo* genomic instability in FA is the presence of an increased frequency of micronuclei in bone marrow and oral mucosal cells and the presence of an increased incidence of blood lymphocytes mutant at the *HPRT* locus. In addition, the frequency of erythrocytes exhibiting evidence of recombination at the glycophorin A locus is significantly elevated. The ease with which clones of cells with translocations are detected in the bone marrow cells also points to the existence of *in vivo* chromosome instability.

Genetic Heterogeneity and Complementation

The clinical variability observed in FA may be related to the genetic heterogeneity that has been identified (19, 20 and 21). The existence of multiple complementation groups in FA was demonstrated by somatic cell hybridization using LCLs derived from unrelated affected individuals. Use of an FA LCL having two selectable markers, ouabain resistance and HPRT deficiency, permitted selection of complemented somatic cell hybrids based on their abilities to grow in medium containing ouabain, hypoxanthine, methotrexate, and thymidine. In the initial study, two hybrid cell lines exhibited a normal incidence of chromosome breakage and a normal growth response following incubation in medium containing MMC whereas the remaining hybrid cell line exhibited hypersensitivity to MMC typical of FA, establishing the existence of at least two complementation groups — groups A and non-A. In subsequent hybridization studies, the number of FA complementation groups was expanded to eight. Thus, FA is as genetically heterogeneous as XP (see below); however, as has been shown to be the case in XP, most individuals with FA belong to just two or three complementation groups; approximately 67% are FAA; 15% are FAC.

Molecular Genetics

The genes mutated in four of the FA complementation groups have been mapped to four different chromosomal regions, confirming the genetic heterogeneity of the disorder: *FAA* to band 16q24.3; FAC to band 9q22.3; *FAD* to band 3p22 -3p26; *FAG* to band 9p13. The genes for complementation groups A, C, and G have been cloned.

FAC was cloned using functional complementation by electroporated cDNAs of the MMC- and DEB-sensitivity of cultured cells (22). *FAC* spans approximately 80 kb of DNA and consists of 14 exons that encode a 558 amino-acid polypeptide with a 63 kDa molecular mass. The relatively few mutations that have been identified in *FAC* generally lead to truncation or internal deletion of the FAC protein. Although an FAC protein may be produced in cells homozygous for the mutation, the abnormal protein fails to correct the hyperresponsiveness of the cells to MMC, showing that intact carboxy- and amino-terminal domains are required for normal protein function.

The two most common mutations in *FAC* are delG322 in exon 1 and IVS4+4 A to T in intron 4. Additional mutations have been detected in exons 1, 6, and 14. Cells with delG322 and IVS4+4 mutations produce 50 kDa and 55 kDa isoforms of the FAC protein, respectively. The mutation delG322 is found in persons of Northern European ancestry. IVS4+4 occurs predominantly in individuals of Ashkenazi Jewish ancestry; IVS4+4 carrier frequency in the Ashkenazi is estimated to be one in 89. The existence of a founder mutation in this major subgroup of Jewry makes possible complementation group C carrier and prenatal screening.

Data from the Registry suggest that individuals homozygous for delG322 have a relatively mild form of the disorder, with few congenital anomalies and a relatively delayed onset of hematological disease, whereas those homozygous for IVS4+4 have a more severe form, with multiple congenital anomalies and an early onset of hematological complications. Individuals with delG322 mutations tend to survive longer than those with IVS4+4 mutations (14.9 years versus 9.7 years, respectively). Quite surprisingly, the degree of DEB hyperresponsiveness is inversely related to the severity of the disease.

FAC protein is expressed in all cells; it corrects the hyperresponsiveness of complementation group C cells following exposure to bifunctional alkylating agents. FAC is found in the cytoplasm of cells at all stages of the cell cycle regardless of whether the cells have been exposed to bifunctional alkylating agents; this observation makes it unlikely that the protein acts directly on the DNA crosslinks induced by these agents. The protein binds to a complex of cytoplasmic proteins with molecular masses of 60, 50, and 35 kDa. It also binds to cyclin-dependent kinase cdc2, suggesting that is plays a role in G_2/M cell-cycle transition, and to GRP94, a protein expressed in the endoplasmic reticulum with a function in protein transport.

Much less is known about FAA, the protein defective in individuals belonging to complementation group A (23). *FAA* spans 80 kb of DNA and consists of 43 exons that encode a 1,455 amino acid polypeptide with a molecular mass of 163 kDa. The FAA protein contains a nuclear localization signal and a leucine zipper motif, suggesting a nuclear localization; however, early studies indicate that it, like FAC, is located in the cytoplasm.

FAA shares no significant homology with other known proteins. Multiple mutations have been identified in FAA. The mutations generally result in cDNA deletions that lead to the production of truncated proteins.

A candidate cDNA for FAG was identified by its ability to eliminate the MMC sensitivity of complementation group G lymphoblasts (24). *FAG* is identical to *XRCC9*, a human gene that complements the MMC sensitivity of Chinese hamster mutant cell line UV40. Neither FAG nor XRCC9 shares homology with any known protein. The protein product has been suggested to function in DNA postreplication repair or cell-cycle checkpoint regulation.

Prenatal Diagnosis

Prenatal diagnosis has been performed on a large number of pregnancies at risk for FA (25). FA amniocytes and chorionic villus cells exhibit hyperresponsiveness to the chromosome breaking action of DEB just as FA lymphocytes do. Prenatal diagnosis of FA is made, therefore, by demonstrating the characteristic increase in chromatid aberrations and interchanges following exposure of fetal cells to DEB.

The recent isolation of the genes that are defective in FAA and FAC permits molecular prenatal diagnosis to be performed in selected situations. FAC carrier detection and prenatal screening currently are commercially available for individuals of Ashkenazi Jewish ancestry. Molecular prenatal diagnosis of FA in other ethnic groups unfortunately will be limited to families ascertained through the birth of an affected child and in which the mutation(s) has been identified.

ATAXIA-TELANGIECTASIA

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

Ataxia-telangiectasia (AT) was described independently by Syllaba and Henner (26) and Louis-Bar (27). It occurs with a frequency of 1/20,000 to 1/40,000; approximately 1% of the population is predicted to be heterozygous for an AT-causing mutation.

The individual with AT appears normal at birth; however, neurological signs and symptoms signaling the presence of the disorder appear early in life (Table 33.3). The first is an unsteady gait, which appears between 1 to 3 years of age at the time the child begins to walk. Cerebellar ataxia then progresses during early childhood, and the affected person becomes unable to walk and usually is confined to a wheelchair or the bed by age 10. Speech progressively becomes slurred, and ocular apraxia develops so that the individual has difficulty following objects moved across the visual field. Choreoathetosis is seen in most individuals. Telangiectasia, another clinical sign of AT, also becomes apparent during the first decade of life. Dilated capillaries characteristically are present on the bulbar conjunctiva, the eyelids, the rims of the ears, and the antecubital areas of the arms.

TABLE 33.3. FEATURES OF ATAXIA-TELANGIECTASIA

Clinical
Neurological abnormalities
Cerebellar ataxia
Telangiectasia, bulbar conjunctiva and skin
Ocular apraxia
Choreoathetosis
Slurred speech
Immunodeficiency
Hypoplastic or missing thymus
Hypogonadism (ovarian dysgenesis)
Elevated serum alpha-fetoprotein
Hypersensitivity to ionizing radiotherapy
Increased neoplasia, especially lymphoid
Cytogenetic
Increased spontaneous chromosome breakage
Telomere association
Clones of T lymphocytes with rearranged chromosomes 7 and 14
Hypersensitivity to ionizing radiation and radiomimetic agents

The thymus is hypoplastic or missing in the majority of cases. Immunodeficiency is a common finding but sometimes is not manifested clinically for a number of years. Serum levels of IgA, IgE, and IgG2 are decreased characteristically while IgM, IgG1, and IgG3 concentrations generally fall within normal ranges. Blood T lymphocytes respond poorly to PHA stimulation *in vitro*. The absolute numbers and proportions of circulating T cells commonly are decreased; for example, CD4+/CD45RA+ and naïve T cells sometimes are decreased. Alpha-fetoprotein (AFP), a serum protein that normally decreases rapidly following birth, remains elevated throughout life in most persons with AT. Hypogonadism, and in females even ovarian dysgenesis, is observed relatively commonly.

In one retrospective study the median age at which AT was diagnosed was 78 months. Gait abnormalities had appeared at 15 months of age on average, but the correct diagnosis typically had not been made until after the appearance of telangiectasia at 72 months. Because of the delay in determining that a child has AT, a couple may initiate another pregnancy before it is realized that the first child is so-affected. To permit earlier diagnosis diagnosis of AT, serum AFP determinations would be useful in a child with unexplained and progressive ataxia. The ataxic child with an increased serum AFP concentration (>10 nanograms/mL) then could have a cytogenetic analysis either of T lymphocytes, determining the level of spontaneous chromosome breakage and looking specifically for clones of cells with rearrangements of chromosomes 7 and 14 (see below), or of cultured fibroblasts, examining them for radiosensitivity. The early diagnosis of AT could be beneficial in the medical management of an affected child and also could permit couples to make informed family planning decisions.

An increased risk of developing lymphoid malignancy is an important feature of AT. 10% to 15% of affected individuals develop cancer by early adulthood. Approximately 85% of the malignancies are leukemias or lymphomas (28). T-cell neoplasia may be diagnosed at any age, e.g., T-cell acute lymphocytic leukemia, T-cell lymphoma, or T-cell prolymphocytic leukemia. The increased predisposition to leukemia and lymphoma in AT has been hypothesized to be related to the increased frequency of chromosomal rearrangements that affect the loci that encode T-cell receptors and the immunoglobulin heavy chain (see below). Solid tumors, including hepatocellular carcinoma, ovarian dysgerminomas,

and breast and gastric carcinoma, also have been observed in older survivors.

Treatment of cancer with ionizing radiation led to the discovery that radiosensitivity is a major feature of AT. Treatment employing standard doses of ionizing radiation leads to the development of life-threatening sequelae that are observed in normal individuals only after much higher doses. Radiosensitivity also readily is demonstrated in cultured AT cells.

The question of cancer proneness in carriers of AT mutations, particularly for breast cancer, has yet to be answered. That a degree of radiation hypersensitivity has been reported in cells of AT heterozygotes has raised this possibility. Disruption of the AT locus through mutation or loss of heterozygosity has been reported in a proportion of T-cell prolymphocytic leukemias, B-cell chronic lymphocytic leukemias, and breast cancers.

Cytogenetics

Evidence for genomic instability in AT was reported in 1966 (29). The instability manifests itself in blood lymphocytes in short-term culture in the form of increased numbers of chromatid and isochromatid gaps and breaks, acentric fragments, dicentric chromosomes, and rearranged monocentric chromosomes. Cultured dermal fibroblasts have a breakage frequency even greater than blood lymphocytes. LCLs generally are reported to exhibit a normal level of spontaneous chromosome breakage; however, clones of cells with stable chromosomal rearrangements may be present in such cultures. The frequency of chromosome breakage in freshly aspirated bone marrow has been variable in the few samples that have been examined and reported, being elevated in some samples but normal in others; however, an increased incidence of micronuclei in exfoliated oral mucosal and bladder cells attests to the existence of an increased amount of chromosome breakage *in vivo*.

Telomere association is increased in both interphase and metaphase cells in AT. Telomere shortening, which takes place in normal cells as they age, is reported to be accelerated in AT cells, possibly providing an explanation for the poor growth of cells *in vitro*.

The most characteristic cytogenetic finding in AT is the presence of clones of circulating T lymphocytes that are identifiable as such by the presence of structural chromosomal abnormalities (30, 31). Clones with chromosomal rearrangements are observed in approximately 10% of affected persons. Such mutant clones may account for anywhere from 10% to 95% of the blood lymphocytes. Clones of cells with structurally abnormal chromosomes also have been observed in dermal fibroblasts and LCLs.

The chromosomal rearrangements in the mutant T-cell clones consist principally of translocations and inversions that have taken place in chromosomes 7 and 14 (Fig. 33.5). The distribution of breakpoints on these two chromosomes is nonrandom, specifically affecting bands 7p13, 7q35, 14q11.2, and 14q32.1. The Ig heavy chain locus maps to band 14q32.1 whereas loci for the alpha, beta, and gamma chains of the T-cell receptor map to bands 14q11.2, 7p13, and 7q35, respectively. All of these genes undergo rearrangement in T lymphocytes. Improper reassociation of the DNA sequences of these loci during rearrangement gives rise to the mutant clones of cells with structurally abnormal chromosomes. Examples of translocations observed are t(7;7)(p13;q35), t(7;14)(q35;q11.2) and t(14;14)(q11.2;q32.1) and of inversions are inv(7)(p13q35) and inv(14)(q11.2q32.1). Long-term study of the blood of affected individuals with clonal chromosomal rearrangements has demonstrated that the clones may increase in frequency over time, suggesting that the mutant clones sometimes have a proliferative advantage. Although the role, if any, of such mutant clones in the development of lymphoid cancer is unknown, it is worth noting that rearrangements of chromosome 14 having breakpoints similar to the benign T lymphocyte clones are observed in T-cell lymphocytic leukemia in AT, which suggests that the benign clones may be forerunners of the malignant ones. Alternatively, rearrangements of the same loci found in the benign clones so characteristic of AT may transform the normal lymphocyte into the neoplastic mode from their inception.

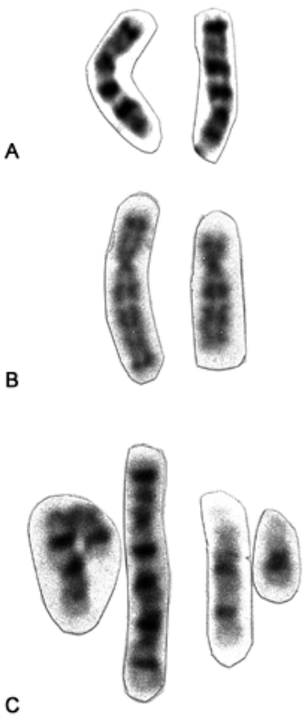


FIGURE 33.5. Partial karyotypes of PHA-stimulated lymphocytes from an individual with AT showing (**A**) inv(7)(p13;q35), (**B**) t(7;7)(p13;q35), and (**C**) t(7;14)(q35;q11.2).

AT cells are strikingly hyperresponsive to the chromosome-breaking action of x-rays and radiomimetic chemicals such as bleomycin. Following irradiation in the G_0 and G_1 stages of the cell cycle cells from normal individuals exhibit increased frequencies of isochromatid gaps, breaks, and interchanges whereas AT cells exhibit increased frequencies of chromatid gaps, breaks,

and interchanges. Following irradiation in the G_2 stage of the cell cycle both normal and AT cells exhibit increased frequencies of chromatid breaks and interchanges; however, the induced aberration frequency is significantly greater in AT. Unirradiated PHA-stimulated AT T lymphocytes exhibit a significantly reduced mitotic index *in vitro* compared to normal cells; following x-irradiation they exhibit significantly less inhibition of mitosis than do normal cells. This results from the failure of AT cells to activate certain cell-cycle checkpoints that normally interrupt DNA synthesis in response to x-irradiation thereby providing cells time to repair DNA damage they have sustained. Unretarded DNA synthesis in AT cells following exposure to x-rays and other radiomimetic chemicals leads to the extensive postradiation chromosomal damage that can be observed cytogenetically.

Genetic Complementation and Heterogeneity

Somatic cell hybridization has been used to demonstrate genetic heterogeneity in AT. The experiments measured radioresistant DNA synthesis in somatic cell hybrids that had been formed by fusion of dermal fibroblasts from individuals with AT from unrelated families. Restoration to normal of DNA synthesis in the x-irradiated hybrid cells was the criterion used for assigning AT cell lines to different complementation groups. Although four complementation groups (A, C, D, and E) were identified in AT (32), molecular techniques have shown the same gene on chromosome 11 to be mutated in each; therefore, the explanation for the existence of multiple complementation groups is obscure.

Molecular Genetics

ATM, the gene for AT, was mapped to chromosome band 11q22 – q23 by genetic linkage studies (33). Microcell-mediated transfer of a normal chromosome 11 tagged with a dominant selectable marker for neomycin resistance into AT complementation groups C and D fibroblasts corrected the radiosensitivity, confirming *ATM*'s location on chromosome 11. ATM was isolated by positional cloning (34). The gene spans 150 kb of DNA and contains 66 exons that produce a 13 kb transcript, which encodes a 350 kDa protein referred to as ATM. ATM is a member of a family of proteins that share a phosphitidylinositol (PI)-3 kinaselike domain at their carboxy termini. ATM also contains a proline-rich region and a leucine zipper motif, suggesting a role in signal transduction. The PI-3 kinases are involved in the regulation of cell-cycle progression, control of telomere length, and response to DNA damage induced by certain DNA-damaging agents.

Mutations throughout *ATM*'s length have been identified in persons with AT. Few recurring mutations have been identified; therefore, most individuals with AT actually are compound heterozygotes, i.e., have different mutations at the two *ATM* loci. DNA haplotyping of unrelated families in whom identical mutations are found generally reveals the families to share a common ancestry. One or more founder mutations have been identified in individuals with AT from Costa Rica, Norway, Poland, Italy, Morocco (in Jews from the Atlas Mountains, now in Israel), Japan, and the United Kingdom. The existence of founder mutations in these populations will provide an opportunity for population screening for AT carrier status.

ATM consistently is mutated in all persons with AT thus far examined regardless of their complementation group, indicating that the basis for the complementation groups does not arise, as had been expected, from mutation at different loci. Most mutations completely inactivate the ATM protein by truncating it, by abolishing its correct initiation or termination, or by deleting large segments. A minority of the mutations result in in-frame deletions and insertions or amino-acid substitutions. Cells from most persons with AT express transcripts for both mutant alleles.

ATM predominantly is a nuclear protein where it interacts with the tumor-suppressor proteins p53 and c-Abl. Following x-irradiation of normal cells the level of tumor-suppressor protein p53 increases in association with a decrease in DNA synthesis. In contrast, AT cells, with their radioresistant DNA synthesis, exhibit a reduced or delayed radiation-induced increase in the level of p53. Interaction of the ATM protein and the protein encoded by c-Abl is thought to alter gene expression in cells. ATM also plays a role in meiosis; by immunofluorescence microscopy, it colocalizes with replication protein A (RPA) along the synaptonemal complex. RPA is a single-strand DNA-binding protein involved in DNA replication, repair, and recombination. The absence of ATM from atm-knockout mice leads to chromosome fragmentation and meiotic arrest in developing gametes. Thus, the functioning of ATM in multiple cellular pathways in multiple cell types may explain the diverse and seemingly unrelated defects that constitute the clinical AT phenotype.

Prenatal Diagnosis

Prenatal diagnosis can be offered to couples in which both partners are obligate heterozygotes for mutation at the AT locus (35). Amniocytes or chorionic villus cells can be examined for increased frequencies of spontaneous and radiation-induced chromosomal aberrations or for radioresistant DNA synthesis. Several affected fetuses have been identified using these methods. With the isolation and characterization of *ATM*, molecular genetic techniques also now can be employed in prenatal diagnosis. Given the fact that few recurring *ATM* mutations have been identified, screening for carrier status is not feasible in most populations; therefore, molecular prenatal diagnosis probably will continue to be limited to couples in which the mutations have been identified following the birth of an affected child.

THE NIJMEGEN BREAKAGE SYNDROME

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

The Nijmegen breakage syndrome (NBS) was described by Weemaes et al. working in Nijmegen (36). The NBS originally was considered to be a variant form of AT because the two disorders share several major abnormal features (see below). Frequency estimates do not exist for the NBS; it has been identified principally in individuals from eastern and central Europe, in particular among individuals of Slavic ancestry.

Birth weight and length are normal in 70% of newborns; however, poor growth during childhood results in short adult stature (Table 33.4). The body proportions are normal of affected individuals with the exception of the head, the circumference of which falls below the third percentile in 75% of newborns; in the remaining 25% progressive and severe microcephaly develops during the first months of life. The facies is characteristic, consisting of a receding forehead, upslanting palpebral fissures, freckling on the nose and cheeks, a prominent nose, a long philtrum, prominent ears, and a receding mandible. Hair on the head is sparse in half of affected individuals.

TABLE 33.4. FEATURES OF THE NIJMEGEN BREAKAGE SYNDROME

Clinical
Postnatal growth deficiency leading to short stature
Characteristic facies
Microcephaly
Receding forehead and mandible
Upslanting palpebral fissures
Prominent nose and long philtrum
Freckling on nose and cheeks
Prominent ears
Scleral telangiectasia
Immunodeficiency
Recurring respiratory, urinary, and gastrointestinal tract infections
Decreased serum concentrations of IgG, IgA, IgG2, and IgG4
Decreased CD3+ and CD4+ cells
Decreased CD4+: CD8+ cell ratio
Increased incidence of lymphomas
Cytogenetic
Lack of spontaneous chromosome breakage
Circulating clones of T lymphocytes with rearranged chromosomes 7 and 14
Hyperresponsiveness to x-rays and radiomimetic agents

Scleral telangiectasia is observed in some affected individuals, and café-au-lait spots may be present on the skin. Cerebellar ataxia, a cardinal feature of AT, is not a feature of the NBS, and serum AFP levels, which are elevated in AT, are normal.

As in AT, both humoral and cellular immunodeficiency commonly are observed in the NBS. Affected individuals have recurring respiratory tract infections and, frequently also, urinary tract and gastrointestinal infections. Serum levels of IgG, IgA, IgG2, and IgG4 are decreased, and CD3+ and CD4+ cells are decreased in frequency as is the CD4+:CD8+ cell ratio. T lymphocytes exhibit a decreased response to PHA stimulation.

The incidence of cancer is increased. Children have a dramatically increased incidence of lymphomas. Other neoplasms that have been observed include rhabdomyosarcoma, meningioma, glioma, and medulloblastoma. Cancer has been the cause of death of approximately 25% of affected individuals.

Cytogenetics

T lymphocyte cultures from individuals with the NBS have not been reported to exhibit an increased frequency of spontaneous chromosome breakage; however, increased chromatid breakage has been observed in NBS fibroblasts. The most characteristic cytogenetic finding is circulating clones of T lymphocytes with translocations and inversions of chromosomes 7 and 14 (36, 37). These chromosomal rearrangements are strikingly similar to those observed in AT (Fig. 33.5). As in AT, the breakpoints involved in the rearrangements coincide with the location of T-cell receptor and immunoglobulin heavy-chain genes — 7p13, 7q35, 14q11.2, and 14q32.1. The incidence of cells with chromosomes 7 and 14 rearrangements is greater in the NBS than in AT, generally being present in 10% to 35% of T lymphocytes. The most common rearrangement in the NBS is inv(7)(p13q35); whereas inv(14)(q11q32), which commonly is encountered in AT, occurs only rarely.

Both T lymphocytes and dermal fibroblasts are hyperresponsive to ionizing radiation and radiomimetic drugs such as bleomycin. The frequency of chromatid aberrations following x-irradiation during S/G_2 is significantly greater than normal in NBS T lymphocytes. In one study, 70% of NBS T lymphocytes exhibited chromatid aberrations following 1 Gy of ionizing radiation as compared to fewer than 10% of T lymphocytes from a normal individual. As in AT, radioresistant DNA synthesis is increased significantly in NBS lymphocytes and fibroblasts following x-irradiation. After a radiation dose of 40 Gy, the rate of DNA replication is 43% to 78% of that in untreated cells whereas the rate in normal cells is reduced to 17% to 38% of that untreated cells.

Genetic Heterogeneity and Complementation

Early complementation studies utilizing radioresistant DNA synthesis as a marker indicated that the NBS was genetically distinct from AT and suggested that affected individuals fell into two complementation groups, V1 and V2 (32). Later studies, utilizing radioresistant DNA synthesis (38) and induced chromosomal aberrations following x-irradiation (39), identified only a single complementation group. Molecular genetic analysis supports the latter findings (see below).

Molecular Genetics

Microcell-mediated transfer of normal human chromosomes into NBS fibroblasts demonstrated that the NBS gene is in bands 8q21-q24 (38), and this regional assignment was confirmed by linkage analysis (40). The gene was isolated by positional cloning (41, 42). It spans 50 kb of DNA and consists of 16 exons that encode a 754 amino-acid polypeptide with a predicted molecular mass of 85 kDa. The protein, named nibrin, has two conserved domains in the amino-terminal region that are found in cell-cycle checkpoint proteins that function in DNA replication and repair — a fork-head-associated domain and a breast-cancer, carboxy-terminal domain. Nibrin associates with Rad50 and Mre11 in a protein complex that functions in DNA double-strand break repair. The complex possesses manganese-dependent, single-stranded DNA endonuclease and 3′->5′ exonuclease activities. The protein is expressed constitutively in all tissues.

The most common mutation in *NBS* is 657del5, which has been suggested to be a founder mutation in Slavs. All the mutations that have been identified in *NBS* lead to the production of truncated proteins.

Prenatal Diagnosis

Analysis of cultured amniocytes and chorionic villus cells, using radioresistant DNA synthesis as a marker, has proven to be a reliable method for prenatal diagnosis of the NBS (43). Now that the *NBS* gene has been isolated, molecular genetic techniques likely will be used for both carrier screening and prenatal diagnosis in families in which the mutation(s) has been identified.

XERODERMA PIGMENTOSUM

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

Xeroderma pigmentosum (XP) was described well over a century ago by von Hebra and Kaposi (44). Frequency estimates range from 1 in 100,000 in Japan and North Africa to 1 in 250,000 to 1 million worldwide.

Affected newborns appear normal, but signs of acute sunlight hypersensitivity of the skin appear in infancy or early childhood (Table 33.5). Minimal exposure to the sun results in unusually prolonged erythema, often accompanied by blistering and peeling. Freckling usually develops in sun-exposed areas by two years of age. The skin becomes dry, scaly, and atrophic, and multiple telangiectatic areas develop as the skin sustains continued damage. Actinic keratoses and other benign forms of neoplasia of the sun-exposed areas of the skin appear, followed or accompanied by the development of cancers in 45% of patients.

TABLE 33.5. FEATURES OF XERODERMA PIGMENTOSUM

Clinical
Skin
Excessive erythema following sun exposure
Freckling and keratoses
Excessive dryness, pigmentation, and atrophy in response to sunlight damage
Benign and malignant neoplasms of many types
Eyes
Photophobia
Conjunctival erythema
Chronic inflammation of bulbar conjunctiva
Scarring of the cornea with blindness, in response to sunlight damage
Nervous system
Loss of deep tendon reflexes
Mental deficiency
Cytogenetic
Lack of spontaneous chromosome breakage
Excessive chromosomal breakage in response to ultraviolet light and certain chemicals

Ocular abnormalities also commonly are observed. Photophobia is present in many affected individuals. Some degree of chronic inflammation of the conjunctiva is characteristic, with atrophy and scarring of the eyelids. Blindness often results from the actinic damage to the cornea. Approximately 20% of individuals with XP have neurological symptoms including areflexia, dysarthria, ataxia, and, in some cases, severe mental deficiency. The progressive neurological deterioration is thought to result from accelerated neuronal death.

Skin cancer, principally basal-cell and squamous-cell carcinoma but also melanoma, is increased enormously in XP. The cancers generally are limited to the face, head, neck, eye, and tip of the tongue, regions of the body that receive excessive sun exposure. Melanoma is increased in both sun-exposed and nonexposed areas. Benign and malignant neoplasms also develop on the eyelids. The mean age of appearance of the first skin cancer is 8 years. Although leukemias, lymphomas, brain tumors, and carcinoma of nonintegumentary structures certainly occur in XP (as in other people), they are impressive by their rarity. The expected life-span of XP patients is reduced by approximately 30 years, the main cause of premature death being complications of skin cancer.

Cytogenetics

Unlike in the other chromosome-breakage syndromes, untreated XP cells fail to exhibit an elevated frequency of spontaneous chromosome breakage; however, they do exhibit cytogenetically detectable hyperresponsiveness following treatment with certain DNA-damaging agents (45). When XP cells are exposed to UV-light (254 nm), the induced chromosomal aberration frequency is significantly greater than in comparably treated normal (nonXP) cells. XP cells also are hyperresponsive with respect to cell killing and to the chromosome-breaking action of 4-NQO, N-acetoxy-2-acetylaminofluorene, HN2, decarbamyl mitomycin C (DMMC), aflatoxin B_1, and sterigmatocystin. The hypersensitivity to these chemicals is accompanied by decreased DNA-repair synthesis.

Untreated XP dermal fibroblasts, blood lymphocytes, and LCLs exhibit a normal baseline SCE frequency; however, the SCE frequency of XP cells of some of the complementation groups (see below) is significantly greater than that of similarly treated nonXP cells following exposure to UV-light. XP cells also exhibit hyperresponsiveness with respect to SCE induction following exposure to 4-NQO, HN2, and DMMC; however, they exhibit a normal SCE response to EMS, methyl methanesulfonate, N-methyl-N'-nitro-N-nitrosoguanidine, ethyl nitrosourea, and dimethyl sulfate. The response of XP cells to SCE induction by different classes of chemicals is related to their inability to repair the specific types of DNA damage caused by the chemicals. DNA damage induced by the chemicals that elicit hyperresponsiveness is defectively repaired by XP cells whereas DNA damage induced by chemicals that elicit a normal response is efficiently repaired.

Genetic Heterogeneity and Complementation

Defective DNA repair in UV-irradiated XP skin fibroblasts was reported in 1968 (46). Following exposure to UV-light different levels of nucleotide-excision repair (NER), detected as so-called unscheduled DNA synthesis (UDS), were observed in different individuals with clinical XP, suggesting very early that heterogeneity might exist at the molecular level. To search for genetic heterogeneity, fibroblasts from unrelated individuals with XP were hybridized, the heterokaryons were exposed to UV-light, and the cells then were exposed for a brief time to tritiated

thymidine (47). The amount of UDS in the nuclei of the heterokaryons was compared to that of diploid nonfused cells on the same slide. Parental cells of the heterokaryons that exhibited normal or near-normal amounts of UDS were assigned to different complementation groups whereas parental cells of the heterokaryons that exhibited low levels of UDS comparable to that of unfused cells were assigned to the same complementation group. Seven complementation groups, XPA-XPG, have been identified using this technique. Individuals belonging to the different complementation groups tend to exhibit different disease severities and different amounts of UDS following exposure to UV-light; however, a striking degree of overlap between groups exists for both the phenotype and DNA-repair level. Besides the seven XP complementation groups that are defective at NER, an XP "variant" exists in which NER is normal but so-called post-replication repair is abnormal. XP variants have the classical clinical XP phenotype.

Molecular Genetics

The genes for all even XP complementation groups and the XP variant have been mapped and cloned (48). The chromosomal location of the gene for each complementation group and the function of the gene is shown in Table 33.6. The proteins that are affected in the different complementation groups are involved in nucleotide excision repair (NER). It is by way of NER that the diverse DNA lesions that result from exposure to chemical and physicial agents are repaired. The NER pathway has two components — transcription-coupled repair (TCR) and global-genome repair (GGR). The repair afforded by proteins involved in TCR is limited to DNA lesions produced on DNA strands that are being transcribed at the time the lesion is formed. Other proteins perform GGR, repairing DNA lesions in the remainder of the genome. All XP complementation groups except XPC are defective in both TCR and GGR; XPC has normal TCR but defective GGR.

TABLE 33.6. MOLECULAR GENETICS OF THE CHROMOSOME BREAKAGE SYNDROMES

Syndrome	Complementation Group	Chromosome Locus	Enzyme/Activity
BS		15q26.1	BLM; RecQ helicase with 3′ ->5′ helicase activity
FA	A	16q24.3	Cytoplasmic protein
FA	C	9q22.3	
FA	D	3p22-p26	Cytoplasmic protein interacting with cyclin-dependent kinase cdc2, endoplasmic reticulum protein GRP94, and other 60, 50, and 35 kDa cytoplasmic proteins
FA	G	9p13	Suspected involvement in DNA post-replication repair or cell-cycle checkpoint control
AT		11q22-q23	Phosphitidylinositol 3-kinase
NBS		8q21	Nibrin; component of protein complex that functions in DNA replication and repair
XP	A	9q34	Recognition of pyrimidine dimers
XP	B	2q21	Forms complex with TFIIH; possesses 3′ ->5′ DNA helicase activity
XP	C	3p25.1	Recognition of lesions on transcriptionally inactive DNA and nontranscribed DNA strands of active genes
XP	D	19q13	Forms complex with TFIIH; possesses 5′ >3′ DNA helicase activity
XP	E	11	Recognition of [6-4] photodimers
XP	F	16p13.3	Endonuclease that makes incision 6 nucleotides distal to 3′ end of adduct
XP	G	13q33	Endonuclease that makes incision 22 nucleotides distal to 5′ end of adduct
XP variant			DNA polymerase v
WS		8p21	WRN; RecQ helicase

Repair processes that are induced following the introduction of lesions into the DNA by a damaging agent are complex, requiring the participation of multiple proteins. The proteins that are mutant in the several XP complementation groups represent only a small number of the proteins required for efficient DNA repair. The protein products for XPA and XPE are damage-recognition factors. XPA protein recognizes cyclobutane dimers formed between adjacent thymine residues; XPE protein recognizes [6-4] photodimers formed between adjacent thymine and cytosine residues. XPA protein additionally recognizes cis-diamminedichloroplatinum(II)-induced adducts. The protein products for XPB and XPD associate with other proteins in the transcription factor IIH complex (TFIIH), a high-molecular-weight protein complex that functions in both NER and in the initiation of mRNA transcription by RNA polymerase II. Both XPB and XPD exhibit DNA helicase activity. XPB possesses 3′-> 5′ DNA helicase activity while XPD's activity is 5′-> 3′. The protein products for XPF and XPG exhibit endonuclease activity. They are involved in the incision steps that take place on either side of the DNA adduct before it is removed. Prior to the incision step, ATP-dependent formation of an open DNA structure of approximately 25 nucleotides in size occurs around the adduct. XPF makes an incision six nucleotides distal to the 3′ end of the adduct; XPG makes an incision 22 nucleotides distal to the 5′ end of the adduct. XPF and XPG act at the juncture between double- and single-stranded DNA. The XPC protein product is involved only in GGR. Lesions on transcriptionally inactive DNA sequences as well as on the nontranscribed strand

of active genes serve as its substrate. The 125 kDa XPC protein forms a complex with HHR23B, a 58 kDA protein. The complex is required for the excision of thymine dimers from DNA. XP variant cells exhibit defective post-replication repair. The gene defective in XP variant is DNA polymerase η (49). DNA polymerase η inserts adenine residues on the complementary DNA strand opposite thymine dimers enabling the dimers to be bypassed during replication. Absence of DNA polymerase η in XP variant cells causes DNA replication to be error-prone, leading to the introduction of many mutations.

Prenatal Diagnosis

Prenatal diagnosis for XP is available to those couples known to be at risk because of the birth to them earlier of an affected child (50). Cultured amniocytes and chorionic villus cells are exposed to UV-light, and the amount of UDS in the fetal cells with the suspected diagnosis of XP is compared to that in cells from a control fetus. An abnormally decreased amount of UDS in the suspected fetus confirms the diagnosis XP. Molecular diagnosis, looking for the presence of specific mutations in the fetus, probably will be available in the future at least for families in which the causative mutations in the XP gene have been identified.

WERNER'S SYNDROME

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Clinical Features

Werner's syndrome (WS) occurs with a frequency of 1 – 22/1,000,000 individuals. The syndrome was described almost a century ago by Werner in a family of two brothers and two sisters who exhibited short stature, premature graying of the hair, cataracts, sclerodermalike skin changes, and hyperkeratotic lesions on the soles of the feet (51). Consanguinity is increased in frequency in parents of affected individuals.

Individuals with WS appear normal throughout childhood and early adolescence (Table 33.7) (52). Short stature, the first symptom to appear, becomes apparent around the time of puberty and is accompanied by low body weight. Average height and weight of adult men with WS is 157 cm and 45 kg, respectively; that of adult women is 146 cm and 40 kg. The trunk is stocky whereas the arms and legs are thin and spindly because of the significant degree of muscle atrophy.

TABLE 33.7. FEATURES OF WERNER'S SYNDROME

Clinical
Short stature
Stocky trunk with thin extremities
High-pitched or hoarse voice
Premature graying and thinning of hair
Cataracts
Scleroderma-like skin changes
Hypogonadism
Decreased fertility
Diabetes mellitus
Atherosclerosis
Increased sarcomas
Cytogenetic
Increased spontaneous chromosomal breakage
Variegated translocation mosaicism

Premature graying of the hair begins in late adolescence, often accompanied by hair loss, particularly on the head, eyelashes, and eyebrows. Axillary and pubic hair also is lost in some patients. Cataracts of the type associated with advanced age develop during the third decade of life and generally are bilateral.

Atrophic skin changes develop during the second decade of life. The sclerodermalike skin changes are noticeable particularly on the face, hands, and feet where the skin becomes tight, shiny, and smooth as result of the replacement of subcutaneous tissue by connective tissue. Hyperkeratoses and calluses develop, particularly on the feet and ankles. The calluses ulcerate and are slow to heal, sometimes making amputation necessary.

The voice is high-pitched or hoarse in approximately half of affected individuals.

Hypogonadism is common. Men generally have small external genitalia while women have infantile external and internal genitalia. Women generally have irregular menses and a premature menopause. Both sexes may be fertile, but fertility apparently is reduced. Sometimes generalized osteoporosis occurs, with involvement particularly of the hands, legs, and feet. Atherosclerosis is common, as is vascular calcification in the lower extremities. Diabetes mellitus is present in approximately 50% of cases.

Individuals with WS have an increased propensity to develop both sarcomas and carcinomas with an overrepresentation of nonepithelial, soft-tissue neoplasia. The most common malignancies are soft-tissue sarcomas, osteosarcomas, melanomas, and thyroid carcinomas; benign meningiomas also occur. Individuals with WS generally die by 47 years of age, the principal causes of death being cancer and vascular accidents.

Cytogenetics

Genomic instability is a striking feature of WS. Although early published cytogenetic reports gave equivocal results, it now generally is agreed that the frequency of spontaneous chromosomal aberrations is increased in both PHA-stimulated peripheral blood lymphocytes and LCLs, as first was reported by Nordenson (53). The most striking cytogenetic feature of WS is the increased frequency of clones of fibroblast cell lines with structurally rearranged chromosomal complements described in 1975 (54). Hoehn et al. reported that 13 of 15 fibroblast cultures established from multiple skin samples taken postmortem from a woman with WS had at least one mutant clone of cells with a unique structural chromosomal rearrangment. Multiple fibroblast cell lines and LCLs established from other persons with WS also have exhibited clones of cells with different structural chromosomal rearrangements.

The term *variegated translocation mosaicism* was coined by Hoehn et al. to describe the existence of multiple clones of cells each having a different structural chromosomal rearrangement in cell lines derived from a single individual. It is unclear whether clones of cells with structurally rearranged chromosomal complements exist *in vivo*; however, the occasional finding of identically rearranged chromosomal complements in different cell lines derived from different biopsies from a single affected individual

supports an *in vivo* origin. The basis for the presence of clonal chromosomal rearrangements in long-term WS cell lines is unclear; however, it may be related to the decreased proliferative response of WS cells *in vitro*.

Fibroblast cultures derived from WS skin have a remarkably shortened life span *in vitro*, typically surviving for only 12.7 population doublings compared to survival for over 50 for normal cells (55). The average cell cycle time of WS cells is prolonged principally because of a longer duration of S.

Genomic instability exists *in vivo* in WS. The frequency of blood lymphocytes mutant at the *HPRT* locus is elevated 8-fold above normal; approximately half of the mutations result from deletions of more than 20 kb of DNA from the *HPRT* locus. *In vitro* DNA ligation studies also have demonstrated that WS cells introduce two to five times more mutations than normal cells during ligation of DNA ends. The mutations mainly are a result of increased numbers of DNA deletions.

Molecular Genetics

WRN, the gene mutated in WS, was mapped to the short arm of chromosome 8 by homozygosity mapping, taking advantage of the fact that both the *WRN* gene locus and its closely flanking polymorphic markers in persons born to consanguineous unions are identical by descent. *WRN* then was isolated by positional cloning (56). The gene spans a genomic region that is at least 100 kb in size and contains 35 exons that encode a 1,432 amino-acid polypeptide.

The function of the WRN protein is unknown. The predicted protein has a central domain containing the seven helicase sequence motifs common to all DNA/RNA helicases, including *BLM*. It shares homology with the RecQ family of DNA helicases. WRN has a nuclear localization signal and has been found only in the cell nucleus. Expression studies have revealed that WRN has an ATP-dependent 3′–> 5′ DNA unwinding activity making it a bona fide DNA helicase. It also has an amino-terminal exonuclease domain with significant similarity to the 3′–> 5′ proofreading domain of *Escherichia coli* DNA polymerase I. Mutations in *WRN* frequently lead to the production of truncated proteins; missense mutations have never been identified. In Japan, where WS is more common than in other countries, analysis of mutations in *WRN* has revealed the existence of a common founder mutation. Elsewhere, persons with WS have different *WRN* mutations so that most of them are compound heterozygotes for the disorder.

Prenatal Diagnosis

With the identification of the gene responsible for WS, prenatal diagnosis now is possible. Prenatal diagnosis probably will be restricted to those families in which the parents are known obligate heterozygotes as result of previously having had an affected child.

SUMMARY

Part of "33 - The Chromosome-Breakage Syndromes: Clinical Features, Cytogenetics, and Molecular Genetics"

Genetically determined genomic instability, first identified in BS, is the subject of basic laboratory investigation in the six rare clinical disorders that comprise the chromosome-breakage syndromes. The chromosome-breakage syndromes differ from one another both clinically and genetically, and their grouping is solely for heuristic purposes. Each disorder is very rare. The pattern of genomic instability in each is unique as are the types and distribution of cancers. It is believed that the explanation for the instability that exists in each disorder will lead to an understanding of various mechanisms by which stability of the genome is maintained. Because of the rarity of the disorders, unraveling the nature of and the basis for the syndromes quite possibly will have a greater impact on human biology that it does on clinical medicine. Yet, the recent attention given this group of rare disorders has resulted in better definition of the clinical syndromes, doubtless of value to affected families.

The past decade has witnessed major advances in our understanding of the molecular bases for the several chromosome-breakage syndromes. The chromosomal locations of 15 of the approximately 20 genes responsible for the disorders have been identified (all except genes for four of the eight FA-complementation groups), and currently the roles that the various newly recognized protein products play in cellular metabolism are subjects of interesting study (Table 33.6). As might have been predicted from the different forms of increased genomic instability that repeatedly have been documented for each of the syndromes, the proteins, which when mutant cause the various syndromes, for the most part have been associated either with DNA replication or with the repair of DNA lesions that are induced following exposure to chemical and physical agents. Some of the proteins interact directly with the damaged DNA; others function indirectly in DNA repair by stimulating the activity of other enzymes. None of the proteins has been identified as being essential for cell survival. Because of the duplicaton that exists in the cell's DNA replication/repair machinery, cells of the chromosome-breakage syndromes are just less efficient at these activities than normal. The multiple activities that some of the proteins have in cells explain how defects in single genes can exert a wide array of phenotypic effects within each of the clinical syndromes. The extensive new information acquired during the study of the chromosome-breakage syndromes has led to a better appreciation of the complexity of the cellular processes that are involved in DNA replication and DNA repair in mammalian cells.

An immediate impact that the advances will have in clinical medicine is in the area of prenatal diagnosis. Prior to the identification of genes responsible for the different disorders, prenatal diagnosis was limited to those families in which couples had been identified as obligate heterozygotes through the birth of an affected child. Individuals belonging to ethnic groups in which one or a very few mutations have been identified, such as blm^{Ash} in BS and IVS4+4 in FA complementation group C individuals who are Ashkenazi Jews, now can learn their carrier status prior to the birth of an affected child. If both partners are carriers of a mutation for a given disorder, prenatal diagnosis can be performed to determine the genetic constitution of the fetus. Testing of carrier status now is commercially available for BS and FA complementation group C individuals of Ashkenazi Jewish ancestry and may become available for ethnic groups in other disorders as founder mutations are identified. Because of the existence

of so many unique mutations in carriers not belonging to specific ethnic groups, recognition that these individuals are obligate heterozygotes for one of the chromosome-breakage syndromes probably will continue to be delayed until the birth of an affected child; once this occurs, family members can be identified unequivocally as being either carriers or noncarriers, and this information can be used in planning future pregnancies.

As a more thorough understanding is gained of how the proteins that are defective in the different syndromes function in normal cells, it may become possible to develop ways to alleviate or eliminate some of the clinical complications that occur during an affected person's lifetime. Such attempts already are being made in FA using genetic engineering and bone marrow transplantation to introduce normal FA genes into affected persons.

REFERENCES

1. Bloom, D. Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs. *Am J Dis Child* 1954;88:754–758.

2. German J, Passarge E. Bloom's syndrome. XII. Report from the Registry for 1987. *Clin Genet* 1989;35:57–69.

3. German J. Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet* 1997;93:101–106.

4. German J. Cytological evidence for crossing-over in vitro in human lymphoidcells. *Science* 1964;144:298–301.

5. German J, Archibald R, Bloom D. Chromosomal breakage in a rare and probably genetically determined syndrome of man. *Science* 1965;148:506–507.

6. Ray JH, German J. The chromosome changes in Bloom's syndrome, ataxia-telangiectasia, and Fanconi's anemia. In: Arrighi FE, Rao PN, Stubblefield E, eds. *Genes, Chromosomes and Neoplasia*. New York: Raven Press, 1981:351–378.

7. Chaganti RSK, Schonberg SA, German J. A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci USA* 1974;71:4508–4512.

8. German J, Schonberg S, Louie E, et al. Bloom's syndrome. Sister-chromatid exchange in lymphocytes. *Am J Hum Genet* 1977;29:248–255.

9. Weksberg R, Smith C, Anson-Cartwright L, et al. Bloom syndrome: a single complementation group defines patients of diverse ethnic origin. *Am J Hum Genet* 1988;42:816–824.

10. McDaniel LD, Schultz RA. Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. *Proc Natl Acad Sci USA* 1992;89:7968–7972.

11. German J, Roe AM, Leppert M, et al. Bloom syndrome: an analysis of consanguineous families assigns the locus mutated to chromosome band 15q26.1. *Proc Natl Acad Sci USA* 1994;91:6669–6673.

12. Ellis NA, Lennon DJ, Proytcheva M, et al. Somatic intragenic recombination within the mutated locus *BLM* can correct the high sister-chromatid exchange phenotype of Bloom syndrome cells. *Am J Hum Genet* 1995;57:1019–1027.

13. Ellis NA, Groden J, Ye T-Z, et al. The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 1995;83:655–666.

14. Auerbach AD, Rogatko A, Schroeder-Kurth TM. International Fanconi Anemia Registry: First Report. In: Schroeder-Kurth TM, Auerbach AD, Obe G, eds. *Fanconi Anemia Clinical, Cytogenetic and Experimental Aspects*. Berlin: Springer-Verlag, 1989:3–17.

15. Fanconi G. Familiare infantile perniziosaartige Anamie (pernizioses Blutbild und Konstitution). *Z Kinderheilkd* 1927;117:257–280.

16. Fanconi G. Familial constitutional panmyelocytopathy, Fanconi's anemia (F.A.): I. Clinical aspects. *Semin Hematol* 1967;4:233–240.

17. Schroeder TM, Anschultz F, Knopp A. Spontane Chromosomenaberrationen bei familiarer Panmyelopathie. *Humangenetik* 1964;1:194–196.

18. Auerbach AD, Adler B, Chaganti RSK. Prenatal and postnatal diagnosis and carrier detection of Fanconi anemia by a cytogenetic method. *Pediatrics* 1981;67:128–135.

19. Duckworth-Rysiecki G, Cornish K, Clarke CA, et al. Identification of two complementation groups in Fanconi anemia. *Somat Cell Genet* 1985;11:35–41.

20. Joenje H, Oostra AB, Wijker M, et al. Evidence for at least eight Fanconi anemia genes. *Am J Hum Genet* 1997;61:940–944.

21. Strathdee CA, Duncan AMV, Buchwald M. Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nature Genet* 1992;1:196–198.

22. Strathdee CA, Gavish H, Shannon W, et al. Cloning of cDNAs for Fanconi anemia by functional complementation. *Nature* 1992;356:763–767.

23. Lo Ten Foe JR, Rooimans MA, Bosnoyan-Collins L, et al. Expression cloning of a cDNA for the major Fanconi anaemia gene FAA. *Nature Genet* 1996;14:320–323.

24. de Winter JP, Waisfisz Q, Rooimans MA, et al. The Fanconi anaemia group G gene *FANCG* is identical to *XRCC9*. *Nature Genet* 1998;20:281–283.

25. Auerbach AD, Sagi M, Adler B. Fanconi anemia: prenatal diagnosis in 30 fetuses at risk. *Pediatrics* 1985;76:794–800.

26. Syllaba L, Henner K. Contribution a l'independence de l-athetose double idiopathique et congenitale. *Rev Neurol* 1926;33:541–562.

27. Louis-Bar D. Sur un syndrome progressif comprenant des telangiectasies capillaires cutanees et conjonctivales symetriques, a disposition naevoide et des troubles cerebelleux. *Confin Neurol* 1941;4:32–42.

28. Morrell D, Cromartie E, Swift M. Mortality and cancer incidence in 263 patients with ataxia-telangiectasia. *J Natl Cancer Inst* 1986;77:89–92.

29. Hecht F, Koler RD, Rigas DA, et al. Leukemia and lymphocytes in ataxia-telangiectasia. *Lancet* 1966;ii:1193.

30. McCaw BK, Hecht F, Harnden DG, et al. Somatic rearrangement of chromosome 14 in human lymphocytes. *Proc Natl Acad Sci USA* 1975;72:2071–2075.

31. Oxford JM, Harnden DG, Parrington JM, et al. Specific chromosome aberrations in ataxia telangiectasia. *J Med Genet* 1975;12:251–262.

32. Jaspers NGJ, Gatti RA, Baan C, et al. Genetic complementation analysis of ataxia telangiectasia and Nijmegen breakage syndrome: a survey of 50 patients. *Cytogenet Cell Genet* 1988;49:259–263.

33. Gatti RA, Berkel I, Boder E, et al. Localization of an ataxia-telangiectasia gene to chromosome 11q22-23. *Nature* 1988;336:577–580.

34. Savitsky K, Bar-Shira A, Gilad S, et al. A single ataxia-telangiectasia gene with a product similar to PI-3 kinase. *Science* 1995;268:1749–1753.

35. Giannelli F, Avery JA, Pembrey ME, et al. Prenatal exclusion of ataxia-telangiectasia. In: Bridges BA, Harnden DG, eds. *Ataxia-Telangiectasia: A Cellular and Molecular Link between Cancer, Neuropathology and Immune Deficiency*. New York: John Wiley and Sons, 1982:393–400.

36. Weemaes CMR, Hustinx TWJ, Scheres JMJC, et al. A new chromosomal instability disorder: The Nijmegen breakage syndrome. *Acta Paediatr Scand* 1981;70:557–564.

37. Hustinx TWJ, Scheres JMJC, Weemaes CMR, et al. Karyotype instability with multiple 7/14 and 7/7 rearrangements. *Hum Genet* 1979;49:199–208.

38. Matsuura S, Weemaes C, Smeets D, et al. Genetic mapping using microcell-mediated chromosome transfer suggests a locus for Nijmegen breakage syndrome at chromosome 8q21-24. *Am J Hum Genet* 1997;60:1487–1494.

39. Stumm M, Sperling K, Wegner RD. Noncomplementation of radiation-induced chromosome aberrations in ataxia-telangiectasia/ataxia-telangiectasia-variant heterodikaryons. *Am J Hum Genet* 1997;60:1246–1251.

40. Saar K, Chrzanowska KH, Stumm M, et al. The gene for ataxia-telan giectasia variant, Nijmegen breakage syndrome, maps to a 1-cM interval on chromosome 8q21. *Am J Hum Genet* 1997;60:605–610.

41. Matsuura S, Tauchi H, Nakamura A, et al. Positional cloning of the gene for Nijmegen breakage syndrome. *Nat. Genet* 1997;19:179–181.

42. Varon R, Vissinga C, Platzer M, et al. Nibrin, a novel DNA double-strand break protein, is mutated in Nijmegen breakage syndrome. *Cell* 1998;93:467–476.

43. Kleijer WJ, van Der Kraan M, Los FJ, et al. Prenatal diagnosis of ataxia-telangiectasia and Nijmegen breakage syndrome by the assay of radioresistant DNA synthesis. *Int J Radiat Biol* 1994;66(suppl 6):s167–174.

44. von Hebra F, Kaposi M. On diseases of the skin, including the exanthemata (Tay W, transl.). *New Sydenham Soc* (Lond) 1874;3:252–258.

45. Parrington JM, Delhanty JDA, Baden HP. Unscheduled DNA synthesis, U.V.-induced chromosome aberrations and SV_{40} transformation in cultured cells from xeroderma pigmentosum. *Ann Hum Genet* (Lond) 1971;35:149–160.

46. Cleaver JE. Defective repair replication of DNA in xeroderma pigmentosum. *Nature* 1968;218:652–656.

47. De Weerd-Kastelein EA, Keijzer W, Bootsma D. Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. *Nature New Biol* 1972;238:80–83.

48. Bootsma D, Kramer KH, Cleaver JE, et al. Nucleotide excision repair syndromes: Xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer*. New York: McGraw-Hill, 1997:245–274.

49. Masutani C, Kusumoto R, Yamada A, et al. The *XPV* (xeroderma pigmentosum variant) Gene encodes human DNA polymerase η. *Nature* 1999;399:700–704.

50. Ramsay CA, Coltart TM, Blunt S, et al. Prenatal diagnosis of xeroderma pigmentosum. Report of the first successful case. *Lancet* 1974;ii:1109–1112.

51. Werner O. (Hoehn H, transl.). On cataract in conjunction with scleroderma. *Adv Exp Med Biol* 1985;190:1–14.

52. Epstein CJ, Martin GM, Schultz AL, et al. Werner's syndrome: a review of its symptomatology, natural history, pathologic features, genetics and relationship to the natural aging process. *Medicine* 1966;45:177–221.

53. Nordenson I. Chromosome breaks in Werner's syndrome and their prevention in vitro by radical-scavenging enzymes. *Hereditas* 1977;87:151–154.

54. Hoehn H, Bryant EM, Au K, et al. Variegated translocation mosaicism in human skin fibroblast cultures. *Cytogenet Cell Genet* 1975;15:282–298.

55. Martin GM, Sprague CA, Epstein CJ. Replicative lifespan of cultivated human cells: effects of donor's age, tissue, and genotype. *Lab Invest* 1970;23:86–92.

56. Yu C-E, Oshima J, Fu Y-H, et al. Positional cloning of the Werner's syndrome gene. *Science* 1996;272:258–262.

SUGGESTED READINGS

Auerbach AD, Buchwald M, Joenje H. Fanconi anemia. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer*. New York: McGraw Hill, 1997;317–332.

Bridges BA, Harnden DG, eds. *Ataxia-Telangiectasia: A Cellular and Molecular Link between Cancer, Neuropathology and Immune Deficiency*. New York: John Wiley and Sons, 1982.

Gatti RA. Ataxia-telangiectasia. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer*. New York: McGraw Hill, 1997:275–300.

German J, ed. *Chromosome Mutation and Neoplasia*. New York: Alan R. Liss, Inc., 1981.

German J, Ellis NA. Bloom syndrome. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer*. New York: McGraw Hill, 1997;301–315.

Ray JH, German J. Sister chromatid exchange in the chromosome breakage syndromes. In: Sandberg, AA, ed. *Sister Chromatid Exchange*. New York: Alan R. Liss, Inc., 1982:553–577.

Salk, D, Fujiwara, Y, Martin, GM, eds. *Werner's Syndrome and Human Aging*. New York: Plenum, 1985.

Schellenberg GD, Miki T, Yu C-E, et al. Werner syndrome. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer*. New York: McGraw Hill, 1997;347–359.

Schroeder-Kurth TM, Auerbach AD, Obe G, eds. *Fanconi Anemia Clinical, Cytogenetic and Experimental Aspects*. Berlin: Springer-Verlag, 1989.

34
Solid Tumor Cytogenetics

Fredrik Mertens

Thomas E. Carey

HISTORICAL PERSPECTIVE

Part of "34 - Solid Tumor Cytogenetics"

The discovery of a consistent, specific chromosome abnormality in chronic myelogenous leukemia (CML) in 1960 by Nowell and Hungerford gave birth to the field of cancer cytogenetics. The visibly altered copy of chromosome 22 that they described became known as the Philadelphia chromosome (Ph). This marker, which is the result of a reciprocal translocation between chromosomes 9 and 22, is present in more than 95% of CML cases and is so commonly known that, even among noncytogeneticists and nonhematologists, Ph is regarded as pathognomonic for CML.

In the years following the description of the Philadelphia translocation in CML, multiple specific chromosome changes have been identified in various hematologic neoplasms. Some chromosome abnormalities are highly specific for a restricted set of malignant diseases; e.g., t(15;17)(q22;q11) is found in nearly all acute nonlymphocytic leukemias of the French-American-British (FAB) M3 type, but it has not been reported in other hematologic diseases (see Chapter 29, Basic Cytogenetics). Other rearrangements such as t(8;14), t(2;8), and t(8;22) involving the loci for MYC (c-myc) on chromosome 8 and the immunoglobulin genes on chromosomes 14, 2, and 22 originally observed in Burkitt's lymphoma are common to several types of B-cell neoplasms.

Thus, as described in Chapter 29, specific chromosome changes can be used to distinguish tumors of a specific type (e.g., B-cell leukemias can be distinguished from T-cell leukemias by rearrangements affecting the immunoglobulin genes and the T-cell receptor genes, respectively). Specific chromosome rearrangements in some cases also have been shown to be independent prognostic indicators. For example, t(8;21) (q22;q22) is associated with a relatively good prognosis in acute nonlymphocytic leukemia, whereas t(4;11) (q21;q23) is associated with a poor prognostic outlook in acute lymphocytic leukemia and acute undifferentiated leukemia.

APPLICATIONS OF TUMOR CYTOGENETICS

Part of "34 - Solid Tumor Cytogenetics"

As indicated above, specific chromosome alterations are useful in classifying hematologic tumors according to type. They also may have prognostic significance and, perhaps most importantly, can provide essential clues to the location of genes that may be mutated, lost, inactivated, or aberrantly expressed as a result of (or in combination with) chromosome rearrangements associated with tumor development and progression. This chapter discusses specific chromosome rearrangements that have been identified in a variety of solid human tumors. By the end of the chapter, we hope to have conveyed to the reader a sense of the emerging picture of nonrandom changes in individual solid tumor types. Knowledge of the consistent chromosome abnormalities in solid tumors will serve as a map to the genes that are altered during the neoplastic process. We will illustrate how some cancer-associated chromosome abnormalities have already led to the isolation of genes and the identification of gene products that are not only important in the cancer process but are essential for normal cell growth and development. The elucidation of how gene expression is altered by chromosome rearrangement will provide the basis for predicting the clinical behavior of individual tumors and should lead to the development of novel strategies for cancer treatment and prevention.

METHODOLOGICAL CONSIDERATIONS

Part of "34 - Solid Tumor Cytogenetics"

Although the role of chromosome changes in hematologic cancers is well established, until recently only a small proportion (6% to 13%) of cancer cytogenetic studies concerned solid tumors. Of these, the majority dealt with benign tumors and pediatric tumors, which generally contain relatively few chromosome abnormalities. Therefore, by 1989, only 1% of the published cytogenetic reports about cancer dealt with common epithelial tumors or carcinomas, which represent 80% of all human cancers. In the intervening 10 years, much progress has been made largely as a result of improved cultivation techniques and new methods of genomic evaluation that do not require chromosome preparations from individual tumor samples.

Complicating Factors in the Cytogenetic Analysis of Solid Tumors

To analyze chromosomal integrity with standard techniques, it is necessary to examine the cells during mitosis. However, in a significant proportion of solid tumors, the mitotic index is low, making it difficult to obtain sufficient mitotic cells for cytogenetic

characterization. This is further complicated by the fact that epithelial tumors are surrounded by stromal components that limit access to the mitotic cancer cell populations. This difficulty in gaining access to the tumor cells often results in chromosome preparations that are fuzzy, incomplete, and less than optimal for complete characterization. Compounding these problems, solid tumors usually show multiple chromosome changes, many of which are complex and some of which have been unidentifiable. Solid tumors also frequently have high ploidy with near triploid or tetraploid counts or more, especially in more advanced tumors. Furthermore, carcinomas, which arise from epithelial surfaces such as the upper aerodigestive tract, the gastrointestinal tract, the lower urogenital tract, and the skin, may be contaminated with microorganisms that can destroy even short-term cultures. Additionally, for a thorough cytogenetic examination, analysis of a minimum number of metaphases is necessary, and without adequate recovery of viable tumor cells, it is difficult to accurately determine the consensus karyotype. For many solid tumors, these problems have now been overcome with the development of more effective culture techniques. As a result, the field of solid cancer cytogenetics is expanding rapidly.

Advances in Culture Techniques for Cytogenetic Analysis of Solid Tumors

Advances in culture methodology include the following: careful selection by the surgeon of viable, uninfected tissue for culture; the adequate washing of tumor specimens with antibiotics, including antifungal agents, to prevent microbial outgrowth; improvements in disaggregation techniques; better media formulations for epithelial cells; use of feeder layers; a better understanding and application of growth factors and hormones; the use of decreased serum supplements for certain tumor types and increased serum supplements for others; and the development of methods for removing stromal cell contamination. For disaggregation of solid tumors, probably the most important advance has been in the commercial development of highly purified enzymes, such as collagenase, to break down tumor stroma without damaging the tumor cells. Early preparations contained other proteolytic enzymes that had detrimental effects on tumor cell viability. Currently available purified enzymes effectively digest stromal components and release viable tumor cells in large numbers for direct study and for *in vitro* cultivation. Other improvements have included modifications of culture media, such as supplementation with hormones, particularly hydrocortisone, insulin, and transferrin; the addition of trace elements such as selenium; and the appropriate use of growth factors in the culture media. The elimination of stromal cell contamination by selective trypsinization or by growing tumor cells in soft agar is important, as is the use of irradiated fibroblast feeder layers or the use of extracellular matrix materials such as collagen, fibronectin, or bovine endothelial matrix for better tumor-cell attachment.

In addition to these improvements in culturing, techniques for better chromosome resolution, such as cell synchronization to increase the mitotic fraction, reducing the time of exposure to Colcemid to improve chromosome morphology, and the use of ethidium bromide to maximize chromosome length, have also played a role in obtaining better results in dealing with solid tumors.

In attempting the *in vitro* cultivation of solid tumors, it is important to recognize that methods that have worked for one histologic type are not necessarily effective on other tumor types. For example, standard culture media were ineffective for small cell carcinoma of the lung (SCLC), but great progress was made in the *in vitro* cultivation of this tumor type when a novel medium formulation was used. Use of the HITES medium containing hydrocortisone, insulin, transferrin, estradiol, and selenium, with low (2.5%) or no serum resulted in a 72% success rate for SCLC, which had previously been resistant to culture. For other tumor types such as squamous cell carcinoma (SCC), early misunderstandings regarding their properties may have led to the widely held belief that these tumors were difficult if not impossible to grow *in vitro*. In fact, when appropriate techniques have been employed, success rates exceeding 30% have been routinely obtained in several laboratories. Explant techniques either on plastic or on irradiated feeder cell layers are effective for this tumor type as are careful disaggregation techniques using purified enzymes. Squamous cells are nearly always substrate dependent, probably because they require attachment molecules for mobility and mitotic activity. Because the attachment to substrate is critical to success with SCC, the simple expedient of using only a small volume of media to cover primary explant cultures prevents the tumor fragments from floating away from the substrate and increases the opportunity for squamous cell outgrowth. Another important factor for cultivating SCC and other tumor types is the application of a satisfactory method for removing fibroblasts. Fibroblasts tend to outgrow the tumor cells in early mixed cultures, surrounding and sometimes undermining the epithelial islands. Removing fibroblasts with brief exposures to trypsin-EDTA while monitoring the culture under the microscope to ensure that tumor cells are not detaching eliminates this problem and increases the tumor cell culture success rate. In a few cases, tumor cell islands have been discovered completely buried under sheets of fibroblasts. The tumor cells were not apparent when the culture was viewed under the microscope; if the fibroblasts had not been removed, these tumor cultures might have been discarded. For different tumor types, slightly different methods of cultivation and isolation have been developed. Good reviews of culture techniques for a variety of human tumor cells are presented in the *Atlas of Human Tumor Cell Lines* (eds. Hay, Park, Gazdar, 1994) and in *Human Cell Culture* Vols. 1 and 2, “Cancer Cell Lines” Parts 1 and 2 (J Masters, ed., 1999). A thorough description of the techniques available for obtaining metaphase chromosome preparations from solid tumors is provided in Mandahl (2001).

Advantages Derived from Cultured Tumor Cells

Improved culture methods have increased not only the yield and quality of solid tumor karyotypes but also the accuracy of the characterization of the breakpoints and rearrangements. This is because, in many cases, the tumor cells persist in culture, making additional harvests possible. As a result, it is possible to accurately determine whether a chromosome is missing because of a

random loss from a cell or because monosomy was one of the events in the genesis of the tumor. Similarly, persistent cell cultures can be harvested repeatedly to obtain early-metaphase chromosomes, which are more informative because they are longer and more bands can be clearly discerned.

Cells in culture provide excellent material for special chromosome staining techniques such as C-banding, Q-banding, and silver staining. These techniques are valuable for identifying rearranged chromosomes, for helping the cytogeneticist to determine breakpoints more accurately, for identifying centromeres, and for detecting dicentric chromosomes. Sometimes, C-banding can even help to identify which chromosome of a pair is present in case of monosomy or to determine if both copies might be the same homologue as a result of loss that has been compensated for by duplication and nondisjunction of the remaining chromosome. Silver staining to detect the nucleolus-organizing regions (NORs) of acrocentric chromosomes is valuable for determining if acrocentrics are involved in rearrangements and for distinguishing NORs from homogeneously staining regions (HSRs). This distinction is important because the latter are considered to be indicative of amplified genes. An example of an HSR in a squamous cell carcinoma that contains amplified copies of cyclin D1 is shown in Fig. 34.1. Note that a FISH probe for the cyclin D1 gene stains the entire region of homogeneous staining on the chromosome with the HSR. Hybridization and staining of tumor tissue also can identify amplified gene segments in tumor cell nuclei.

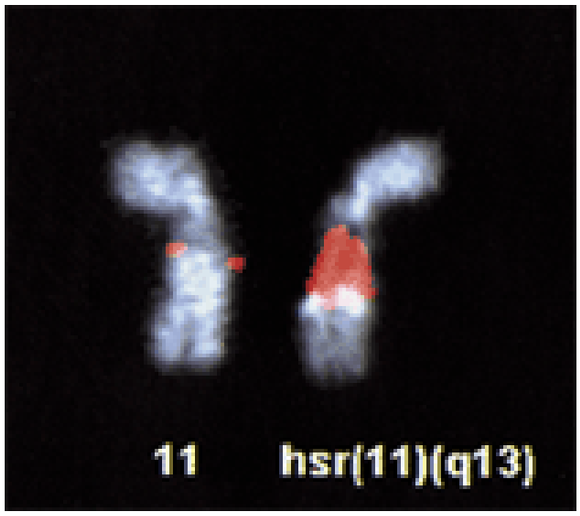


FIGURE 34.1. Partial karyotype from a squamous cell carcinoma of the oral cavity. **Top:** G-banding revealed a homogeneously staining region in band 11q13, (**right**) corresponding to massive amplification of the cell cycle regulatory *CCND1*, as detected by FISH (Courtesy of Dr. Yuesheng Jin, Dept. of Clinical Genetics, Lund University Hospital, Sweden).

In vitro cultures also provide pure populations of tumor cells for studies to further investigate the mechanisms at work at the level of the DNA and RNA molecules in tumor cells with rearranged or deleted chromosomes. Just a few examples of such studies include (i) Southern analysis to detect and study amplified, mutated, or deleted genes, (ii) analysis of simple sequence repeat polymorphisms to detect loss of heterozygosity and homozygous deletion in tumor lines, (iii) Northern analysis to detect aberrant gene expression or inactivation of alleles, and (iv) gene isolation and sequencing to determine the nature of mutations that alter genes or affect gene expression. Finally, cultured cells provide a test system for examining the cell biology associated with altered gene expression.

Controversy Over the Use of Cell Cultures

In spite of the many advantages gained through the use of cultured cells for studies of the genetic changes in cancer, there is controversy over the validity of the karyotypic changes that are defined using cultured cells. Some investigators suggest that culture conditions are not physiologic and may introduce selective pressures that result in clones of cells that are not representative of the *in vivo* tumor population. It also has been suggested that cells grown in culture are unstable and undergo frequent changes or continuous karyotypic evolution *in vitro*.

Addressing the Cell Culture Controversy

To analyze these objections, several points should be considered. As a starting point, it should be noted that following amniocentesis, short-term culture is employed for expanding mitotic cells to detect chromosome abnormalities *in utero* (see Chapter 31), yet accurate results are obtained consistently, suggesting that short-term culture does not routinely introduce karyotypic artifacts. With solid tumors, it is difficult to entirely eliminate some time in culture from the process of karyotyping tumors because it is necessary to harvest mitotic cells to observe the chromosomes in their condensed state. Nevertheless, in some cases it is possible to make “direct preparations” in which karyotypes are prepared from cells released directly from tumor tissue. Direct preparations can serve as controls for evaluating the effects of culture on the tumor populations by comparing results obtained by this method to those obtained after culture. To minimize the possible selective effects of *in vitro* culture, we have also used an “in situ” technique, in which metaphase spreads are prepared from relatively few mitotic cells in short-term cultures by directly treating primary cultures on coverslips with hypotonic solution. In addition, conventional harvests from very early cultures also can minimize selective or evolutionary effects of *in vitro* culture. If random chromosome aberrations arise as a result of *in vitro* culture, then *in vitro* events should accumulate with time in culture. To evaluate this, *in situ* and direct preparations can be compared to early passages and to long-term cultures from the same tumors to document whether *in vitro* changes are significantly skewing the results. In our laboratories, we have observed that the consensus karyotype is represented in both early and later cultures. When a cell culture of an established SCC line was monitored long-term only one chromosome change was detected. In repeated studies over a period that spanned more than 60 *in vitro* passages, a normal 7 was replaced by an i(7p). Even in this case it is possible that the population of cells containing the i(7p) were present all along but were too infrequent to be detected in earlier preparations (Carey et al. 1989). Results from direct preparations, *in situ* preparations on early cultures, and long-term culture have shown that the karyotypes obtained immediately or after days, weeks, and months of *in vitro* culture are remarkably consistent (Worsham et al. 1991). Thus, in our experience, tumor cultures examined after different periods of time in culture nearly always contain the same chromosome rearrangements and the same numerical changes and, in most cases, even maintain similar fractions of near-diploid and near-tetraploid populations. Incidentally, we also have noted that when there are near-diploid and near-tetraploid populations

within the same tumor, each population contains the same rearrangements, and they usually differ from one another only in copy number (Carey et al. 1993, Worsham et al. 1993).

If *in vitro* culture selects for nonrepresentative cell populations with certain rearrangements, then there should be frequent discrepancies between *in vitro* and *in vivo* findings. In earlier studies of leukemias and lymphomas, when the chromosome abnormalities obtained from leukemia cell lines in long-term culture were compared to those found in marrow or blood samples from patients, the same consistent changes and rearrangements for each leukemic type were noted. For example, the original studies of Burkitt's lymphoma showed the same t(8;14) translocation in both fresh specimens and Burkitt lymphoma cell lines.

If tumor cells grown in culture are not representative of the cells that persist in the patient, then we would expect that separate tumor samples from the same patient would represent different clones. When we examine karyotypes from cell cultures taken from primary tumors and metastatic or recurrent tumors from the same donors, the same primary clonal abnormalities are present in both primary and metastatic or recurrent cultures. (For examples of this type of analysis, see the discussion of chromosome changes in squamous carcinomas of the head and neck below). Taken together, these observations support the concept that *in vitro* cultures are a reasonable model of the *in vivo* tumor.

Molecular Allelotyping

It has become possible to detect chromosome losses using polymorphic markers on each chromosome arm. For such studies, it is unnecessary to culture cells for metaphase chromosome analysis. Molecular genetic methods also allow one to determine if karyotypic analyses on cultured tumor cells accurately represent the cells in the tumor specimen. For molecular allelotyping, DNA is harvested from tumor cells and normal cells from the same individual and assessed for loss of heterozygosity using polymorphic markers. Originally this type of analysis employed restriction fragment length polymorphisms (RFLP), but subsequently simple sequence repeat polymorphisms (SSRPs) were discovered that are more numerous and more polymorphic than RFLPs. The simple sequence repeats consist of di-, tri- or tetra-nucleotide repeats that have also been referred to as microsatellite repeat polymorphisms (MSRPs) because satellite DNA surrounding the centromeres also consists of nucleotide repeats. SSRPs appear frequently throughout the genome and have been used very effectively to supplement and confirm consistent chromosome rearrangements identified by conventional cytogenetic methods. Allelotyping employs primer sets that flank individual polymorphic SSRPs that have already been mapped to each chromosome arm. The available sets of primers and SSRPs that are known are very numerous because these have been identified in the course of gene mapping studies and as a by-product of the genome sequencing project. Multiple polymorphic probes are used for each arm of the chromosome to detect loss of heterozygosity (LOH) that corresponds to chromosome deletion. Because the SSRPs on the maternal and paternal chromosome differ in length in 60% to 80% of cases these markers are often informative for evaluating LOH. DNA from normal tissue e.g., normal blood cells and from a tumor-enriched area of the tumor is amplified with the polymerase chain reaction (PCR) using primer sets for an SSRP. These PCR products are separated by size on a DNA sequencing gel. If the repeats are different in the maternal and paternal chromosomes then two bands will be visualized. If one of the chromosomes has been lost in the tumor then only the allele from the intact chromosome will be present in the PCR product. This method has been used to great advantage to define chromosomal regions that have been lost in tumors for which it has not been possible to obtain good chromosome data by karyotyping (Fearon and Vogelstein 1990). Essential to this kind of analysis is good separation of normal and tumor tissue. If normal cells are included in the tumor DNA extractions, there will be significant loss of resolution because the contaminating normal cells contribute a signal from both chromosomes. The allelotyping method has allowed comparison of cell lines to the original tumor tissue from which the lines were established. Frank et al. (1997) compared DNA isolated from 10 head and neck tumors to the cell lines established from each for LOH at multiple polymorphic markers on chromosome arm 18q. In every case the loss or retention of individual alleles was the same in the tumor and the cell line from which it was established. Because multiple markers were assessed in each pair it was also possible to confirm that the breakpoints were the same in the tumor and the cell lines and if part of 18q was retained in the tumor, it was also retained in the cell line. Cell lines also have proved very valuable for defining the minimum region of loss first detected by chromosome studies. Homozygous deletions have been noted to occur in some tumors and these can have the effect of inactivating both copies of tumor suppressor genes. Homozygous deletions affecting both chromosomes are difficult to detect in DNA isolated from tumor tissue because of the presence of contaminating normal cells. Cell lines lack normal cells and can be screened for homozygous deletions using both polymorphic and nonpolymorphic markers such as STS (sequence tagged site) markers. Loss of chromosome 18q arm has been observed by karyotyping in multiple tumor types. Screening pancreatic tumor cell lines for homozygous deletions on chromosome arm 18q resulted in the discovery that *DPC4/SMAD4* is a tumor suppressor gene in pancreatic carcinoma (Hahn et al. 1996). Similarly, screening of human squamous carcinoma cell lines allowed Sun et al. (1999) to define a minimal region of homozygous loss on 8p23.2 that presumably houses a tumor suppressor gene. These observations contribute to the body of data showing that cell lines are highly representative of tumors from which they are derived and strengthen the concept that cytogenetic data obtained from short and even long-term cultures of human tumors accurately portray the situation in the original tumor.

Fluorescence In Situ Hybridization (FISH) — Interphase Cytogenetics and Chromosome Painting

Interphase cytogenetics or chromosome painting, which employs fluorescent, chromosome-specific probes, can be used on cells in interphase to identify individual chromosomes either in tissue sections or in cultures. As discussed above, it is not unusual to obtain only a few metaphase cells in some solid tumors, although there may be abundant interphase cells. Fluorescence *in situ* hybridization (FISH) of chromosome-specific probes to the

nuclei of nonmitotic cells makes it possible to examine the interphase cells for specific genetic changes. For numerical changes, satellite DNA probes that are specific for the centromere of a given chromosome can be used. This method is already effective for characterizing ploidy changes, trisomies, and monosomies, and it may become useful for identifying specific rearrangements as well. As shown in Fig. 34.1 FISH can be used with specific gene probes to detect amplified genes on chromosomes as well as in tumor tissue. Worsham et al. (1993), Worsham et al. (1999) have used FISH to investigate genetic changes in multiple cases of tumor cell lines and the original tumor tissue. Those results have shown identity of both copy number and marker content in tumor tissue and the corresponding cell line. In Fig. 34.2 an example of FISH using centromere specific probes allows the enumeration of chromosome copy number in individual nuclei within tumor tissue (Worsham et al. 1995). In the same study, FISH for a rearranged chromosome was used on a cultured cell line metaphase spread and on tissue sections from both tumors to demonstrate identical clonal origin of synchronous primary tumors in a patient originally thought to have had separate carcinogenic events in the floor of mouth and the pyriform sinus.

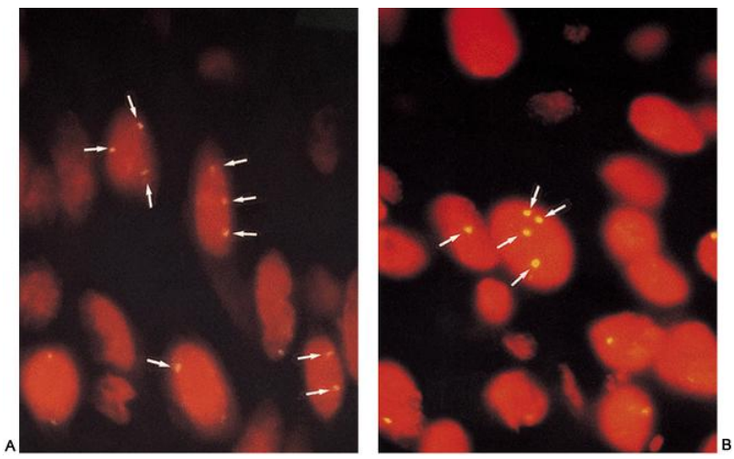


FIGURE 34.2. FISH analysis of a squamous cell carcinoma of the pyriform sinus using centromere-specific probes for chromosomes 7 (**A**) and 17 (**B**). Note that the tumor cell nuclei contain two, three, and four signals each for chromosome in the plane of the photo. (Reproduced from Worsham MJ, Wolman SR, Carey TE, et al. Common clonal origin of synchronous primary head and neck squamous cell carcinomas: analysis by tumor karyotype and fluorescence in situ hybridization. *Human Path* 1995;26:251–261).

Whole chromosome specific probes have been developed and are available commercially. These tools are very useful when a metaphase chromosome spread can be obtained. To detect specific translocations, a mixture of DNA probes for different arms of the chromosomes can be used. An example of multicolor FISH analysis (also called spectral karyotyping) of a hypodiploid myxoid malignant fibrous histiocytoma is shown in Fig. 34.3. Note that the complex rearrangement involving chromosomes 1, 3, 6, and 10 are easily resolved using this technology.

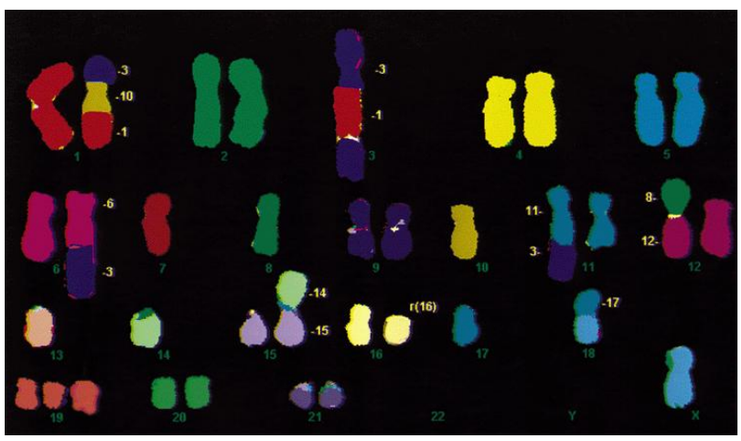


FIGURE 34.3. Complex structural rearrangements in a hypodiploid myxoid malignant fibrous histiocytoma, as detected by multicolor FISH (COBRA FISH, Applied Imaging) analysis (Courtesy of Dr. David Gisselsson, Dept. of Clinical Genetics, Lund University Hospital, Sweden).

Comparative Genomic Hybridization

Comparative genomic hybridization (CGH) combines fluorescence hybridization techniques with a conventional metaphase spread to examine the entire genome from an abnormal cell such as a tumor cell to a normal cell with a normal complement of intact chromosomes. As in the case of allelotyping careful isolation of pure tumor cell DNA is a prerequisite for success. Unlike allelotyping, there is no requirement for normal DNA from the same donor, normal DNA from any individual of the same sex can be used. The DNA from the tumor is labeled by nick translation with biotinylated nucleotides and the DNA from the normal cells is labeled with digoxigenin-labeled nucleotides. These samples are then mixed in equal amounts and hybridized to normal metaphase spreads where the labeled DNA competes for

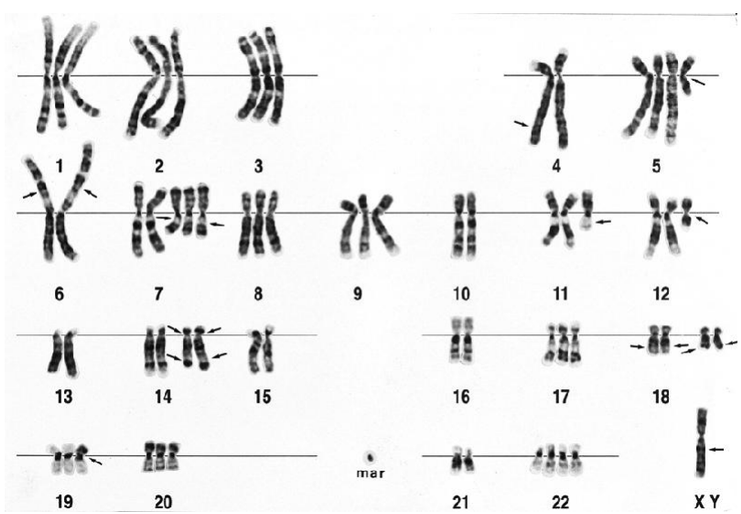


FIGURE 34.4. G-banded karyogram from a pancreatic carcinoma showing multiple numerical and structural rearrangements (Courtesy of Dr. Ludmila Gorunova, Dept. of Clinical Genetics, Lund University Hospital, Sweden).

binding. The hybridization signals are developed with fluorescein-avidin and rhodamine-labeled antidigoxigenin. If there are equal copy numbers in the normal and tumor DNA, then the hybridization signals will be equal and the fluorescein and rhodamine signals will be balanced. However, if there is loss of a chromosome segment in the tumor then the corresponding region of the target metaphase chromosome will be hybridized predominantly with the rhodamine-labeled normal DNA and this will appear red. Similarly if there is a region of chromosomal gain in the tumor, then there will be more hybridization of the corresponding region of the target chromosome with the fluorescein-labeled tumor DNA and this region will appear green. CGH has confirmed results of conventional cytogenetic analysis and has extended the ability to investigate genetic lesions in tumors for which no chromosome spreads can be obtained. Using CGH, Van Dekken et al. (1999) have defined specific regions of gain and deletion in gastroesophageal carcinomas. The current level of technology being developed will use the principle of CGH to assess chromosome gain and loss by hybridization of tumor DNA to gene probes arrayed on slides. As the number of known genes expands it is envisioned that a single hybridization to a gene array chip will identify every example of loss and gain in the entire genome of an individual tumor. The power of this technology will allow customization of therapy to maximize effectiveness based on the genetic changes expressed in an individual patient's tumor.

CYTOGENETIC CHANGES IN SOLID TUMORS

Part of "34 - Solid Tumor Cytogenetics"

Types of Cytogenetic Changes in Solid Tumors

In general, solid tumors from adults and advanced tumors have more numerous and more complicated chromosome changes than hematologic tumors or early pediatric tumors. Many hematological and some pediatric neoplasms are characterized by relatively simple primary abnormalities such as balanced translocations. In contrast, the most common chromosome changes in solid tumors of adults are unbalanced translocations that result in the loss of genetic material (i.e., chromosome deletions). Gains also are frequent and include trisomies of various chromosomes, duplication/amplification of chromosome segments, or amplification of activated protooncogenes. The frequency and extent of ploidy change also is often much greater in solid tumors than in hematologic tumors. An example of a complex karyotype in a solid tumor is shown in Fig. 34.4. This G-banded karyotype from a pancreatic carcinoma shows complex rearrangements, increased copy number of multiple chromosomes, isochromosome formation i(5p) and i(12p), and deletions. In Fig. 34.5, the use of G-banding and FISH to investigate complex alterations in an atypical lipomatous tumor are illustrated. Note that the ring chromosomes contain amplified sequences for both *MDM2* and *CDK4*, genes that can lead to loss of cell cycle control. A summary of the solid tumors known to have clonal chromosome aberrations is given in Table 34.1.

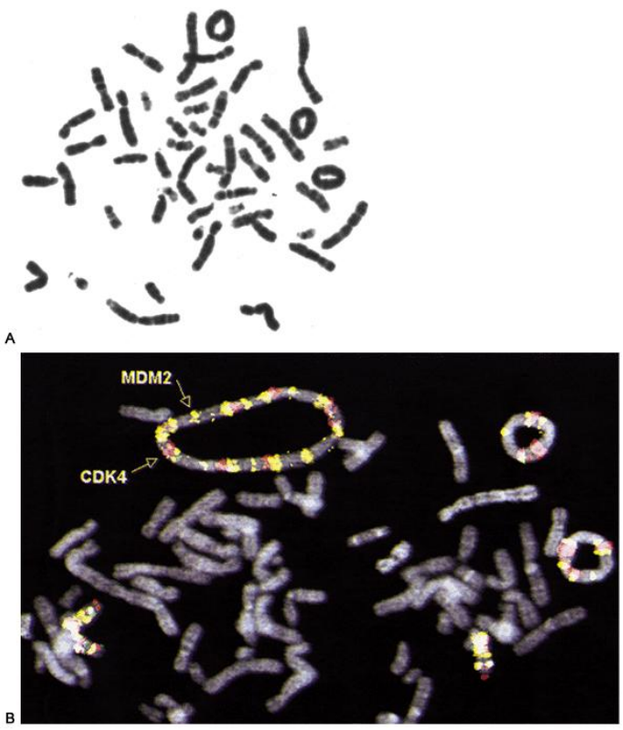


FIGURE 34.5. A: G-banded metaphase spread showing supernumerary ring chromosomes in an atypical lipomatous tumor. These ring chromosomes characteristically contain amplified chromosome 12 sequences (**B**) as shown by two-color FISH using probes for *MDM2* and *CDK4* (Courtesy of Dr. David Gisselsson, Dept. of Clinical Genetics, Lund University Hospital, Sweden).

SIGNIFICANCE OF CHROMOSOME REARRANGEMENTS

Part of "34 - Solid Tumor Cytogenetics"

TABLE 34.1. DISTRIBUTION OF SOLID TUMORS WITH CLONAL CHROMOSOMAL ABERRATIONS BY SITE

Anatomical site	No. of Cases	Anatomical site	No. of Cases
Adrenal gland	112	Parathyroid gland	1
Anus	8	Penis	19
Bladder	142	Peritoneum	8
Brain	1181	Pineal gland	16
Brain stem	23	Pituitary gland	20
Breast	622	Pleura	92
Cerebellum	186	Prostate	180
Eye	198	Salivary glands	364
Fallopian tube	1	Skeleton	421
Gallbladder	5	Skin	203
Heart	13	Small intestine	19
Hypothalamus	2	Soft tissues	1265
Kidney	884	Spinal cord	54
Large intestine	460	Stomach	150
Larynx	53	Teeth	6
Liver	73	Testis	264
Lung	497	Thyroid gland	245
Nasal cavity	24	Tongue	43
Nasopharynx	41	Ureter	3
Oesophagus	51	Urethra	1
Oral cavity	61	Uterus, cervix	90
Oro- and hypopharynx	32	Uterus corpus	525
Ovary	491	Vagina	15
Pancreas	109	**Total**	**9273**

Figures obtained in February 1999 from the database for Mitelman's *Catalog of Chromosome Aberrations in Cancer*.

Translocations, Point Mutations, and Amplification of Protooncogenes

Activation of Protooncogenes

Protooncogenes are normal cellular genes involved in the process of cell replication. When the protooncogenes go awry by any of several mechanisms, the involved cell acquires increased ability to undergo autonomous replication, which results in uncontrolled proliferation. As discussed in Chapter 29, Basic Cytogenetics, in hematological neoplasms, protooncogene activation commonly occurs by chromosome translocations.

Translocations

The classic example of oncogene activation by this type of mechanism is the translocation of the *MYC* protooncogene from chromosome 8 to the immunoglobulin heavy-chain (*IGH*) locus on chromosome 14. As a result of the juxtaposition of *MYC* to the active *IGH* gene, there is increased MYC expression and deregulation of cell growth, driven by the promoter and enhancer elements of the *IGH* gene.

As discussed in Chapter 38, Bone Marrow Transplantation, the translocation of chromosomes 9 and 22 in chronic myelogenous leukemia (CML) provides another example. This translocation fuses the *ABL* protooncogene locus on 9 and a gene locus called breakpoint cluster region (*BCR*) on 22. This translocation

results in the active production of both message and a protein called $P210^{bcr/abl}$ that is similar to the oncogene product of the Abelson murine leukemia virus designated $P160^{gag/v\text{-}abl}$. Both of these rearranged genes encode fusion proteins with deregulated tyrosine-kinase activity. Expression of the $P160^{gag/v\text{-}abl}$ fusion protein induces acute leukemia in mice, and studies from David Baltimore's laboratory demonstrated that mice given murine bone marrow cells transfected with the human P210bcr/abl hybrid construct also develop hematologic tumors and a myeloproliferative disorder that closely resembles CML in humans. This demonstrates most conclusively the significance of the translocation that gives rise to the Philadelphia chromosome.

The consistent translocations in T-cell leukemias have been shown by Carlo Croce and others to often involve the T-cell receptor gene loci and a breakpoint region containing genes that become activated by the translocation. By analogy to the B-cell neoplasms, we can predict that this type of rearrangement involves an active gene, the T-cell receptor, whose promoters and enhancing elements can presumably affect the expression of a gene related to growth regulation on the recipient chromosome. Recent evidence confirms that this is the case and novel genes activated by translocation with the T-cell receptor have been identified.

Based on these observations, it generally is accepted that consistent chromosome rearrangements in cancers are likely to occur at the sites of protooncogenes. Several authors have collected consistent breakpoint data from the literature and compared these sites to known oncogene loci. These comparisons show a strong concordance, which adds further support to the concept that chromosome rearrangements in cancer are functionally related to the cancerous behavior of the tumor cells.

Oncogene Activation

In solid tumors, a major mechanism of oncogene activation appears to be activation by point mutations and gene amplification, rather than activation by translocation. Among the best-studied point mutations and most strongly implicated in the genesis of solid tumors are those of the *RAS* gene family, which consists of three members: *HRAS*, *NRAS*, and *KRAS*. In these cases, specific activating mutations at codons 12, 13, and 61 have been identified. Point mutation of an oncogene has been implicated in the tumorigenesis of several cancers, including bladder (*HRAS*), colon (*KRAS*), pancreas (*KRAS*), lung (*KRAS*), seminoma, and melanoma. *RAS* protooncogene proteins act by binding to GTP, after which the normal RAS product is converted to an inactive form that binds GDP. Point mutations activate *RAS* by eliminating the site necessary for conversion to the inactive GDP-binding form, thus the activated gene product allows unregulated cell growth.

Protooncogene Amplification

Amplifications of HER-2/neu (*ERBB2*), and epidermal growth-factor receptor (*EGFR* or *MYCN*) protooncogenes have been implicated as being prognostically important in neuroblastoma, breast and ovarian cancers, and glial tumors, respectively. Protooncogene amplification also has been reported in several other tumor types as well, but prognostic implications have not been well established in these cases. In neuroblastoma, *MYCN* oncogene amplification has been demonstrated by cytogenetic studies in the form of double minutes (DMs), HSRs, and, more recently, autonomously replicating, small, circular episomes. Double minutes are acentric chromatin bodies that segregate randomly during cell division. When they integrate randomly into a chromosomal

TABLE 34.2. DISTRIBUTION OF EPITHELIAL AND GERM-CELL TUMORS WITH CLONAL CHROMOSOMAL ABERRATIONS BY MORPHOLOGY

Tumor Type	No. of Cases	Tumor Type	No. of Cases
Benign Epithelial Tumors	***653***	Squamous-cell carcinoma	444
Adenoma	571	Basal-cell carcinoma	42
Papilloma	12	Merkel-cell carcinoma	25
Oncocytoma	33	Transitional-cell carcinoma	128
Warthin's tumor	19	Wilms' tumor	175
Various	18	Hepatoblastoma	33
Malignant Epithelial Tumors	***3871***	Carcinoid tumor	21
Carcinoma in situ	30	Chordoma	11
Unclassified	149	Various	46
Undifferentiated carcinoma	45	***Germ Cell and Gonadal***	***395***
Large-cell carcinoma	41	***Stromal Cell Tumors***	
Small-cell carcinoma	56	Unclassified	15
Adenocarcinoma	2543	Seminoma/dysgerminoma	70
Adenoid cystic carcinoma	28	Teratoma	217
Mucoepidermoid carcinoma	27	Combined tumors	54
Acinic-cell carcinoma	10	Gonadal stromal-cell tumors	33
Adenosquamous carcinoma	17	Various	6

Figures obtained in February 1999 from the database for Mitelman's *Catalog of Chromosome Aberrations in Cancer*.

site, an HSR is formed. The size and the number of DMs vary from cell to cell. The presence of DMs in neuroblastoma has been correlated with tumor grade and amplified MYCN expression, which is strongly correlated with rapid tumor progression. In breast cancer and ovarian cancer, amplification and overexpression of the HER-2/neu (*ERBB2*) oncogene has been demonstrated mainly by molecular studies. However, careful examination by cytogenetic methods in a large series of patients might also reveal gene amplification in the form of DMs and HSRs in these two tumor types. As with NMYC in neuroblastoma, there is a correlation between *ERBB2* amplification and elevated *ERBB2* expression and highly malignant behavior in breast cancer and ovarian cancer. In glial tumors, amplification of *EGFR* has been observed in the form of DMs and HSRs, and this is associated with an advanced grade and a poor prognosis.

Chromosome Deletion and Inactivation of Tumor-Suppressor Genes

Activation of an oncogene by translocation, point mutation, or amplification is a dominant phenomenon because malfunction of only one allele is involved in the tumorigenesis. In contrast, inactivation of a tumor-suppressor gene is a recessive mechanism at the cellular level because both alleles must be affected for tumorigenesis. The careful study of tumor-suppressor genes has provided significant insight into the mechanisms that control the cell cycle and into how transforming viruses can induce unregulated cell growth. In cases where there is evidence for genetic losses, one must consider the possibility that loss of a gene that regulates cell growth may have occurred.

CONSISTENT CHROMOSOME REARRANGEMENTS IN SOLID TUMORS

Part of "34 - Solid Tumor Cytogenetics"

In the following survey of the cytogenetic features of solid tumors, the number of references has been limited to recent review articles and, in a few cases, to articles dealing with specific matters of biological or clinical significance. More detailed information about the cytogenetics of solid tumors can be obtained in review articles (Mertens et al. 1997; Mitelman et al. 1997a, Mitelman et al. 1997b), specialized books (Heim and Mitelman, 1995; Mitelman, 1998), or at web sites [Recurrent Chromosome Abberations in Cancer] (http://www.cgap.nci.nih.gov/Chromosomes/RecurrentAberrations); [Mitelman Database of Chromosome Abberrations in Cancer] (http://cgap.nci.nih/gov/Chromosome/Mitelman). Table 34.2, Table 34.3, and Table 34.4 list the numbers of solid tumors by type for which clonal aberrations have been identified.

TABLE 34.3. DISTRIBUTION OF MESENCHYMAL TUMORS WITH CLONAL CHROMOSOMAL ABERRATIONS BY MORPHOLOGY

Tumor Type	No. of Cases	Tumor Type	No. of Cases
Benign Mesenchymal Tumors	***1480***	***Malignant Mesenchymal Tumors***	***1086***
Unclassified	1	Unclassified	14
Fibroma	4	Fibrosarcoma	24
Fibromatosis	85	Malignant fibrous histiocytoma	59
Mesoblastic nephroma	14	Dermatofibrosarcoma protuberans	27
Chondroma/chondroblastoma/chondromyxoid fibroma	18	Clear-cell sarcoma	21
		Liposarcoma	141
Myxoma	10	Leiomyosarcoma	63
Lipoma	257	Rhabdomyosarcoma	106
Lipoblastoma	9	Angiosarcoma	7
Hibernoma	7	Hemangiopericytoma	16
Leiomyoma	375	Kaposi sarcoma	6
Rhabdomyoma	1	Chondrosarcoma	83
Hemangioma	1	Osteosarcoma	103
Lymphangioma	1	Giant-cell tumor of bone	34
Osteocartilaginous exostosis	15	Synovial sarcoma	104
Osteoblastoma	1	Neurofibrosarcoma	46
Tenosynovial giant-cell tumors	14	Epithelioid-cell sarcoma	7
Meningioma	519	Mixed mesodermal tumor	22
Neurofibroma	4	Mesothelioma	95
Neurilemoma	57	Desmoplastic small round-cell tumor	8
Various	87	Various	100

Figures obtained in February 1999 from the database for *Mitelman's Catalog of Chromosome Aberrations in Cancer*.

TUMORS OF THE NERVOUS SYSTEM

Part of "34 - Solid Tumor Cytogenetics"

Tumors of the central nervous system and the surrounding meninges are among the most extensively analyzed solid tumors. Altogether, some 1,500 cases, slightly more than one half of which are gliomas, with clonal chromosomal changes have been reported. In astrocytomas, the most common form of glioma, a nonrandom pattern of chromosomal changes has been established. The karyotypes often are quite complex, with structural rearrangements clustering to chromosome arms 1p, 1q, 6q, 9p, and 19q, and multiple numerical changes, the most common of which are loss of chromosomes X, Y, 10, 13, and 22, and gain of chromosomes 7, 19, and 20. Some of these aberrations, in particular +7, -Y, -10, and -22, may occur as the sole anomalies, suggesting that they may be of particular relevance in the early steps of tumorigenesis. The significance of -Y and +7 as the sole anomalies is, however, debated because both have been detected in nonneoplastic brain tissue as well (Johansson et al. 1993).

Molecular genetic investigations have implicated a great number of oncogenes and tumor suppressor genes in astrocytoma development. Of the suppressor genes that seem to be most commonly deleted and/or mutated — *CDKN2A* in 9p21, *PTEN* in 10q23, *RB1*> in 13q14, and *TP53* in 17p13 — several are located in chromosome segments that characteristically are found

to be deleted at cytogenetic analysis. Frequently amplified oncogenes include *CDK4* in 12q and the epidermal growth factor receptor gene (*EGFR*) in 7p. Amplification of the latter is particularly common in tumors showing double-minute chromosomes, i.e., a cytogenetic sign of gene amplification found in approximately 10% of adult astrocytomas.

Cytogenetic and CGH analyses, as well as molecular investigations, have shown that low-grade astrocytomas have less complex genomic changes than do high-grade tumors, but that the genotypic differences are quantitative rather than qualitative. In a study of 206 patients with gliomas, most of which were astrocytic tumors, the presence of clonal chromosome aberrations was significantly associated with poor clinical outcome. Furthermore, in univariate analysis, rearrangements (most of which were gains and losses, respectively) of chromosomes 7 and 10 were strongly associated with reduced survival (Ganju et al. 1994). In children, low-grade and cerebellar astrocytomas usually have a normal karyotype, whereas high-grade and supratentorial tumors closely resemble the astrocytic tumors in adults. Regarding molecular genetic correlations, *TP53* mutations and loss of 22q material have been found in primary low-grade tumors, whereas most other investigated genes seem to be more commonly involved in high-grade and/or recurrent astrocytomas.

Far less is known about the other main subtypes of glioma, i.e., oligodendrogliomas and ependymomas. The karyotypic pattern among the roughly 50 oligodendrogliomas that have been

TABLE 34.4. DISTRIBUTION OF NEUROGLIAL, NEURONAL, PRIMITIVE NEUROECTODERMAL, AND MELANOCYTIC TUMORS WITH CLONAL CHROMOSOMAL ABERRATION BY MORPHOLOGY

Tumor Type	No. of Cases	Tumor Type	No. of Cases
Neuroglial Tumors	711	***Peripheral Primitive Neuroectodermal Tumors***	***213***
Glioma, NOS	44		
Astrocytoma	511	Unclassified	32
Oligodendroglioma	46	Ewing sarcoma	147
Ependymoma	66	Askin tumor	11
Choroid plexus tumors	14	Peripheral neuroepithelioma	20
Various	30	Various	3
Neuronal Tumors	***427***	***Melanocytic Tumors***	***277***
Retinoblastoma	82	Benign nevus	6
Neuroblastoma	182	Dysplastic nevus	15
PNET/medulloblastoma	153	Malignant melanoma	256
Various	10		

Figures obtained in February 1999 from the database for Mitelman's *Catalog of Chromosome Aberrations in Cancer*.

reported to date closely resembles that in low-grade astrocytomas, with predominance of loss of X, Y, and 22, and trisomy 7. The karyotypic spectrum in ependymoma is slightly different: loss of chromosome 22 is the single most frequent change, occurring in one third of the 70 cases that have been reported. The frequent finding of monosmy 22, and the fact that the incidence of ependymomas is higher among patients with neurofibromatosis type 2, makes the tumor suppressor gene *NF2* a possible target suppressor gene. Mutations of *NF2* seem very rare, however, in sporadic ependymomas.

Medulloblastomas, or primitive neuroectodermal neoplasms (PNET), are the most frequent CNS tumors in children. The cytogenetic hallmark of these neoplasms is the presence of an i(17q), which is found in one third of the cases. Other recurrent, but less frequent (about 10%), aberrations include loss of a sex chromosome, del(6q), +7, -10, -17, -22, and double-minute chromosomes. FISH and molecular studies have revealed that most i(17q) actually are isodicentric chromosomes (idic), with a breakpoint in proximal 17p that lies within the critical region for the Smith-Magenis syndrome (Scheurlen et al. 1997). The mechanisms behind the frequent idic(17) formation or the critical outcome of loss of distal 17p remain unknown. However, molecular analyses have shown that *TP53* probably is not the target gene.

Meningioma was the first solid tumor type in which a characteristic chromosome abnormality — monosomy 22 — was detected. At present, more than 500 cases with clonal aberrations have been reported, providing a solid basis for cytogenetic correlations with molecular, histopathologic, and clinical parameters. Cytogenetic and molecular genetic data strongly indicate that loss of chromosome 22 material is a primary event in the development of meningiomas and that more than one gene may be involved. Cytogenetically, loss of one copy of chromosome 22 is seen in three fourths of the cases, often as the sole anomaly. Furthermore, several cases with deletions, typically of the distal half, of 22q have been described, and molecular analyses have pinpointed a commonly deleted region to 22q12-qter. The target gene for many of these rearrangements is *NF2*, located in 22q12, which has been shown to be deleted and/or mutated in the majority of cases. Still, molecular data indicate that some 40% of the cases do not have functional inactivation of the *NF2* gene, suggesting that alternative mechanisms are operative in a substantial subset of meningiomas. Other chromosomes/chromosome arms that have been implicated by cytogenetic or CGH results are X, 1p, 8, 11p, and 14, all of which are nonrandomly lost.

Correlations with clinical parameters have shown that high-grade lesions have more complex karyotypes than do low-grade meningiomas, that abnormal karyotypes are more frequent among suprasellar meningiomas than at other sites, and that recurrent or invasive meningiomas have more complex karyotypes than the benign ones. In particular, loss of 1p and hyperdiploidy seem to be associated with high risk of recurrence (Steudel et al. 1996).

Acoustic neuromas also have been linked to a locus on chromosome 22. This tumor type occurs in both a sporadic form in which tumors are unilateral and a familial form characterized by bilateral tumors. The familial form also is referred to as bilateral acoustic neurofibromatosis or neurofibromatosis type 2 (NF-2). lnitially, NF-2 was thought to be a variant of von Recklinghausen's disease, which also is called neurofibromatosis type 1 (NF-1). However, studies of families with NF-2 have shown that acoustic neuromas are genetically more closely linked to meningiomas than to NF-1. There are two sets of evidence for this distinction. First, some members of the NF-2 pedigrees also develop meningioma, but individuals with NF-1 do not develop acoustic neuromas. Second, consistent loss of alleles from the region of chromosome 22 affected in the meningioma tumors also has been documented in tumor DNA from patients with bilateral acoustic neuroma, even though cytogenetic evidence linking chromosome 22 to acoustic neuromas is sparse.

Retinoblastoma, a rare malignancy that arises from cells of the embryonal neural retina, develops only in infants and young children. In most cases, it is a sporadic disorder, but approximately one third of the tumors are familial, with the predisposition being inherited as an autosomal-dominant trait. Based on clinical differences between sporadic and inherited forms, Knudson (1971) postulated that retinoblastoma develops as a result of two separate mutations of the same gene. In sporadic tumors, both occur somatically, whereas in heritable retinoblastoma the first mutation is present in the germ line and the second is somatic. The identification of constitutional deletions of 13q14 in a subset of retinoblastoma patients led to the molecular identification of the *RB1* gene, which has come to serve as the prototype for tumor suppressor genes. Subsequent molecular analyses have shown that most retinoblastomas have functionally inactivated *RB1* genes, and that introduction of wild-type *RB1* into retinoblastoma cell lines suppresses tumorigenicity.

At the cytogenetic level, 13q14 aberrations are relatively infrequent. Only 10% of the almost 100 cases with clonal aberrations that have been reported have had deletions or unbalanced translocations leading to loss of this band, and an additional 10% have shown monosomy 13. Instead, i(6p), mostly seen as a supernumerary isochromosome, and gain of 1q material are the most common acquired chromosome changes, each occurring in about one-third of the cases. Nothing is known about the significance of these secondary changes.

Roughly 100 uveal melanomas with chromosome aberrations have been reported. The cytogenetic features are quite different from those detected in malignant melanomas of the skin (see below). They usually are near diploid, and the most frequent alterations are monosomy for chromosome 3 and gain of extra copies, usually through isochromosome formation, of the long arm of chromosome 8. These two imbalances tend to be present simultaneously, and are found in approximately 50% of the cases. Also, chromosome 6 is commonly involved (45% of the cases), with deletions of 6q occurring in one third of the cases. Nothing is known about the essential molecular consequences of -3, i(8q), or del(6q). From a clinical point of view, however, it could be noted that -3 and gain of 8q have been associated with ciliary body involvement and poor clinical outcome, suggesting that cytogenetic analysis might add valuable prognostic information (Sisley et al. 1997).

Neuroblastoma arises from postganglionic sympathetic neurons, often in one of the adrenal glands. Although neuroblastoma

cells are notoriously difficult to culture *in vitro*, close to 200 cases with clonal karyotypic changes have been reported. The modal chromosome number has ranged from hypodiploid to hypertetraploid. The most frequent abnormality, found in approximately 75% of the cases, is partial 1p monosomy, with frequent deletion of the region between 1p32 and 1p36. LOH studies indicate a lower — 30% to 40% — frequency of 1p deletions, a discrepancy that may be from selective outgrowth *in vitro* of del(1p) carrying cells. The molecular studies have identified one commonly deleted segment in band 1p36 and, possibly, one additional target for deletions in 1p32. Several candidate tumor suppressor genes for distal 1p have been tested, but so far without any success.

The other cytogenetic hallmark of neuroblastoma is the presence of double minute (dmin) chromosomes or, less frequently, HSRs, found in one third of short-term cultured cells and in even higher frequencies among established cell lines. At the molecular level, the dmin and HSR correspond to large-scale amplification of the *MYCN* oncogene, and it seems as if the gene amplification occurs early in tumorigenesis; if absent at the time of diagnosis, it does not develop during tumor progression.

Cytogenetic deletions of 1p and dmin/HSR/*MYCN* amplification, which typically occur in tumors with near-diploid karyotypes, are more frequent among older children and have been associated with poor clinical outcome, whereas karyotypes with near-triploid chromosome counts usually lack these aberrations and are associated with young age of onset (<1 year) and good prognosis. Using molecular techniques, however, the prognostic significance of 1p deletions has been questioned (Schwab et al. 1996).

CGH analyses have identified frequently lost or gained chromosome segments that have been detected at far lower frequencies after banding analysis of cultured cells. The most striking examples

are gain of 17q, present in 70% to 85% of the cases, and loss of chromosome 14, present in almost one third of the cases.

TUMORS OF THE HEAD AND NECK REGION

Part of "34 - Solid Tumor Cytogenetics"

SCCs account for 90% to 95% of all malignancies in the upper aerodigestive tract. Some 200 cases with clonal abnormalities have been reported after short-term culturing. The majority of these have had fairly complex karyotypes, typically showing a large number of numerical and unbalanced structural rearrangements. The most common imbalances detected at banding analysis are loss of 2q34-qter, 3p, 4p, 4q28-qter, 8p, 9p13-pter, 10p, chromosomes 13 and 14, 15p, 17q23-qter, 18q21-qter, 21p, 22p, and the Y chromosome, and gain of 3q, chromosome 7, 8q, and 11q13. The still relatively few cases that have been analyzed by CGH have shown a similar distribution of gains and losses, albeit at slightly different frequencies. At the cytogenetic level, more than half of the breakpoints involved in structural rearrangements affect centromeres or the near-centromeric regions of the chromosomes. The only euchromatic segments that are repeatedly involved are 1p11–13 and 11q13, the latter often in the form of HSR. FISH studies have shown that the breakpoints in proximal 1p are widely scattered, usually resulting in loss of material from distal 1p, whereas HSR affecting 11q13 (See Fig. 34.1) always represents amplification of the cell cycle regulating *CCND1* gene (Jin et al. 1998a, Jin et al. 1998b). Molecular studies have shown that other genes are also included in these amplicons, but that only *CCND1* and the neighboring *EMS1* gene are consistently overexpressed as a result of the amplification. Frequent inactivation of *CDKN2A* correlates well with the frequent loss of 9p21. The molecular significance of other recurrent changes in SCC is largely unknown.

A subset of SCC have shown fairly simple karyotypes, often

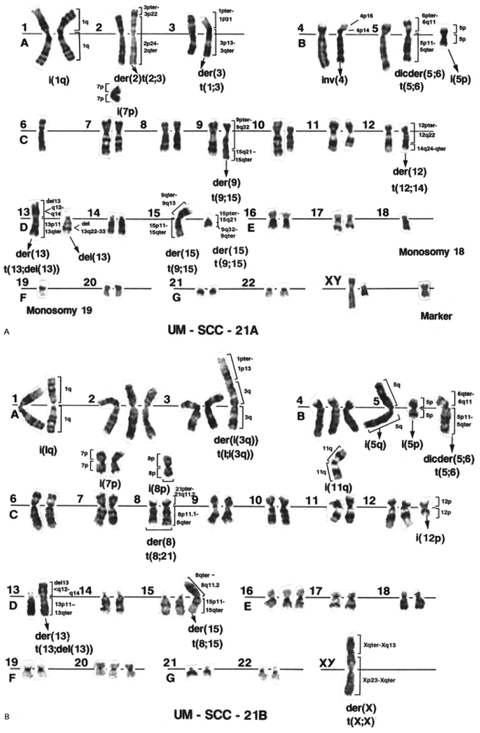


FIGURE 34.6. G-banded karyotypes from head and neck cell lines UM-SCC-21A (**A**) from the primary site and (**B**) UM-SCC-21B from a submental metastasis of a baso-squamous-cell carcinoma showing both common clonal markers and new chromosome changes that arose in each tumor after the two populations separated. The metastatic tumor was removed from the patient 14 months after the primary tumor was excised. Of note, there are numerous changes that arose in the primary tumor that are not present in the cell line from the metastasis demonstrating that the metastatic population had left the primary site long before the surgery on the primary site occurred. *(continued)*

with multiple, unrelated clones, with only a few numerical and/or structural rearrangements. The significance, if any, of these aberrations is questionable. First, established cell lines invariably display more complex changes; second, there are no recurrent rearrangements among the karyotypes with simple structural rearrangements and; third, similar changes have been detected also in nonneoplastic mucosa from elderly individuals (Jin et al. 1997a). Thus, it may well be that most of these changes detected in short-term cultured tumor specimens represent accumulation of genetic damage in stromal or nonneoplastic epithelial cells.

The only thorough evaluation of the clinical significance of cytogenetic findings in SCC showed that complex karyotype in general, and 11q13 rearrangements in particular, were associated with poor clinical outcome (Åkervall et al., 1995). These findings are in line with molecular genetic data indicating that amplification and/or overexpression of *CCND1* confer a poor prognosis. An analysis of a smaller number of stage 3 head and neck cancer patients demonstrated that loss of 18q was significantly associated with poor survival (Pearlstein et al. 1998).

In some cases, tumor progression may be accompanied by chromosomal gains that can obscure the loss of whole chromosomes. Fig. 34.6 presents a set of karyotypes from cultures of a primary tumor and a metastatic tumor in the same patient. UM-SCC-21A is derived from a baso-squamous carcinoma of the skin that was placed in culture more than a year before a metastasis became palpable in the soft tissues of the neck. UM-SCC-21B was derived from that metastatic tumor. The karyotypes show that both tumors are derived from the same clone. Note the presence in both tumors of i(1q), i(5p), t(5;6), i(7p), and t(13;del(13)). However, several changes from the primary tumor are not present in the metastasis and other new changes not found in the primary have developed in the metastatic tumor. Of particular interest are monosomy 6, 18, and 19 in the primary tumor, all of which are represented by two copies in the metastasis. This suggests that these losses occurred with further progression at the primary site after the metastatic population had diverged.

Salivary Gland Neoplasms

The non-squamous cell carcinomas of the head and neck region that arise predominantly from major and minor salivary glands, which include adenocarcinoma, acinic-cell carcinoma, mucoepidermoid carcinoma, adenoid cystic carcinoma, and carcinoma ex pleomorphic adenoma, are less well characterized, with less than 25 informative cases per tumor type. The karyotypes of these malignancies usually are less complex than those typically encountered in SCC. Adenocarcinomas and acinic cell carcinomas often show deletions of 6q and trisomies for chromosomes 7 or 8, in mucoepidermoid carcinomas a recurrent t(11;19)(q14-21;p12) has been described, adenoid cystic carcinomas are characterized by a recurrent t(6;9)(q21-24;p22-24) or del(6q), and carcinomas arising from pleomorphic adenomas usually have structural rearrangements resembling those in pleomorphic adenomas (see below) together with other structural or numerical abnormalities, such as polyploidization, indicating that the malignant transformation often is paralleled by cytogenetically visible clonal evolution.

More than 250 pleomorphic adenomas of the salivary glands with clonal chromosome aberrations have been reported. Three cytogenetic subgroups have been recognized: one third display 8q12 rearrangements, usually as a t(3;8)(p21;q12), 15% have different types of 12q13-15 rearrangements, and the remaining cases have apparently normal karyotypes or simple numerical changes, such as -Y and +8. The molecular consequences of the t(3;8)(p21;q12) recently were clarified. The translocation results in an exchange of promoter sequences between the zinc finger gene *PLAG1* in 8q12 and the constitutively expressed **B**-catenin *CTNNB1* gene in 3p21, which in turn leads to activation of *PLAG1* (Kas et al. 1997). Activation of *PLAG1* through other similar mechanisms also has been demonstrated for some of the variant translocations that are known to occur. The 12q13-15 rearrangements have been shown to involve the *HMGIC* gene, which is commonly rearranged also in a variety of benign mesenchymal tumors (see below). In line with this finding is the observation that adenomas with 12q13-15 rearrangements have a more pronounced stroma component *in vivo* and that cell cultures from these lesions tend to have a more fibroblastlike appearance.

Of the different subtypes of malignant thyroid tumors, only papillary and follicular carcinomas have been investigated in any detail. Still, there are only about 30 cases of each tumor type in the literature. The papillary carcinomas often have a quite simple karyotype, with rearrangements of 10q as the only characteristic feature, found in one third of the cases. The most frequent structural aberration is a paracentric inversion inv(10)(q11q21), which at the DNA level results in a fusion transcript consisting of parts of the *RET* protooncogene in 10q11 and parts of either *PTC1* or *PTC3* in 10q21. Other less frequent mechanisms behind *RET* activation in papillary carcinomas include fusion with *PTC2* through a translocation with chromosome 17.

The follicular carcinomas have had karyotypes of varying complexity, although most have had a chromosome count in the diploid range. Structural rearrangements leading to loss of 3p material is the most common finding. This is in line with molecular studies, which have identified a commonly deleted region in 3p25-pter. A few cases have had karyotypes with simple numerical (+7) or structural rearrangements [del(13q)], indistinguishable from those seen in benign thyroid lesions.

Consistent patterns of chromosome changes have been detected also in benign thyroid adenomas and hyperplasias (Belge et al., 1998). From cytogenetic studies of more than 200 such cases, two main chromosomal subgroups have emerged: one with rearrangements of 19q, accounting for 20% of the cases with clonal changes, and one with a combination of trisomies, almost always including chromosome 7, seen in one third of the cases. Less frequent cytogenetic subgroups are characterized by deletions of 2p or 13q.

Parathyroid adenomas are often characterized by an inv(11) that translocates the parathyroid hormone receptor gene at 11p15 to the region of 11q13. This rearrangement led to the cloning of a gene on the long arm that resides near the BCL1 breakpoint. This gene was originally called *PRAD1* (parathyroid adenoma gene 1). However, when its sequence was elucidated, it was found to be homologous to the cyclin gene involved in cell-cycle progression in xenopus oocytes (Motokura et al. 1991).

Later it was identified as *Cyclin D1*, which is overexpressed in parathyroid tumors with the inv(11). This provides another example in which translocation of an active gene, in this case *PTH*, causes unscheduled expression of a cell regulatory gene (*cyclin D1/CCND1*) as a result of chromosome rearrangement in neoplasia.

TUMORS OF THE BREAST

Part of "34 - Solid Tumor Cytogenetics"

Clonal chromosomal changes have been documented in over 500 breast carcinomas. A wide spectrum of karyotypic changes have been detected among them, with some tumors showing a single numerical or structural alteration and some tumors having highly complex karyotypes with numerous unidentifiable marker chromosomes. Part of these differences may be a result of the analysis of different subsets of breast carcinoma. Initially, mainly pleural effusions from highly advanced-stage carcinomas were used, but with the development of improved cell culture techniques, it also became possible to analyze less aggressive, primary lesions. Another factor influencing the cytogenetic result is the culture technique used; direct harvesting is more likely to yield complex karyotypes, but often fails to provide any information, whereas short-term culturing tends to favor the outgrowth of cells with pseudodiploid or near-diploid karyotypes (Steinarsdóttir et al. 1995).

In spite of these technical problems, a nonrandom pattern of chromosomal aberrations has been detected. No recurrent balanced rearrangement has been detected, but several imbalances are frequent: loss of 1p, 3p, 6q, 8p, 11q, 13q, 16q, and 17p, and gain of 1q and 8q. Some changes have been suggested, on the basis of their appearance both as sole anomalies and together with other aberrations, to be of particular importance in breast cancer development: del(1)(q11-12), del(1)(q42), i(1)(q10), der(1;16)(q10;p10), del(3)(p12-13p14-21), del(6)(q21-22), +7, +8, +12, +18, and +20. However, none of these aberrations is specific for breast carcinoma, and most of them have been identified in benign breast lesions as well (see below). The understanding of the role of these imbalances in breast cancer development is further hampered by the fact that they involve large chromosomal segments, making it difficult to identify the essential molecular genetic consequences.

Several cytogenetic-clinicopathologic correlations have been reported for breast carcinomas, but few have been confirmed in independent studies. It seems, however, as if loss of 16q, as detected by cytogenetic, FISH, or molecular techniques, is associated with node-negative tumors and good prognosis (Hansen et al. 1998; Adeyinka et al. 1999).

Chromosomal aberrations have been reported in almost 100 benign breast lesions (fibroadenomas, phyllodes tumors, adenolipomas, hamartomas, nodular fasciitis, and fibrocystic disease), as well as in prophylactic mastectomies from women with familial predisposition to breast cancer development (Lundin and Mertens, 1998). No disease-specific aberration has been detected among these lesions, almost all of which have displayed a near-diploid chromosome count. Many of the recurrent changes that have been detected, e.g., del(1)(q11-12), der(1;16)(q10;p10), del(3)(p12-13p14-21), del(6q), +7, +18, and +20, coincide with those that have been detected among the invasive breast carcinomas, suggesting that they are not sufficient for malignant transformation. Furthermore, the finding of shared abnormalities in benign and malignant breast tumors and the fact that breast carcinomas often have simple karyotypes imply that cytogenetics is of limited value in the differential diagnosis. One exception would be the phyllodes tumors, where malignant tumors consistently have shown complex karyotypes and the benign variants have had simple karyotypes, which is in line with recent CGH data (Lu et al. 1997; Lundin and Mertens 1998).

The finding of multiple, unrelated cytogenetically abnormal clones is a common phenomenon (up to 50% of the cases after short-term culturing) in both benign and malignant breast lesions. In contrast to other tumors frequently displaying unrelated clones, i.e., skin tumors and head and neck carcinomas, there is a regularity to the pattern of aberrations observed. The significance of such polyclonality is still unclear, but strongly implies that a large proportion of breast tumors are of multicellular origin (Teixeira et al. 1996) or that an early clonal change, which may lack a visible rearrangement, causes genetic instability that drives rapid evolution of clonal diversity. However, like for many other tumors in which multiple clones are observed, most are not long-term survivors and the progressive tumor is a stemline clone. A definite answer to this question of polyclonality will, however, require combined molecular and cytogenetic approaches as well as longitudinal studies with early and late samples of the same tumors.

TUMORS OF THE LUNG

Part of "34 - Solid Tumor Cytogenetics"

The brunt of the cytogenetic information on lung carcinomas stems from the analysis of adenocarcinomas and squamous cell carcinomas. For each of these histopathologic subtypes, approximately 150 cases with clonal aberrations have been reported. The other main categories are large-cell (40 cases) and small-cell lung carcinomas (SCLC, 50 cases), whereas less than 20 informative adenosquamous-cell carcinomas or carcinoids have been published. No aberration specific to any of the different subtypes has so far been described. Generally, all types of lung carcinoma have a highly complex karyotype, with a near-triploid chromosome count and massive numerical and structural changes. The perhaps most conspicuous finding is the frequent loss of 3p material, which initially was thought to be particular for SCLC. Indeed, combined cytogenetic and molecular genetic results indicate that 3p deletions occur in 90% to 95% of all SCLC, but later studies have shown that the majority of non-SCLC also display this aberration. The 3p deletions, which according to chromosome banding analyses have a maximum frequency at bands 3p21-22, are likely to result in the inactivation of a tumor suppressor gene. Despite intensive efforts to identify the putative target gene, no convincing candidate has been identified to date.

Other frequent (at least 25% of the cases) chromosomal imbalances among lung carcinomas include loss of chromosomes X, Y, 4, 9 (in particular 9p), 13, 15, 18, 22, and chromosome arms 1p (1p22-1p36), 6q, 8p, and 17p, and gain of chromosome 7 and chromosome arm 1q. It is still too early to conclude whether any of these imbalances may be particularly characteristic for any of the different histologic subtypes.

Almost 100 mesotheliomas, the major malignancy of pleural origin, with abnormal karyotypes have been described. According to Hagemeijer et al. (1990), who presented the hitherto largest series comprising 30 informative tumors, the cytogenetic findings suggested a dichotomization of the tumors into those with hypodiploid/hypotetraploid chromosome numbers and frequent -4, 9p-, and -22, and those with hyperdiploid karyotypes and frequent 3p-, +5, +7, and +20. No significant association was, however, found between cytogenetic pattern and clinical outcome or histologic subtype of the tumor.

Pulmonary chondroid hamartoma of the lung is a benign neoplasm composed of cartilage, adipose tissue, undifferentiated mesenchymal cells, and epithelium-lined, cleftlike spaces. Almost 200 cases with clonal aberrations have been reported, most of which were part of a large single-institutional series encompassing 191 tumors (Kazmierczak et al. 1999). In that cohort, 70% of the short-term cultured samples showed clonal aberrations, with 21 tumors showing 6p21 rearrangements and 95 tumors showing 12q14-15 rearrangements. The most frequent recurrent translocations were t(12;14)(q15;q24) and t(6;14)(p21;q24). FISH and molecular genetic studies have revealed that the target genes in 6p21 and 12q15 are *HMGIY* and *HMGIC*, respectively, which are rearranged in many other benign mesenchymal tumors (see below). The target gene in 14q24 remains to be identified, but is likely to be the recombination repair gene *RAD51B*, which was shown to be involved in uterine leiomyomas with t(12;14) (Schoenmakers et al. 1999).

TUMORS OF THE GASTROINTESTINAL TRACT

Part of "34 - Solid Tumor Cytogenetics"

The majority of the more than 100 reported ventricular carcinomas, almost all of which were adenocarcinomas, have had highly complex karyotypes. Different research groups have reached slightly different results, and the karyotypes often have included numerous unclassifiable marker chromosomes, making it difficult to evaluate the significance of individual chromosome aberrations. Nevertheless, frequently reported rearrangements include gain of chromosomes X, 8, and 9, loss of the Y, deletions of 3p, 6q, and 7q, isochromosomes for the long and short arms of chromosomes 8 and 5, respectively, and various rearrangements involving 11p.

Of the more than 100 pancreatic malignancies that have been reported to date, the vast majority have been carcinomas of the exocrine pancreas. A review of the cytogenetic data on this tumor type was presented recently by Gorunova et al. (1998). The karyotypes usually are complex with a chromosome number in the triploid-tetraploid range (See Fig. 34.4). Several recurrent changes, all of them leading to imbalances, have been identified: del(1)(q11-12), del(1)(q21), i(1)(q10), del(3)(p11), i(3)(q10), del(4)(q21), del(4)(q25), i(5)(p10), del(6)(q15), i(8)(q10), del(9)(p13), del(11)(p13), dup(11)(q13q23), i(12)(p10), der(13;15)(q10;q10), del(18)(q12), del(18)(q21), and i(19)(q10). Common numerical aberrations include -Y, +2, +7, -18, +20, and -21. These findings are in agreement with CGH and molecular studies, showing frequent loss of 1p, 3p, 6q, 9p, 13q, 15q, 18q, 19p and gain of 5p, 7q, 8q, 11q, 12p, 19q. Several potential target genes for commonly deleted (*CDKN2A* in 9p, *SMAD4* in 18q) or amplified (*CCND1* in 11q, *KRAS2* in 12p, and *AKT2* in 19q) regions have been identified by molecular analyses.

A striking feature of pancreatic carcinomas is the pronounced intratumor heterogeneity seen as multiple related or unrelated clones. In the series of 25 tumors reported by Gorunova et al. (1998), 19 tumors displayed from 2 to 58 clones. As in head and neck carcinomas (see above), unrelated clones usually are near-diploid and sometimes occur together with complex clones. The significance of this cytogenetic polyclonality is unknown.

Very few primary liver tumors have been subjected to chromosome-banding analysis. Hepatoblastomas, the most frequent liver tumor in childhood, have, however, a nonrandom pattern of cytogenetic changes with trisomy 2, or gain of 2q, i(8q), and gain of chromosome 20 as the most common features.

The cytogenetic database on colorectal carcinomas comprises more than 300 cases, mostly adenocarcinomas, with clonal chromosome aberrations. Approximately two thirds of the tumors have displayed a near-diploid chromosome count, often with relatively few numerical and/or structural rearrangements. The remaining one third have had more complex karyotypes, sometimes with more than 100 chromosomes. The most frequent numerical aberrations are gain of chromosome 7 and loss of the Y and chromosome 18 (each detected in 30% to 35% of the cases) followed by trisomy for chromosomes 13 and 20 and monosomy for chromosomes 14, 17, and 22 (each found in 15% to 20%). The most common structural aberrations, all of which lead to loss and/or gain of genetic material, are isochromosomes for the long arms of chromosomes 1, 8, 13, and 17, and deletions of 1p and 17p.

Hereditary non polyposis colon cancer (HNPCC) also known as Lynch syndrome is an inherited predisposition to colorectal cancer that is characterized by microsatellite instability as a result of DNA replication errors (Aaltonen et al. 1993). In initial studies, two chromosome regions were discovered by kindred analysis of inherited linked marker haplotypes. The first region was on chromosome 2p16. In families in which a positive LOD score was not found for this locus, subsequent analysis revealed a locus at 3p21-23. These regions, which are frequently lost in karyotypes from colon cancers as well as other tumors, are the sites of *hMSH2* on 2p (Kohonen-Corish et al. 1995) and *hMLH1* on 3p (Bronner et al. 1994). Hemizygous deletion eliminates the wild-type copy of the gene leaving only the mutated form of the enzyme in the tumor cells, which then have increased DNA replication errors. Mutations of these genes are common in Lynch syndrome patients. Further analysis has led to the identification of other mismatch repair genes. Correction of mismatch repair has been demonstrated *in vitro* by microcell-mediated transfer of a normal chromosome 3 to a human colorectal carcinoma cell line that contains a defect in *hMSH2*. This experiment demonstrates that the defective mismatch genes are responsible for the DNA repair errors and microsatellite instability observed in HNPCC cells (Koi et al. 1994).

Bardi et al. (1995) have reported a correlation between karyotype and tumor grade and site. Structural chromosome aberrations were more common among poorly differentiated carcinomas, whereas well and moderately differentiated carcinomas more often had numerical changes only. Abnormal karyotypes in

general, and complex ones in particular, were more common in rectal tumors than in tumors of more proximal location. Furthermore, it has been suggested that patients with complex tumor karyotypes have a worse clinical outcome than other patients.

In contrast to most other malignancies, carcinomas of the large bowel usually develop through histologically defined premalignant stages, the so-called adenoma-carcinoma sequence. Thus, the study of benign colorectal neoplasms is potentially of utmost importance for our understanding of colorectal tumorigenesis. More than 100 adenomas with clonal abnormalities have been reported. In general, these tumors have had fairly simple karyotypes, but there is a clear tendency for increasing karyotypic complexity with increasing degree of dysplasia. It also should be noted that some cases have displayed highly aneuploid karyotypes. The spectrum of chromosomal gains and losses is very similar to that in colorectal carcinomas, but with gains dominating over losses. The most frequent gains are +7 (50% of the cases), +13 (30%), and +20 (20%), and -18 is the most common monosomy (15%). The breakpoints involved in structural rearrangements cluster to chromosome 1, with loss of distal 1p as the most frequent aberration. Deletions of 1p has in many cases been the sole cytogenetic aberration, indicating that loss of a, so far, unknown gene in distal 1p is of primary importance in colorectal tumorigenesis. Guided by an interstitial deletion of 5q in a patient with familial adenomatous polyposis (FAP, also called familial polyposis coli or FPC), an autosomal-dominant condition in which affected individuals develop hundreds to thousands of colonic polyps led to the discovery of the APC gene at 5q21-q22. The adenomatous polyps are a premalignant condition and nearly all affected patients develop colon cancer by or in their 5th decade. APC mutations are present in at least 95% of FAP families and screening for the mutations can exclude nonaffected family members (Laken et al. 1999).

Many of the chromosome segments that are often involved in imbalances in adenomas and carcinomas are known to harbor genes that from molecular genetic studies have been implicated in colorectal carcinogenesis (Fearon and Vogelstein 1990). Whereas the cytogenetic data may be interpreted to confirm the previous conclusion that it is the net accumulation of genetic changes, rather than the order in which they appear, that is crucial for determining the tumor phenotype, they also show that the molecular picture of colorectal tumor development is far from complete.

TUMORS OF THE GENITOURINARY TRACT

Part of "34 - Solid Tumor Cytogenetics"

Cytogenetic investigations of RCC have resulted in more than 600 reported cases with clonal chromosome abnormalities. Several histopathologic subtypes are recognized, which may be broadly categorized into papillary and nonpapillary RCC. Among the latter, which constitute about 85% of all RCC, the clear cell type is the by far best characterized by chromosome banding techniques. The single most common change (in some series up to 90% of the cases) in this subtype is loss of 3p material. Often this occurs through terminal deletions at bands 3p12-13, but other chromosomal mechanisms, such as interstitial deletions, loss of the entire chromosome, or unbalanced translocations, in particular with chromosome arm 5q, also have been described. Although the cytogenetic data indicate a maximum deletion frequency at band 3p21, molecular investigations suggest that the salient outcome of 3p deletions concerns a more distal locus, i.e., the *VHL* gene in 3p25. This gene, which is causally related to the inherited predisposition to kidney cancer seen in the von Hippel-Lindau disease, is functionally inactivated (lost, mutated, or methylated) in close to 70% of all RCC. Of course, one cannot exclude the possibility that other, more proximal tumor suppressor genes that are deleted concomitantly with the *VHL* gene are of pathogenetic importance, as well. Indeed, loss of heterozygosity studies implicate three distinct, commonly deleted regions: one in 3p13-14, one in 3p21, and one in 3p25-26. Furthermore, similar 3p deletions are common in several other tumor types for which the *VHL* gene is not involved, suggesting that this is a region with several important tumor genes.

Other frequent (more than 15% of the cases) rearrangements in clear-cell RCC include gain of 5q, deletions of 6q, trisomy 7, and loss of chromosomes Y, 8, 9, and 14. Most of these changes probably represent secondary, progression-related alterations, although so far no clear-cut evidence has been provided that the presence or absence of any of them should correlate with clinical outcome. Trisomy 7 presents particular interpretational problems in this context. Not only is gain of one or more copies of chromosome 7 commonly found in the same clone as 3p-, but it can also occur as the sole change, either alone or together with clones carrying 3p- and/or other changes. Furthermore, it has been clearly established that trisomy 7-carrying cells may be present also in nonneoplastic kidney tissue. Thus, the impact of +7 on renal-cell carcinogenesis is presently elusive.

Chromophilic RCC, most of which have a papillary growth pattern, has a distinct cytogenetic profile. The adenomas and low-grade variants usually show only a few numerical changes of which +7, +17 and -Y, are by far most common. In contrast, the more malignant cases display additional trisomies for chromosomes 12, 16, and 20, and loss of the supernumerary chromosome 17 or deletions of 17p. Another subset of malignant papillary/chromophilic RCC is characterized by a pathognomonic t(X;1)(p11;q21), resulting in the fusion of the transcription factor gene *TFE3* on the X-chromosome with the *PRCC* gene on chromosome 1.

Nonrandom karyotypic patterns also have been detected in oncocytomas, one third of which have displayed -Y and -1, chromophobic-cell carcinomas, which usually have hypodiploid chromosome numbers, and collecting duct carcinomas, which are characterized by loss of chromosomes 1, 6, 14, 15, and 22.

Wilms' tumor, the most common renal tumor in childhood, is cytogenetically totally different from adult forms of kidney cancer. Although the inherited forms, and their association with congenital malformations and constitutional mutations in the *WT1* (11p13) gene and *WT2* region (11p15), have gained much attention, more than 95% of the tumors are believed to be sporadic in nature. The karyotypic pattern among the 175 reported cases is clearly nonrandom. Most tumors have had a near-diploid chromosome count. Numerical aberrations mainly involve gain of chromosomes, trisomy 12 being particularly frequent followed by +8, +6, +7, +13, +20, and +17; several of these trisomies

have been detected as the sole karyotypic deviation. Also, structural rearrangements typically result in imbalances, the most common being gain of 1q and loss of 11p. One particular aberration, der(16)t(1;16)(q10-21;q10-13), has been detected in more than 10 cases. The same aberration has been described in many other neoplasms, notably Ewing sarcoma and breast cancer, and probably represents a secondary change. A few cases with triploid-tetraploid karyotypes have been reported, and these tumors tend to have an anaplastic morphology. The same histologic subset, which carries a poor prognosis, also is characterized by *TP53* mutations, which are rare in other subtypes of Wilms' tumor.

A less frequent pediatric kidney tumor is mesoblastic nephroma. Cytogenetically, these low-malignant tumors are characterized by simple numerical changes, the most consistent being +11 followed by gain of chromosomes 8, 10, 17, and 20. The numerical changes, as well as the morphology of the tumors, are essentially identical to those in infantile fibrosarcomas, and it was recently shown that the two tumor types share also a t(12;15)(p13;q25)-associated *ETV6/NTRK3* fusion, which presumably constitutes the primary, tumorigenic event (Knezevich et al. 1998).

No specific chromosomal abnormality has been detected among transitional cell carcinomas of the bladder. Some of the more than 100 reported cases have had a simple karyotype, with only a few numerical and/or structural rearrangements, whereas others have had highly complex changes. Frequent aberrations include various rearrangements of chromosome 1, del(3p), i(5p), trisomy 7, monosomy for or deletion of chromosome 9, and del(10q). Both cytogenetic and molecular genetic data have implicated deletions of 9p and 9q as important in the earliest phases of tumorigenesis, whereas 17p deletions are associated with tumor progression. Cytogenetic studies also have convincingly demonstrated that multifocal uroepithelial tumors usually have a monoclonal origin.

Prostate cancer cells have turned out to be very difficult to maintain in culture, which in part may explain the relative lack of cytogenetic data on this common tumor type. Furthermore, about 100 of the 180 reported cases with acquired chromosome changes have had -Y, +7, and/or a “normal” tetraploid — 92,XXYY — karyotype as the sole anomalies, all of which are of doubtful significance in prostate carcinogenesis. It could be noted, though, that -Y and +7 are common also in samples from benign prostatic hyperplasia. Among the remaining carcinomas, karyotypes of varying complexity have been identified, and it seems as if the structural rearrangements in these tumors cluster to chromosomes 1, 7, 8, and 10. The most consistent outcomes of such rearrangements are deletions of 7q, often in band 7q22, 8p, and 10q. In one study containing a mere 15 cases with clonal aberrations and 42 with normal karyotypes, the former were found to have a significantly shortened survival (Lundgren et al. 1992).

Several series of cytogenetically analyzed ovarian carcinomas have been published, and the literature now contains more than 400 cases, the majority of which were classified as adenocarcinomas, with clonal chromosome abnormalities. Most tumors have had fairly complex karyotypes, with near-triploid or hypodiploid stemlines. The most frequent numerical changes are loss of X, 4, 8, 13, 14, 15, 17, and 22, and gain of chromosome 12, each detected in at least 15% of the cases. Structural rearrangements cluster to chromosomes 1, 3, 6, 9, 11, and 19, leading to loss of 1p, 1q, 3p, 6q, 9p, 11p, 19p, and 19q material.

A few carcinomas have karyotypes with only numerical changes or a single structural rearrangement. Such karyotypes have been associated with well-differentiated tumors, whereas poorly differentiated carcinomas have complex karyotypes. Furthermore, complex karyotypes, as well as several nonrandomly rearranged regions, have been correlated with poor clinical outcome (Taetle et al. 1999). Cytogenetic studies also have shown that bilateral ovarian carcinomas usually have identical karyotypes, strongly implicating that the multifocality is from tumor spreading rather than independent tumorigenesis.

Benign ovarian tumors are less well characterized, but a recurrent pattern of numerical changes, with +12 as the most common variant, has been detected among the roughly 40 cases reported. Trisomy 12, often as the sole anomaly, has been detected in slightly more than one half of the cases, and seems to be particularly frequent in the subgroup of sex cord stromal tumors.

More than 600 uterine tumors with clonal chromosome changes have been reported. The by far most extensively analyzed subtype is leiomyoma, a benign myometrial proliferation, for which information on some 325 tumors exist. Abnormal karyotypes are found in approximately one third of the cases, and several karyotypic subgroups have been established (Morton 1998). The most frequent, each accounting for 10% to 20% of the cases, are characterized by deletions of the long arm of chromosome 7 and structural rearrangements of 12q13-15, the most common being a t(12;14)(q14-15;q23-24), respectively. Cytogenetic and molecular genetic investigations indicate that the minimal deleted segment on 7q maps to 7q21-22, but so far no target gene has been identified. The 12q rearrangements usually involve the *HMGIC* gene, which encodes an architectural transcription factor, leading to aberrant expression of the protein. In contrast to other mesenchymal tumors with frequent aberrations of the *HMGIC* gene, such as lipoma (see below), the breakpoints in leiomyomata often occur upstream of the gene. The importance of altered expression in leiomyomas of proteins belonging to the high mobility group (HMG) family is further emphasized by the fact that a subset of the tumors without involvement of *HMGIC* have rearrangements of band 6p21, harboring the closely related *HMGIY* gene.

The other main subtype of benign uterine tumors, endometrial polyps, also display nonrandom chromosomal changes. These overgrowths of endometrial tissue, for which 25 cases with clonal abnormalities have been described, have at least three different characteristic karyotypic changes. Rearrangements of 6p21, often in the form of a t(6;14)(p21;q24), are the most common, followed by rearrangements of 12q13-15, and deletions of 7q. The cytogenetic picture thus is very similar to the one in leimyomas, and molecular studies have shown that in the polyps, too, the 6p and 12q rearrangements affect the *HMGIY* and *HMGIC* genes, respectively. Most interestingly, a recent FISH study revealed that tamoxifen-induced polyps carry *HMGIC* and *HMGIY* rearrangements that are indistinguishable from those in sporadic polyps (Dal Cin et al. 1998).

The cytogenetic information on malignant uterine tumors is less detailed. The majority of the close to 100 reported endometrial carcinomas, most of which are adenocarcinomas, with clonal chromosome aberrations have had a near-diploid chromosome count, with gain of 1q material, trisomy 10, and deletions of 6q as the most frequent changes. The latter imbalance has been suggested to be particularly frequent among serous papillary, endometrioid, and mixed müllerian endometrial carcinomas. Cytogenetic analysis of a small number of cultured endometrial carcinomas revealed several chromosomal locations that appear to be consistent. These include deletions of distal 18q, rearrangements with breakpoints involving the short arm of chromosome 22, and breakpoints involving 12q. Loss of chromosome 4, deletion 9q and 13q, and a breakpoint on 8p may be associated with aggressive tumors. An example of a karyotype from an endometrial carcinoma containing the common changes is shown in Fig. 34.7.

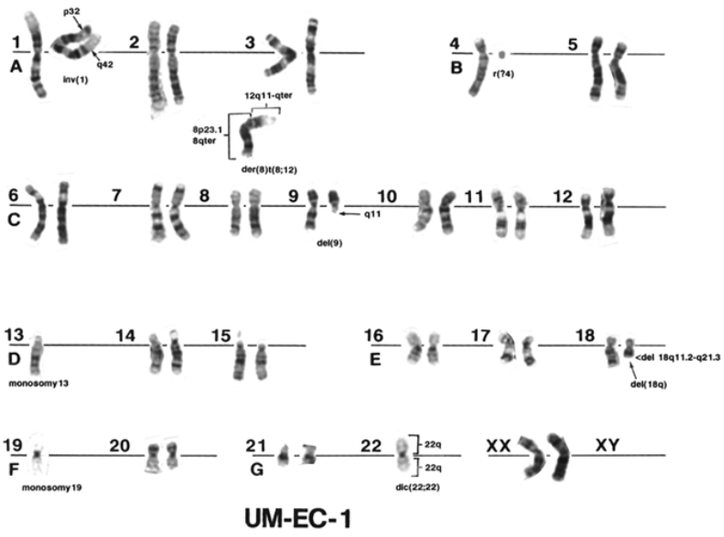


FIGURE 34.7. Karyotype from a hypodiploid endometrial carcinoma demonstrated several of the consistent chromosome changes found in this tumor type. Note that there is monosomy 4, 13, and 19, a pericentric inversion of chromosome 1, del(9q), del(18q), i(22q) and t(8;12) (p23.1;q22). The ring chromosome is most likely derived from the missing chromosome 4. (Reprinted with permission from Grenman SE, Van Dyke DL, Worsham MJ, et al. UM-EC-1, a new hypodiploid human cell line derived from a poorly differentiated endometrial cancer. *Cancer Res* 1988;48:1864–1873.)

Less than 30 malignant uterine tumors showing mesenchymal differentiation have been reported. The leiomyosarcomas have complex karyotypes, usually without any similarities to their benign counterparts, whereas a recurrent t(7;17)(p15-21;q12-21) has been found in stromal sarcomas.

Close to 100 karyotypically abnormal carcinomas of the uterine cervix have been reported. However, many of these karyotypes have been incompletely described, which makes the cytogenetic information an unreliable source for identifying pathogenetically important rearrangements. Recent CGH studies have, however, displayed a nonrandom pattern of imbalances with gain of 3q as the most consistent change, present in two thirds of the invasive cases and one third of *in situ* carcinomas (Kirchhoff et al. 1999). Other frequent (>30% of the cases) changes detected by CGH in invasive carcinomas were gains on Xq, 1q, 5p, 8q, and 15q, and losses on 3p, 4p, 6q, 11q, and 13q.

Squamous-cell carcinomas of the vulva (vulvar carcinoma), are relatively rare tumors, with an incidence of 1 or 2 tumors per 100,000 women. As a result, there are relatively few studies of this tumor type. The etiologic factors for these tumors are not well understood and may differ from squamous cancers that arise at other sites after prolonged exposure to carcinogens. Worsham et al. (1991) studied a consecutive series of six patients and found each tumor to have a unique and complex karyotype. However, there was a surprising degree of similarity in the rearrangements. Five of the six tumors had losses of 3p14-cen, 8pter-p11, 22q13.1-q13.2, and the short arm of the inactive X; the same proportion of tumors had gains of 3q25-qter, band 11q21, and rearrangement breakpoints in 5cen-q12. In this same series, two regions of relative chromosome loss involving 10q and 18q were observed in all four tumors in patients who died from progression of their tumors, whereas the two patients whose tumors did not have these losses were alive and well several years later. In some cases, the loss affected the whole chromosome, but in others there were specific deletions in which the smallest region of loss corresponded to del(10) (q23q25) and del(18) (q22q23). These regions provide targets for additional study as possible sites of genes important in prognosis. An example of a hypotetraploid karyotype from a vulvar carcinoma is presented in Fig. 34.8. Note that many of the regions already mentioned as important in other tumors also are affected in this squamous cancer. For example, i(5p) and loss of 5q as well as loss of 18q are changes in this tumor that are also associated with colon cancer and squamous cancers at other sites such as head and neck.

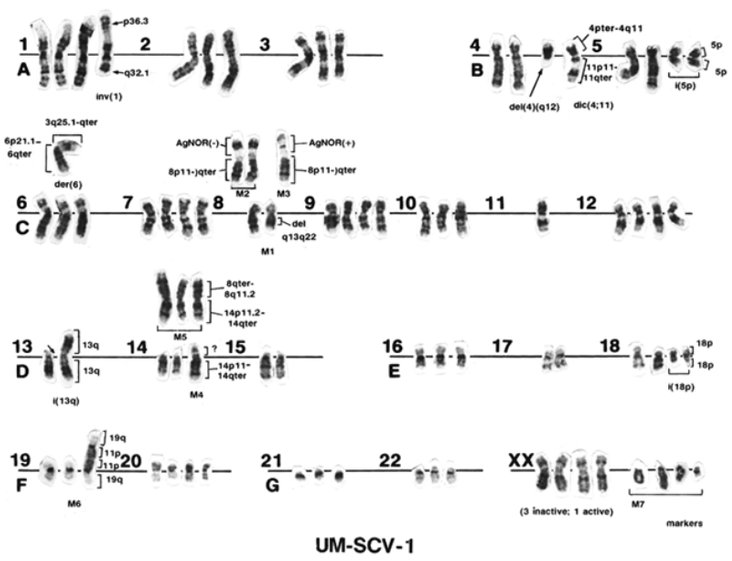


FIGURE 34.8. Representative karyotype from a hypotetraploid squamous carcinoma of the vulva. In this figure, the rearranged chromosomes are identified. Prominent abnormalities against the tetraploid background are the loss of 4q, 5q, 8p, 11 (there is only one intact 11, a total of three copies of 11p, and two copies of 11q), 15, 17, and 18q, as well as the gain of 5p, 8q, 14q, and 18p. (Reprinted from Grenman SE, Van Dyke DL, Worsham MJ, et al. Phenotypic characterization, karyotype analysis and in vitro tamoxifen sensitivity of new ER-negative vulvar carcinoma cell lines, UM-SCV-1A and UM-SCV-1B. *Int J Cancer* 1990;45:920–927.)

Germ-cell tumors (GCT) are a heterogeneous group of neoplasms that may be broadly classified into seminomatous and

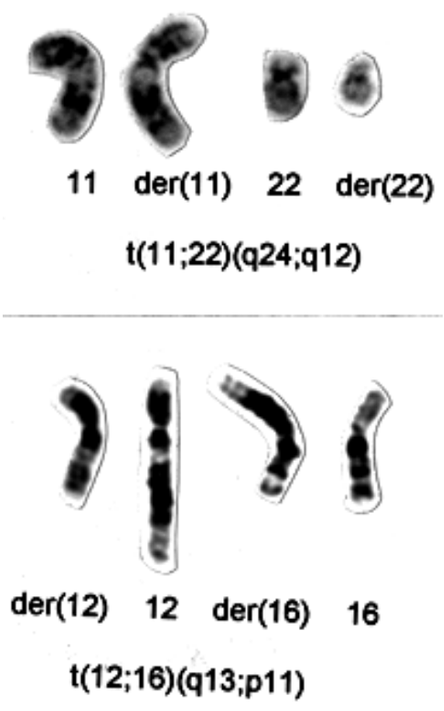


FIGURE 34.9. Partial karyotypes illustrating two sarcoma-associated chromosomal translocations: **(top)** t(11;22)(q24;q12), leading to the *EWS/FLI1* fusion gene, in a Ewing sarcoma, and **(bottom)** t(12;16)(q13;p11), leading to the *FUS/CHOP* fusion gene, in a myxoid liposarcoma.

nonseminomatous tumors. They usually have a testicular or ovarian origin, but are sometimes detected as extragonadal lesions, and may affect all age groups. However, of the more than 350 cases with abnormal karyotypes that exist in the literature, the majority has been testicular GCT from adult men. In this setting, the most frequent aberration, found in more than 80% of cases in some series, is one or more copies of an isochromosome for the short arm of chromosome 12. As the i(12p) usually is supernumerary, together with the fact that other rearrangements leading to gain of 12p sequences have been described, it seems reasonable to assume that the crucial molecular outcome of the isochromosome formation is gain of genes on 12p rather than loss of 12q material. The relevant loci on 12p have not yet been identified, but both cytogenetic and CGH data indicate that these are to be found in the proximal part of 12p (12p11-12). Apart from the i(12p), several other nonrandom chromosomal imbalances have been detected in testicular GCT. In relation to the ploidy level, which usually is near-triploid, gain of 1q, 7p, and chromosome 8, and loss of chromosomes 4, 5, 10, 11, 13, 18, and 19 and chromosome arm 9p have all been detected in at least 30% of the cases.

The i(12p) chromosome seems to be only slightly less common (40% of the cases) in extragonadal GCT in adults, but far less so in pediatric GCT. Instead, the majority of the pediatric tumors have had quite simple karyotypes with only a few numerical and/or structural rearrangements. This pattern resembles the situation in ovarian GCT in adult women. Most of these tumors have been classified as mature or immature teratomas and have recurrently shown trisomies for chromosomes 3, 8, 12, and 14.

SKIN TUMORS

Part of "34 - Solid Tumor Cytogenetics"

Clonal chromosomal aberrations have been described in less than 50 basal cell carcinomas. Almost all of them have shown a near-diploid chromosome count, which is in line with flow cytometry data. In general, simple numerical changes, in particular +18, +9, +20, +7, and +5, predominate (Jin et al., 1998c). The only chromosome arm that appears to be nonrandomly involved in structural rearrangements, usually deletions, is 9q. These deletions are likely to be part of the functional inactivation of the *PTCH* gene, a tumor suppressor gene. *PTCH* is lost or inactivated in both sporadic and familial basal cell carcinomas as determined by molecular genetic methods. Interestingly, a recurrent t(9;16)(q22;p13) has been found in three cases (Jin et al. 1997b). Whether this translocation represents another mechanism for *PTCH* deregulation remains to be elucidated.

Merkel-cell carcinoma is a rare, aggressive skin tumor of neuroendocrine origin. Only 25 cases with clonal chromosome aberrations have been reported, and no specific change has been detected among them. Chromosomes 1 and 11 have been most frequently involved, but apart from implicating the potential importance of deletions of distal 1p, the cytogenetic data are, so far, not helpful in directing future molecular studies.

Approximately 100 malignant melanomas of the skin with clonal chromosome abnormalities have been reported. In contrast to uveal melanomas (see above), most cases have had very complex karyotypes, which may reflect the fact that a large subset of the tumors has been highly advanced metastatic lesions (Thompson et al. 1995). Balanced structural aberrations occurring in more than one case are rare. Instead, the karyotypes are dominated by chromosomal imbalances, in particular gain of chromosome 7 and chromosome arms 1q, 6p, and 8q, and loss of chromosomes 3, 10, and 16 and chromosome arms 6q, 8p, and 9p, all of which have been detected in at least 15% to 40% of the cases. Trent and his coworkers noted a consistent t(1;6) (p22,q12-q21;q11-q13) in several melanomas and after review of the literature observed that band region 6q11-q13 was consistently the recipient of non-reciprocal translocations from several other chromosomes as well. These included (in order of frequency) two regions of chromosome 1 (p22,q12-q21), chromosomes 5, 3, 4, and 17. Subsequent analysis has led to the cloning and the identification of a gene identified by subtractive hybridization of expressed sequences suppressed by chromosome 6 transfer in human melanoma cells. *AIM1*, (absent in melanoma 1) is a novel non-lens member of the ß-γ crystallin superfamily (Ray et al. 1997) that is localized to 6q21. *AIM1* expression reverses tumorigenicity in melanoma cells and represents a candidate tumor suppressor gene associated with consistent 6q rearrangement. The only other imbalance for which a candidate target gene has been identified is loss of distal 9p. Band 9p21 harbors the tumor suppressor gene *CDKN2A* (*MTS*/*p16*), which is inactivated through deletion, mutation, or hypermethylation, in a large subset of sporadic melanomas. Furthermore, *CDKN2A* also is known to be constitutionally mutated in several kindreds with familial predisposition to malignant melanoma.

BONE AND SOFT TISSUE NEOPLASMS

Part of "34 - Solid Tumor Cytogenetics"

The group of musculoskeletal neoplasms comprises a wide variety of histopathologic entities that clinically cover the spectrum from totally innocuous lesions to highly aggressive malignancies. Because of the relative ease with which the cells from these mesenchymally or neuroectodermally derived tumors may be propagated *in vitro*, the cytogenetic literature is quite extensive, with roughly 1,000 cases with clonal changes having been reported for both the benign and the malignant variants. The interest in cytogenetic analysis of these tumors also has been spurred by the finding of highly consistent, sometimes pathognomonic, karyotypic patterns in many of them. Still, however, each histopathologic entity remains poorly investigated, and only for lipomas, liposarcomas, osteosarcomas, rhabdomyosarcomas, and synovial sarcomas have more than 100 cases with clonal chromosome aberrations been published. For recent, more detailed reviews of the karyotypic findings in musculoskeletal tumors, see Mandahl (1996), Sreekantaiah (1998), and Mandahl et al. (1999).

The clinical and biological variation among the musculoskeletal neoplasms is reflected by their karyotypes. The malignant tumors, the sarcomas, may be broadly dichotomized into one group characterized by specific, balanced translocations, and one group typically showing massive chromosomal rearrangements leading to recurrent, but nonspecific, chromosomal gains and losses. The former group includes myxoid liposarcoma, synovial sarcoma, clear-cell sarcoma, alveolar rhabdomyosarcoma,

extraskeletal myxoid chondrosarcoma, Ewing sarcoma, desmoplastic small-cell round-cell tumor, and infantile fibrosarcoma (Table 34.5). In all of these sarcoma types, the characteristic translocation leads to the creation of a fusion gene, which in turn will encode a novel hybrid protein. Because one of the two genes involved in each translocation always encodes a transcription factor, it is believed that the oncogenic effect is exerted by influencing the expression of other genes. Examples of consistent translocations in liposarcoma and Ewing's sarcoma are illustrated in Fig. 34.9. Also some unbalanced rearrangements are strongly associated with certain sarcoma types, such as the presence of one or more supernumerary ring chromosomes in dermatofibrosarcoma protuberans, parosteal osteosarcoma, and other sarcomas of low or borderline malignancy (Table 34.5).

TABLE 34.5. BONE AND SOFT TISSUE TUMORS WITH CHARACTERISTIC CHROMOSOMAL FINDINGS

Tumor Type	Chromosomal Aberrations	Molecular Genetic Aberrations
Sarcomas		
Alveolar rhabdomyosarcoma	t(1;13)(p36;q14)	*FKHR/PAX7* fusion
	t(2;13)(q35;q14)	*FKHR/PAX3* fusion
Clear-cell sarcoma	t(12;22)(q13;q12)	*EWS/ATF1* fusion
Dermatofibrosarcoma	r(17;22)	*COL1A1/PDGFB* fusion
protuberans	t(17;22)(q22;q13)	
Desmoplastic small-cell, round-cell tumor	t(11;22)(p13;q12)	*EWS/WT1* fusion
Extraskeletal myxoid	t(9;22)(q22;q12)	*EWS/CHN(TEC)* fusion
chondrosarcoma	t(9;17)(q22;q11)	*RBP56/CHN(TEL)* fusion
Ewing tumors	t(2;22)(q33;q12)	*EWS/FEV* fusion
	t(7;22)(p22;q12)	*EWS/ETV1* fusion
	t(11;22)(q24;q12)	*EWS/FLI1* fusion
	t(17;22)(q12;q12)	*EWS/EIAF* fusion
	t(21;22)(q22;q12)	*EWS/ERG* fusion
Infantile fibrosarcoma	+8,+11,+17,+20,	
	t(12;15)(p13;q25-26)	*ETV6/NTRK3* fusion
Myxoid liposarcoma	t(12;16)(q13;p11)	*FUS/CHOP* fusion
	t(12;22)(q13;q12)	*EWS/CHOP* fusion
Parosteal osteosarcoma	supernumerary ring chromosomes	amplification of 12q sequences
Synovial sarcoma	t(X;18)(p11;q11)	*SYT/SSX1 or SYT/SSX2* fusion
Benign Tumors		
Adipose-tissue tumors		
-atypical lipomatous tumors	supernumerary ring chromosomes or giant markers	amplification of 12q sequences, including *MDM2*
-hibernoma	rearrangements of 11q13	deletion of 11q13 sequences, including *MEN1*
-lipoblastoma	rearrangements of 8q11-13	rearrangements of *PLAG1*
-lipoma	rearrangements of 6p21 or 12q13-15	rearrangements of *HMGIY* and *HMGIC*
	t(3;12)(q27-28;q14-15)	*HMGIC/LPP* fusion
-spindle cell/pleomorphic lipoma	del(13q), del(16q)	unknown
Osteocartilaginous exostoses	del(8)(q24)	loss of *EXT1*
Schwannoma	del(22q), -22	loss of *SCH (NF2)*

The subgroup characterized by more complex karyotypes comprises the majority of the sarcomas, including, e.g., leiomyosarcoma, malignant fibrous histiocytoma, embryonal rhabdomyosarcoma, malignant peripheral nerve-sheath tumors, chondrosarcoma, and osteosarcoma. Although no specific aberration has been detected among these sarcomas, they all show nonrandom, but largely overlapping, patterns of chromosomal gains and losses. For none of the recurrent imbalances has the crucial molecular consequence been pinpointed, but, as for other tumor types, it is assumed that repeatedly deleted regions indicate the location of tumor suppressor genes, and that frequently amplified regions harbor dominantly acting oncogenes.

The benign musculoskeletal tumors are as cytogenetically heterogeneous as the sarcomas (Table 34.5). However, they usually have near-diploid chromosome counts with only a few structural and/or numerical aberrations, distinguishing them from many of their malignant counterparts. The lipomatous tumors may serve as a good example of the excellent correlations that exist between histopathologic and cytogenetic subgroups: ordinary lipomas typically show translocations involving 6p21 or 12q13-15, lipoblastomas are characterized by rearrangements of 8q11-13, pleomorphic and spindle cell lipomas show deletions of 13q and/or 16q, hibernomas have rearrangements of 11q13, and atypical lipomatous tumors have supernumerary ring chromosomes and/or giant marker chromosomes. Also for some of the benign lesions, the molecular consequences of the chromosomal rearrangements have been clarified. For instance, the 6p21 and 12q13-15 rearrangements, usually translocations, in lipomas lead to the disruption, sometimes through fusion with other genes, of the *HMGIY* and *HMGIC* genes, respectively.

NONNEOPLASTIC LESIONS

Part of "34 - Solid Tumor Cytogenetics"

The interpretation of cytogenetic findings in cell cultures from suspected solid tumors presents problems that are rarely encountered in the analysis of neoplastic hematologic disorders. While

the detection of clonal, acquired chromosome abnormalities in a peripheral blood or bone marrow sample, with the well-known exception of -Y in elderly men, is a highly reliable sign of neoplasia, chromosomal changes in other tissues do not necessarily represent neoplastic cells. Although relatively few studies on the cytogenetic picture in tissue samples from adults without neoplasia have been conducted, preliminary data do indicate that accumulation of chromosome damage might be a common phenomenon also in nonneoplastic cells. Not only have clonal changes, in particular numerical changes, been detected in normal brain, kidney, lung, synovia, and oral mucosa tissue, but there also is an emerging picture that somatic mutations, as detected at the chromosome level, may be nonrandomly associated with proliferative lesions that traditionally are regarded as nonneoplastic. For instance, in osteoarthritis, a degenerative joint disease characterized by cartilage destruction, synovial hyperproliferation, and bony outgrowths known as osteophytes, close to 90% of osteophytes and 95% of synovial samples show clonal chromosome aberrations (Mertens et al. 1996). Even more peculiar, almost all of them have the same abnormality, i.e., trisomy 7, and it has been established that the trisomic cells, compared to diploid cells, have an increased proliferative capacity *in vitro*. Furthermore, some 10% of the cases have breakpoints in 12q13-15, and it has been shown by FISH that these rearrangements may lead to disruption of the *HMGIC* gene in a way that seemingly is indistinguishable from what is thought to be essential in the pathogenesis of a large variety of benign mesenchymal neoplasms (Broberg et al. 1999). Finally, in rheumatoid arthritis, another inflammatory joint disease characterized by trisomy 7, molecular genetic studies have shown frequent *TP53* mutations and altered expression of several oncogenes, features that otherwise are associated with malignant cell proliferation (Aupperle et al. 1998). Thus, it can no longer be argued that somatic mutations are restricted to tumor cells.

How then, should one interpret the frequent finding in some tumor types, notably those of epithelial origin, of pseudodiploid karyotypes with a single or only a few numerical and/or structural changes? As has been shown to be the case in squamous cell carcinomas of the head and neck region and in breast carcinomas, the cytogenetic finding is often at odds with flow cytometric, molecular, or CGH data, which clearly demonstrate the presence of a highly aneuploid cell population within the same tumor specimen (Persson et al. 1999). Bearing in mind that most tumor samples used for cytogenetic analysis contain a mixture of tumor parenchymal and stromal cells and, furthermore, that it is impossible to distinguish between these two cell types after hypotonic treatment and fixation, one possible interpretation of the discrepancies would be that the clonal chromosome aberrations sometimes could be part of the stroma. Whereas this seems to be a reasonable explanation for some tumors, e.g., head and neck carcinomas, the consistency of the detected aberrations in tumor samples and the lack of such aberrations in normal tissue, argue against this hypothesis for other tumors, such as breast carcinomas. The alternative explanation would be that the tumor is composed of more than one, related or unrelated, neoplastic cell population. However, the majority of the data from studies looking specifically at this question indicate monoclonal origin. Most probably, these questions regarding the significance of some genetic findings cannot be resolved until they are addressed by combined cytogenetic, molecular genetic, and immunohistochemical approaches.

CLINICAL SIGNIFICANCE OF CYTOGENETIC FINDINGS IN SOLID TUMORS

Part of "34 - Solid Tumor Cytogenetics"

Ultimately, there are two aims with characterizing the cytogenetic profiles of solid tumors. First, the hope is of course that by learning more about the chromosomal mechanisms that are operative in tumor development one will also obtain a better understanding of the cellular processes that distinguish the neoplastic cells from their normal progenitors. This, in turn, would open up possibilities for designing new treatment strategies, or perhaps even preventing malignant transformation. Second, even before such detailed knowledge about the genetic features of tumor cells has accrued, the cytogenetic results could be useful as diagnostic and prognostic markers. Certainly, genetic investigations, be they at the chromosome or DNA level, are already an integral part of the battery of diagnostic tools that are employed in the classification of small cell, round cell tumors in children, and during recent years the relevance of cytogenetic analysis in the differential diagnosis of musculoskeletal tumors has become increasingly acknowledged. Other situations where cytogenetic analysis may be of help in providing a specific diagnosis or in distinguishing between malignant and benign tumors are discussed in a recent review article by Mitelman et al. (1997b). A few relevant examples are provided below.

Differential Diagnosis

Small cell and round cell tumors in children pose a considerable diagnostic dilemma because it is often impossible to make a diagnosis based on histology alone. The presence of a specific translocation, t(11;22), will make a definite diagnosis of Ewing's sarcoma, whereas the presence of a 1p deletion will establish the diagnosis of neuroblastoma. Similarly, neuroblastoma can be distinguished from Askin's tumor and neuroepithelioma by the presence of t(11;22) in the latter tumors. For a long time, neuroepithelioma and Askin's tumor were considered to be atypical neuroblastomas and were treated accordingly. However, because, like Ewing's sarcoma, they contain t(11;22), they since have been reclassified as primitive neuroectodermal tumors, and they now receive the more favorable therapy used for Ewing's sarcoma.

The presence of a specific change, t(12;16), can be used to make the diagnosis of a subset of liposarcomas, i.e., myxoid liposarcoma. Even though the distinction between benign and malignant tumors is clear in most soft-tissue tumors, the t(12;16) also can be utilized to make this distinction because the benign tumors have chromosome 12q13 involvement but do not show the chromosome 16 involvement. Abnormalities in chromosomes 2, 13, and 1 can be used to distinguish rhabdomyosarcoma from other neuroepithelial tumors. A specific translocation, t(2;13), might even distinguish the alveolar subtype.

Cytogenetics has been clearly underutilized in the area of diagnosis of malignant disease in body fluids such as pleural and peritoneal effusions. Several studies have provided evidence for the correlation between the presence of a clonal abnormality and the presence of neoplastic cells. Thus, cytogenetic analysis can be a useful adjunct to standard cytology in the differential diagnosis of neoplasia versus inflammation or other causes.

Prognostic Implications

In bladder cancer, chromosome abnormalities were recognized as predictors of clinical outcome even before the banding era. Banding techniques have refined the accuracy of identifying chromosome changes in this tumor type, and some specific changes, such as the deletion of 11p, have been shown to be associated with tumor progression. In general, low-grade, noninvasive bladder tumors have diploid karyotypes with few markers, and the higher-grade, invasive tumors have polyploid karyotypes with multiple marker chromosomes. The presence of marker chromosomes also is correlated with recurrence and mortality. Patients whose tumors did not have marker chromosomes had only a 5% rate of recurrence and achieved a 5-year survival rate of 90%, whereas patients whose tumors had markers had a 50% recurrence rate and a 5-year survival rate of only 75%. Thus, the karyotype can be a prognostic indicator in bladder cancer.

In neuroblastoma, the presence and the copy number of double minutes (*MYCN* amplification) has clinical implications. Patients with low-grade tumors and low copy number of double minutes have a much better likelihood of survival compared to patients with high-grade tumors and higher double minutes. Even among patients with high-grade tumors, those with low numbers of double minutes do well compared to those with high numbers. In general, patients with *MYCN* amplification also have chromosome 1 deletions, and this group of patients has the poorest prognosis.

Another prognostic indicator already mentioned above is the amplification of HER-2/neu (*ERBB2*) oncogene at the DNA, mRNA, and protein level that has been correlated with a shortened disease-free interval and decreased overall survival in breast and ovarian cancers. Of particular interest is that this prognostic indicator has now led to the development of a new therapeutic agent consisting of a humanized anti-HER2/Neu antibody. This agent shows great promise as it has caused regression and inhibition of tumor progression even in far advanced cases of metastatic disease.

In colon cancer, tumors from the left side often have more genetic changes, including chromosomal as well as allelic losses, than tumors from the right side. These genetic changes are correlated with disease-free interval and ultimate survival. Interestingly, loss of chromosome arm 18q has been linked to poorer prognosis in colon cancer. Using microsatellite markers and DNA from formalin-fixed, archived tumors, the investigators found that LOH on 18q was linked to poorer survival within stage II and stage III patients. For stage II patients, 5-year survival was 93% in cases where there was no LOH on 18q but only 54% in cases with LOH. Similarly for stage III, the 5-year survival decreased from 52% to 38% if 18q LOH was present.

In head and neck squamous-cell carcinoma, LOH on distal 18q also is associated with poorer survival. In multivariate analysis LOH on 18q was the strongest predictor of decreased survival when compared to age, black race, male gender, advanced stage, tumor histology, and the presence of neck disease. Of these indicators, only black race and advanced stage also were significant (Pearlstein et al. 1998). In head and neck cancer, rearrangement affecting 11q13 also is associated with reduced survival compared to those cases without 11q13 chromosome change (Åkervall et al. 1995). Many of these cases are likely a result of CCND1 amplification, which also is associated with poorer outcome.

In principle, there is no reason to expect that the cytogenetic profile would be less informative for solid tumors than for hematologic malignancies, but the database needed to validate such an assumption is presently far too limited. As a general phenomenon, however, it seems, not unexpectedly, as if the degree of cytogenetic complexity is correlated with clinical outcome (Mitelman et al. 1997b). Significantly shortened survival for patients with complex tumor karyotypes compared to patients whose tumors display only simple abnormalities has been reported for carcinomas of the breast, colon, head and neck region, kidney, ovary, and pancreas. These findings should be kept in mind when considering the many reports, usually including far less than 100 patients, claiming prognostic relevance for a particular chromosome or gene abnormality. Furthermore, with few exceptions, none of the reported correlations between cytogenetics and clinical outcome has been corroborated in independent series. Bearing in mind that some tumor types are rare and that many of them now are analyzed only by alternative techniques, it may well turn out to be difficult to reach any firm conclusions regarding the prognostic importance of cytogenetic findings in solid tumors unless several cytogenetic laboratories combine their results.

SUGGESTED READINGS

Aaltonen LA, Peltomaki P, Leach F, et al. Clues to the pathogenesis of familial colorectal cancer. *Science* 1993;260:812–816.

Adeyinka A, Mertens F, Idvall I, et al. Different patterns of chromosomal imbalances in metastasising and non-metastasising primary breast carcinomas. *Int J Cancer* 1999;84:370–375.

Åkervall J, Jin Y, Wennerberg J, et al. Chromosomal abnormalities involving 11q13 are associated with poor prognosis in squamous cell carcinoma of the head and neck. *Cancer* 1995;76:853–859.

Aupperle KR, Boyle DL, Hendrix M, et al. Regulation of synoviocyte proliferation, apoptosis, and invasion by the p53 tumor suppressor gene. *Am J Pathol* 1998;152:1091–1098.

Bardi G, Sukhikh T, Pandis N, et al. Karyotypic characterization of colorectal adenocarcinomas. *Genes Chromosomes Cancer* 1995;12:97–109.

Belge G, Roque L, Soares J, et al. Cytogenetic investigations of 340 thyroid hyperplasias and adenomas revealing correlations between cytogenetic findings and histology. *Cancer Genet Cytogenet* 1998;101:42–48.

Broberg K, Höglund M, Limon J, et al. Rearrangement of the neoplasia-associated gene HMGIC in synovia from patients with osteoarthritis. *Genes Chromosomes Cancer* 1999;24:278–282.

Bronner CE, Baker, SM, Morrison PT, et al. Mutation in the mismatch repair gene homologue hMLH1 associated with hereditary non-polyposis colon cancer. *Nature* 1994;368:258–261.

Carey TE, Van Dyke DL, Worsham MJ, et al. Characterization of human laryngeal primary and metastatic squamous cell carcinoma cell lines UM-SCC-17A and UM-SCC-17B. *Cancer Res* 1989;49:6098–6107.

Carey TE, Van Dyke DL, Worsham MJ. Nonrandom chromosome aberrations and clonal populations in head and neck cancer. *Anticancer Res* 1993;13:2561–2568.

Dal Cin P, Timmerman D, Van den Berghe I, et al. Genomic changes in endometrial polyps associated with tamoxifen show no evidence for its action as an external carcinogen. *Cancer Res* 1998;58:2278–2281.

Fearon ER and Vogelstein B. A genetic model of colorectal tumorigenesis. *Cell* 1990;61:759–767.

Frank CJ, McClatchey KD, Devaney KO, et al. Evidence that loss of chromosome 18q is associated with tumor progression. *Cancer Res* 1997;57:824–827.

Ganju V, Jenkins RB, O'Fallon JR, et al. Prognostic factors in gliomas. A multivariate analysis of clinical, pathologic, flow cytometric, cytogenetic, and molecular markers. *Cancer* 1994;74:920–927.

Gorunova L, Höglund M, Andrén-Sandberg Å, et al. Cytogenetic analysis of pancreatic carcinomas: intratumor heterogeneity and nonrandom pattern of chromosome aberrations. *Genes Chromosomes Cancer* 1998;23:81–99.

Grenman SE, Van Dyke DL, Worsham MJ, et al. UM-EC-1, a new hypodiploid human cell line derived from a poorly differentiated endometrial cancer. *Cancer Res* 1988;48:1864–1873.

Hagemeijer A, Versnel MA, Van Drunen E, et al. Cytogenetic analysis of malignant mesothelioma. *Cancer Genet Cytogenet* 1990;47:1–28.

Hahn SA, Schutte M, Shamsul Hoque ATM, et al. DPC4, a candidate tumor suppressor gene at human chromosome18q21.1. *Science* 1996;271:350–353.

Hansen LL, Yilmaz M, Overgaard J, et al. Allelic loss of 16q23-2-24.2 is an independent marker of good prognosis in primary breast cancer. *Cancer Res* 1998;58:2166–2169.

Hay R, Gazdar A, Park J-G, eds. *Atlas of Human Tumor Cell Lines*. Orlando, Fl: Academic Press, Inc., Harcourt Brace Jovanovich, 1994.

Heim S and Mitelman F. *Cancer Cytogenetics*, 2nd ed. New York: Wiley-Liss, 1995.

Jin C, Jin Y, Wennerberg J, et al. Clonal chromosome aberrations accumulate with age in upper aerodigestive tract mucosa. *Mutat Res* 1997a;374:63–72.

Jin Y, Mertens F, Persson B, et al. The reciprocal translocation t(9;16)(q22;p13) is a primary chromosome abnormality in basal cell carcinomas. *Cancer Res* 1997b;57:404–406.

Jin Y, Höglund M, Jin C, et al. FISH characterization of head and neck carcinomas reveals that amplification of band 11q13 is associated with deletion of distal 11q. *Genes Chromosomes Cancer* 1998a;22:312–320.

Jin Y, Jin C, Wennerberg J, et al. Cytogenetic and FISH characterization of chromosome 1 rearrangements in head and neck carcinomas delineate a target region for deletions within 1p11-1p13. *Cancer Res* 1998b;58:5859–5865.

Jin Y, Mertens F, Persson B, et al. Nonrandom numerical chromosome abnormalities in basal cell carcinomas. *Cancer Genet Cytogenet* 1998c;103:35–42.

Johansson B, Heim S, Mandahl N, et al. Trisomy 7 in nonneoplastic cells. *Genes Chromosomes Cancer* 1993;6:199–205.

Kas K, Voz ML, Röijer E, et al. Promoter swapping between the genes for a novel zinc finger protein and ß-catenin in pleiomorphic adenomas with t(3;8)(p21;q12) translocations. *Nat Genet* 1997;15:170–174.

Kazmierczak B, Meyer-Bolte K, Hue Tran K, et al. A high frequency of tumors with rearrangements of genes of the HMGI(Y) family in a series of 191 pulmonary chondroid hamartomas. *Genes Chromosomes Cancer* 1999; 26:125–133.

Kirchhoff M, Rose H, Laub Petersen B, et al. Comparative genomic hybridization reveals a recurrent pattern of chromosomal aberrations in severe dysplasia/carcinoma in situ of the cervix and in advanced stage cervical carcinoma. *Genes Chromosomes Cancer* 1999;24:144–150.

Knezevich SR, Garnett MJ, Pysher TJ, et al. ETV6-NTRK3 gene fusions and trisomy 11 establish a histogenetic link between mesoblastic nephroma and congenital fibrosarcoma. *Cancer Res* 1998;58:5046–5048.

Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA* 1971;68:820–823.

Kohonen-Corish MR, Doe WF, St. John DJ, et al. Chromosome 2p linkage analysis in hereditary non-polyposis colon cancer. *J Gastroenterol Hepatol* 1995;10:76–80.

Koi M, Umar A, Chauhan DP, et al. Human chromosome 3 corrects mismatch repair deficiency and microsatellite instability and reduces N-methyl-n'-nitro-nitrosoguanidine tolerance in colon tumor cells with homologous hMLH1 mutation. *Cancer Res* 1994;54:4308–4312.

Laken SJ, Papdopoulos N, Petersen GM, et al. Analysis of masked mutations in familial adenomatous polyposis. *Proc Natl Acad Sci USA* 1999;96:2322–2326.

Lu Y-J, Birdsall S, Osin P, et al. Phyllodes tumors of the breast analyzed by comparative genomic hybridization and association of increased 1q copy number with stromal overgrowth and recurrence. *Genes Chromosomes Cancer* 1997;20:275–281.

Lundgren R, Heim S, Mandahl N, et al. Chromosome abnormalities are associated with unfavorable outcome in prostatic cancer patients. *J Urol* 1992;147:784–788.

Lundin C and Mertens F. Cytogenetics of benign breast lesions. *Breast Cancer Res Treat* 1998;51:1–15.

Mandahl N. Cytogenetics and molecular genetics of bone and soft tissue tumors. *Adv Cancer Res* 1996;69:63–99.

Mandahl N. Methods in solid tumour cytogenetics. In: Rooney DE and Czepulkowski BH, eds. *Human Cytogenetics: A Practical Approach*, Vol. 2: Malignancy and Acquired Abnormalities, 3rd ed. Oxford: Oxford University Press, 2000.

Mandahl N, Mertens F, Mitelman F. Genetic changes in bone and soft tissue tumors. *Acta Orthop Scand* 1999;70(suppl 285):30–40.

Masters J (ed.). *Human Cell Culture* Vols. 1 and 2, Cancer Cell Lines. Dordrecht, Holland: Kluwer Academic Press, 1999.

Mertens F, Pålsson E, Lindstrand A, et al. Evidence of somatic mutations in osteoarthritis. *Hum Genet* 1996;98:651–656.

Mertens F, Johansson B, Höglund M, et al. Chromosomal imbalance maps of malignant solid tumors: A cytogenetic survey of 3185 neoplasms. *Cancer Res* 1997;57:2765–2780.

Mitelman F. *Catalog of Chromosome Aberrations in Cancer '98*, CD version 1. Johansson B and Mertens F, eds. New York: Wiley-Liss, 1998.

Mitelman F, Mertens F, Johansson B. Breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nat Genet* 1997a;15:417–474.

Mitelman F, Johansson B, Mandahl N, et al. Clinical significance of cytogenetic findings in solid tumors. *Cancer Genet Cytogenet* 1997b;95:1–8.

Morton CC. Many tumors and many genes. Genetics of uterine leiomyomata. *Am J Pathol* 1998;153:1015–1020.

Motokura T, Bloom T, Kim HG, et al. A novel cyclin encoded by a bcl-1-linked candidate oncogene. *Nature* 1991;350:512–515.

Pearlstein RP, Benninger MS, Carey TE, et al. Loss of 18q predicts poor survival in patients with squamous cell carcinoma of the head and neck. *Genes Chromosomes Cancer* 1998;21:333–339.

Persson K, Pandis N, Mertens F, et al. Chromosomal aberrations in breast cancer: a comparison between cytogenetics and comparative genomic hybridization. *Genes Chromosomes Cancer* 1999;25:115–122.

Ray ME, Wistow G, Su YA, et al. AIM1, a novel non-lens member of the beta gamma crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc Natl Acad Sci USA* 1997;94:3229–3234.

Scheurlen WG, Seranski P, Mincheva A, et al. High-resolution deletion mapping of chromosome arm 17p in childhood primitive neuroectodermal tumors reveals a common chromosomal disruption within the Smith-Magenis region, an unstable region in chromosome band 17p11.2. *Genes Chromosomes Cancer* 1997;18:50–58.

Schoenmakers EFPM, Huysmans C, Van de Ven WJM. Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. *Cancer Res* 1999;59:19–23.

Schwab M, Praml C, Amler LC. Genomic instability in 1p in human malignancies. *Genes Chromosomes Cancer* 1996;16:211–229.

Sisley K, Rennie IG, Parsons MA, et al. Abnormalities of chromosomes 3 and 8 in posterior uveal melanoma correlate with prognosis. *Genes Chromosomes Cancer* 1997;19:22–28.

Sreekantaiah C. The cytogenetic and molecular characterization of benign and malignant soft tissue tumors. *Cytogenet Cell Genet* 1998;82:13–29.

Steinarsdóttir M, Pétursdóttir I, Snorradóttir S, et al. Cytogenetic studies of breast carcinomas: Different karyotypic profiles detected by direct harvesting and short-term culture. *Cenes Chromosomes Cancer* 1995;13:239–248.

Steudel WI, Feld R, Henn W, et al. Correlation between cytogenetic and clinical findings in 215 human meningiomas. *Acta Neurochir* 1996;65:73–76.

Sun PC, Schmidt AP, Pashia ME, et al. Homozygous deletions define a region of 8p23.2 containing a putative tumor suppressor gene. *Genomics* 1999;62:184–188.

Taetle R, Aickin M, Panda L, et al. Chromosome abnormalities in ovarian adenocarcinoma: II. Prognostic impact of nonrandom chromosome abnormalities in 244 cases. *Genes Chromosomes Cancer* 1999;25:46–52.

Teixeira MR, Pandis N, Bardi G, et al. Karyotypic comparisons of multiple tumorous and macroscopically normal surrounding tissue samples from patients with breast cancer. *Cancer Res* 1997;56:855–859.

Thompson FH, Emerson J, Olson S, et al. Cytogenetics of 158 patients with regional or disseminated melanoma. Subset analysis of near-diploid and simple karyotypes. *Cancer Genet Cytogenet* 1995;83:93–104.

Van Dekken, Geelen E, Dinjens WNM, et al. Comparative genomic hybridization of cancer of the gastroesophageal junction: Deletion of 14q31-32.1 discriminates between esophageal (Barrett's) and gastric cardia adenocarcinomas. *Cancer Res* 1999;59:748–752.

Worsham MJ, Van Dyke DL, Grenman SE, et al. Consistent chromosome abnormalities in squamous cell carcinoma of the vulva. *Genes Chromosomes Cancer* 1991;3:420–432.

Worsham MJ, Carey TE, Benninger MS, et al. Clonal cytogenetic evolution in a squamous cell carcinoma of the skin from a xeroderma pigmentosum patient, HFH-SCC-XP-1. *Genes Chromosomes Cancer* 1993;7:158–164.

Worsham MJ, Wolman SR, Carey TE, et al. Common clonal origin of synchronous primary head and neck squamous cell carcinomas: analysis by tumor karyotype and fluorescence in situ hybridization. *Human Path* 1995;26:251–261.

Worsham MJ, Wolman SR, Carey TE, et al. Chromosomal aberrations identified in culture of squamous carcinomas are confirmed by fluorescence in situ hybridization. *J Clin Pathol: Mol Pathol* 1999;52:42–46.

Section 6 The Major Histocompatibility Complex

The Major Histocompatibility Complex - Introduction

Emanuel Hackel

Section Chief

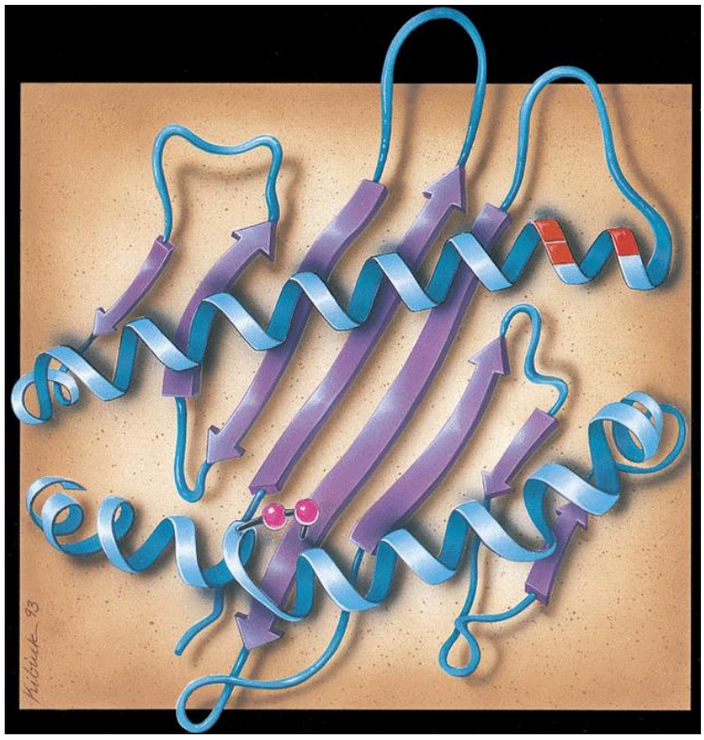


Figure.

35 HLA: Structure, Function, and Methodologies

36 Molecular HLA Typing

37 HLA: Applications

38 Bone Marrow Transplantation

As details of the molecular and immunological phenomena known collectively as HLA continue to unfold, application of these in clinical medicine, in forensics, and in a myriad of other areas follow close behind. Indeed, sometimes the details get in the way of seemingly beneficial applications, but these ultimately sort themselves out, and our understanding continues to evolve. Thus, the chapters on HLA methods and applications, modified and updated, are herewith presented with the full knowledge that, as in the past, changes are afoot. Current practices and procedures are widely used and lead to useful interpretations with effective results in clinical medicine as well as in research. The certainty that evolving knowledge will produce methodological changes in the near future does not detract at all from their importance as clinical tools and their fruitfulness in research.

HLA testing has evolved over the years through a variety of different methods. Some of these have persisted, with minor variations, for a long time. Given the diverse nature of the demonstrated structural differences between the class I, II, and III HLA antigens, it is not surprising that a number of diverse typing methods should be required to identify them. These methods are subject to constant review and are revised when it is appropriate. The methods presented here are those in use at the time of writing, but it should come as no surprise that variations of these may become standard. The use of lymphocytotoxicity is diminishing, as is the use of DNA probes. Use of reference strands — heteroduplexing — is emerging as a possible method. Class III antigens are found to be burgeoning with polymorphisms, but the role of these, clinically, remains unclear.

Newer specificities, identified by polymerase chain reaction–sequence-specific primer (PCR-SSP) techniques or by sequence-based typing (SBT) are increasingly in use, providing exquisite matching. The use of other methods, leading to less exquisite matching is on the wane. Yet, given the increased use of immunomodulation, and immunosuppression in solid organ transplantation, it may well be that the extreme exquisiteness becoming possible in HLA complex matching is not required for clinical success in these areas. This, however, does not diminish the need for maximum precision in HLA typing given its many other roles. As has always been the case with HLA, as we hone our understanding of the major histocompatibility complex ever more sharply, the need for more precise typing methods will be the driving force behind the development of these methods.

35
HLA: Structure, Function, and Methodologies

Malek Kamoun

Chester M. Zmijewski

Organ transplantation has long held fascination for scientists of many persuasions. Workers in diverse specialties have been attracted to this area for many reasons and it has burgeoned into activity requiring input from these diverse specialties. Attempts to transplant tissues and organs for the purpose of studying oncogenesis or ageing were thwarted by the failure of the engrafted tissues or organs to thrive in the body of the host. Characteristically grafts of skin exchanged between animals belonging to the same species begin to heal shortly after surgery. Blood vessels from the host extend into the grafted tissue and circulation is established. After 10 to 12 days, the vessels become occluded, circulation ceases, and necrosis begins. This culminates in the formation of an eschar and rejection of the graft at about day 14.

When skin grafts were exchanged between animals of the same inbred strain, which are akin to identical twins, healing occurred without incident and survived for the life of the animal. However, grafts exchanged between animals of different strains were rejected as expected. Finally, the F1 offspring from the mating of two different inbred strains would accept grafts from animals of either parental strain. On the other hand, either parent rejected grafts from an F1 donor. This offered experimental evidence indicating the phenomenon of graft rejection was under genetic control.

As early as 1944, Medawar clearly demonstrated that the “homograft” (currently referred to as “allograft”) reaction had an immunologic basis (1). He showed that although the rejection of skin did occur in 14 days, a second graft from the same donor was rejected much more rapidly. Such a secondary response implied immunologic memory. Furthermore, subsequent grafts were rejected even more rapidly until finally a situation took place in which vascularization failed to occur giving rise to a so-called “white graft.”

Microscopic examination of biopsies taken from the site of these grafts revealed the presence of a large number of small lymphocytes clustered around capillary venules, inflammatory cells, and, in the case of white grafts, massive thrombi with relatively little lymphocytic involvement.

Based on these observations and other experiments involving the generation of tolerance, Medawar postulated that there existed a set of alloantigens under genetic control that are present on donor tissues and recognized by the host. The host mounts an immune response to these antigens, which is not unlike a delayed hypersensitivity response consisting of immune lymphocytes as effectors. Upon continued exposure to antigen or after repeated challenges, antibodies directed against these antigens are formed. Indeed, it was eventually demonstrated that the features of the white graft were caused primarily by circulating antibodies that had been formed in response to previous grafts. This was the origin of the idea of histocompatibility antigens, which were first described by Snell on inbred strains of mice and later by Dausset, Payne, Van Rood, and others, in humans (2, 3).

Histocompatibility antigens were conceived as being a set of alloantigens, i.e., tissue surface antigens that differ among members of the same species, which were significant only as targets for the allograft reaction. However, shortly after their definition, it became clear that these antigens, which are controlled by a genetic complex referred to collectively as the Major Histocompatibility Complex (MHC), are involved in immunologic self recognition and therefore have a much deeper biologic function. It is now clear that histocompatibility antigens serve as self-recognition molecules. In this capacity, they help to trigger the T-cell–dependent immune response and actively participate in the immune elimination of cells infested with foreign substances.

The antigens of the HLA system comprise the histocompatibility antigens in man. The genetic complex controlling the production of these antigens is located on the distal one third of the short arm of chromosome 6. Counterparts of this system exhibiting the same general characteristics have been found in all mammalian species studied to date and are similarly genetically controlled. For example, the mouse equivalent, called H-2, is found on mouse chromosome 17. This system has been studied extensively and continues to serve as a model for the investigation of basic questions in immunology centered on immune responsiveness and tolerance. Among the other species in which the MHC systems have been studied are horses (ELA), cows (BLA), cats (FLA), dogs (DLA), rats (RtLA), and rhesus monkeys (RhLA).

GENETICS AND SPECIAL FEATURES OF HLA

Part of "35 - HLA: Structure, Function, and Methodologies"

Genes of the HLA Complex

The HLA region on chromosome 6 is composed of a set of very closely linked loci. The general arrangement of these loci is

shown in Fig. 35.1 (4). This region extends over approximately 4 centimorgans of DNA or about 4×10^6 base pairs, and contains more than 200 genes. The particular combination of HLA alleles found on an individual chromosome is known as an HLA haplotype. The genetic complex can be divided into three major regions, each of which controls the production of molecules that have distinct biologic functions. These are called class I, which encompasses the HLA-A, -B, and -C loci; class II, which includes the HLA-D region loci HLA-DR, DQ, and DP; and class III, which are loci whose gene products are molecules that belong to the complement system and include C2, C4, and Bf as well as the cytokines referred to as tumor necrosis factors (TNF). The class II region also includes the genes for the TAP-1 and TAP-2 peptide transporter, the LMP genes that encode proteasome subunits, and the gene encoding the DMα and DMß chains. The function of DM genes is to catalyze peptide binding to HLA class II molecules. The DOA and DOB genes, which encode the DO molecule, a negative regulator of DM, are also related to the class II genes. A listing of the loci in the HLA region arranged in order from the telomeric to the centromeric end is given in Table 35.1 (5).

TABLE 35.1. NAMES FOR GENES IN THE HLA REGION, ORDERED FROM TELOMERE TO CENTROMERE

Name[a]	Molecular Characteristics
HLA-F	Class I – like molecule
HLA-G	Class I – like molecule
HLA-H	Class I pseudogene
HLA-A	Class I α-chain
HLA-J	Class I pseudogene
HLA-E	Class I – like molecule
HLA-C	Class I α-chain
HLA-B	Class I α-chain
HLA-DRA	DR α chain
HLA-DRB1	DR ß1 chain determining specificities DR1, DR2, DR3, DR4, DR5 etc.
HLA-DRB2	pseudogene with DR ß – like sequences
HLA-DRB3	DR ß3 chain determining DR52 and Dw24, Dw25, Dw26 specificities
HLA-DRB4	DR ß4 chain determining DR53
HLA-DRB5	DR ß5 chain determining DR51
HLA-DRB6	DRB pseudogene found on DR1, DR2 and DR10 haplotypes
HLA-DRB7	DRB pseudogene found on DR4, DR7 and DR9 haplotypes
HLA-DRB8	DRB pseudogene found on DR4, DR7 and DR9 haplotypes
HLA-DRB9	DRB pseudogene, isolated fragment
HLA-DQA1	DQ α chain as expressed
HLA-DQB1	DQ ß chain as expressed
HLA-DQB3	DQ ß-chain-related sequence, not known to be expressed
HLA-DQA2	DQ α-chain-related sequence, not known to be expressed
HLA-DQB2	DQ ß-chain-related sequence, not known to be expressed
HLA-DOB	DO ß chain
TAP2	ABC (ATP binding cassette) transporter
LMP7	Proteasome-related sequence
TAP1	ABC (ATP binding cassette) transporter
LMP2	Proteasome-related sequence
HLA-DMB	DM ß chain
HLA-DMA	DM α chain
HLA-DOA	DO α chain
HLA-DPA1	DP α chain as expressed
HLA-DPB1	DP ß chain as expressed
HLA-DPA2	DP α-chain-related pseudogene
HLA-DPB2	DP ß-chain-related pseudogene

[a] Gene names given in bold type have been assigned since the 1996 Nomenclature Report.

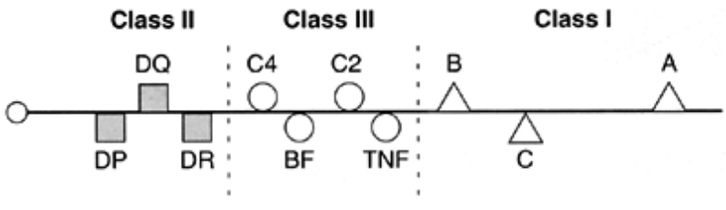


FIGURE 35.1. A diagrammatic representation of the distal third of the short arm of human chromosome 6 carrying the major loci of HLA. The region is divided into three areas that control the production of the class I, class II, and class III molecules.

Other Genes with Specialized Immune Functions Encoded in the HLA Region

Although the most important known function of the HLA gene products is the processing and presentation of antigens to T cells, many other genes map within the region; some of these are known to have other roles in the immune system, but many have yet to be characterized functionally. In addition to the highly

polymorphic HLA class I and class II genes, there are a several HLA class I genes encoding variants of these that show little polymorphism. Some of these genes function to inhibit cell killing by NK cells, a role played by several class I HLA molecules. Such a role has been suggested for HLA-G, which is expressed on fetus-derived placental cells. These cells express no HLA class I molecules and the expression of HLA-G may protect the fetus from the attack by either CD8 T cells or NK cells. Likewise, HLA-E has a specialized role in the recognition of NK cells. HLA-E binds to a very restricted subset of peptides, derived from the leader peptide of other HLA class I molecules. It is known that these peptide:HLA-E complexes can bind to CD94, one of the receptors on NK cells. HLA-E recognition by NK cells blocks killing. The class I and class II regions also contain a number of pseudogenes, i.e., genes that are not expressed as detectible peptides. These genes currently are under study as possible regulators of the intensity of immune responses.

HLA Polymorphism

The HLA system is highly polymorphic. Indeed it is one of the most polymorphic genetic systems in man. Associated with each of the loci is a set of alleles, each of which gives rise to the production of a unique antigenic specificity that is expressed on the cell surface and that can be detected by specific antibodies and/or immunologically activated T cells. The recognizable antigens of the HLA system are shown in Table 35.2 (6). The nomenclature for the alleles of the HLA system has been established and is controlled by a select committee of the World Health Organization (WHO). The antigen name is composed of a letter designating the locus e.g., A, B, C, DR, DP, DQ, and a number. Originally, before it was known that multiple loci controlled HLA, the numbers were assigned in the order of their discovery. Therefore, some inconsistencies occur such as 1, 2, and 3 assigned to the A locus, 5, 7, and 8 assigned to B, and then 9, 10, and 11 reverting to A. The more recently described antigens now follow in logical sequence. Another noticeable peculiarity is the absence of antigens 4 and 6 from the original list. These numerical assignments were reserved for some very special antigens that did not seem to follow the accepted patterns of inheritance or serology. Since then, they have been shown to be so-called “supertypic” or public antigens, which will be described in a subsequent section and are now called Bw4 and Bw6.

TABLE 35.2. THE HLA SPECIFICITIES 1998

		D			B		C
DPw1	DQ1	DR1	Dw1	B5	B40	B59	Cw1
DPw2	DQ2	DR103	Dw2	B7	B4005	B60(40)	Cw2
DPw3	DQ3	DR2	Dw3	B703	B41	B61(40)	Cw3
DPw4	DQ4	DR3	Dw4	B8	B42	B62(15)	Cw4
DPw5	DQ5(1)	DR4	Dw5	B12	B44(12)	B63(15)	Cw5
DPw6	DQ6(1)	DR5	Dw6	B13	B45(12)	B64(14)	Cw6
	DQ7(3)	DR6	Dw7	B14	B46	B65(14)	Cw7
	DQ8(3)	DR7	Dw8	B15	B47	B67	Cw8
	DQ9(3)	DR8	Dw9	B16	B48	B70	Cw9(3)
		DR9	Dw10	B17	B49(21)	B71(70)	Cw10(3)
		DR10	Dw11(7)	B18	B50(21)	B72(70)	
		DR11(5)	Dw12	B21	B51(5)	B73	
		DR12(5)	Dw13	B22	B5102	B75(15)	
		DR13(6)	Dw14	B27	B5103	B76(15)	
		DR14(6)	Dw15	B35	B52(5)	B77(15)	
		DR15(2)	Dw16	B37	B53	B78	
		DR17(3)	Dw18(6)	B39(16).	B55(22)	B81	
		DR18(3)	Dw19(6)	B3901	B56(22)		
			Dw20	B3902	B57(17)	Bw4	
		DR51	Dw21		B58(17)	Bw6	
		DR52	Dw22				
		DR53	Dw23				
			Dw24				
			Dw25				
			Dw26				

In some of the older literature, certain HLA specificities may carry a number prefixed with a “w.” In previous versions of the HLA nomenclature, this prefix stood for “workshop” and indicated a provisional designation pending confirmation based on extensive serologic testing. This was often a very labor-intensive process that was clouded by the fact that the antisera used for the testing were frequently multispecific. Further, these antisera were products of an immune response and as such carried the expected individual variations associated with such responses.

Current practice calls for serologic specificities to be named based on their correlation with an identified DNA sequence (6). Therefore, there is no need for provisional designations because the new antigenic specificities will have a primary biochemical definition that is devoid of serological vagaries. The “w” prefix, however, is retained in the case of the Bw4 and Bw6 epitopes to distinguish them from alleles; as well as the C locus specificities to distinguish them from the complement components.

The new system of nomenclature also allows for the differentiation between the currently accepted classic serologic splits and what are better termed allelic variants. Thus, for example, B51 and B52 are splits of the B5 specificity, distinguishable serologically,

while B*5101 and B*5102 are allelic variants of B51, discernable by biochemical techniques. (Alleles are designated by the locus letter followed by an asterisk and the allele's number.) Currently, however, new serologically defined splits that are a result of epitopes corresponding to combinations of alleles are not given official names. A comprehensive listing of serologically defined specificities and their allele equivalents has been published by Hurley et al. (7).

The individual loci of the HLA complex are tightly linked. Therefore, genetic recombination or crossing over is a rare event, occurring for the most part no more than 1% of the time between A and B and between B and DR and 0.6% between A and C and 0.2% between B and C. Thus, the alleles of the HLA system for the most part are inherited *en bloc* from each parent as a unit called a haplotype.

Fig. 35.2 shows the typical inheritance pattern of the HLA antigens in a hypothetical family with six siblings. Each of the first four siblings is different with respect to the HLA antigens he or she has inherited, however the group can be divided into pairs that share either identical maternal or identical paternal chromosomes. However, the fifth sibling must be identical to one of the others, because there are only four combinations possible. Thus, in any family having more than one sibling, there is a one in four chance that two siblings will be HLA-identical and a one in two chance that they will share one haplotype.

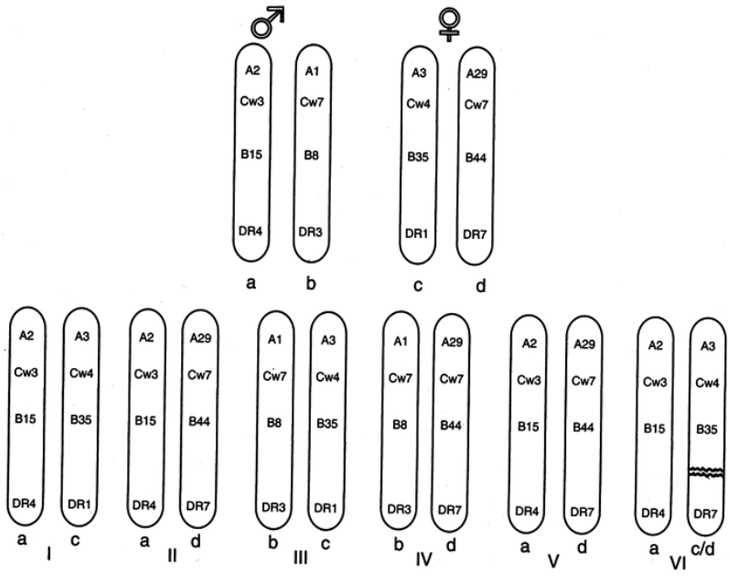


FIGURE 35.2. A pedigree of a family consisting of two parents and six siblings. The HLA haplotypes are labeled c and d. Siblings I through IV represent the different combinations. Sibling V is HLA-identical to sibling II. Sibling VI carries a maternal haplotype that is a recombinant in which the A, C, and B loci come from chromosome c and the DR locus comes from chromosome d.

The HLA-identical siblings in a family may be identical without being identical twins. The same combinations of gametes bearing the MHC of chromosome 6 may be expected in 25% of all offspring on the basis of chance combination alone. Nevertheless, the relatively low frequency of genetic recombination within the HLA region resulting from tight linkage is the phenomenon that permits transplantation between siblings other than identical twins.

Although recombination in the HLA system is an infrequent event, it does occur and usually between the B and the DR loci. An example of this type of recombination is given by sibling six in Fig. 35.2. In this case, offspring number six is identical to siblings numbered four and five with respect to class I antigens but differs with respect to class II antigens. Such recombination can have important clinical implications for transplantation.

Highly polymorphic multilocus systems such as HLA usually are found in the population at large in a state of genetic equilibrium. This equilibrium is such that the frequency with which two or three alleles from the linked loci occur together in a single haplotype is equal to the product of the individual gene frequencies in the population being observed. Although this holds true to a certain extent in the HLA system, there are some notable exceptions. The phenomenon of genetic disequilibrium or more properly gametic association is one of the hallmarks of the HLA system.

Gametic associations are regularly found between certain alleles of A and B, C and B, B and DR, and DR and DQ. In addition, there are instances where certain haplotypes contain even those sequences controlling the class III antigens in a unique gametic

TABLE 35.3. MAJOR GAMETIC ASSOCIATIONS[a]

Whites	Blacks
A1,B8	A36,B44
A30,B13	A30,B42
A25,B18	A29,B44
A3,B7	A2,B45
Cw4,B35	Cw4,B35
Cw2,B27	
Cw5,B44	
DR2,B7	DR2,B7
DR3,B8	DR3,B8
DR7,B13	DR7,B13
	DR11,B18

[a]From Baur MP, Neugebauer M, Albert ED.
Reference tables of two-locus haplotype frequencies for all MHC marker loci. In: Albert ED, Baur MP, Mayr WR, eds. *Histocompatability testing 1984*. New York: Springer-Verlag, 1984:677–755.

association, forming an extended haplotype. The measure of gametic association is given by the expression, "Δ" or delta and represents the difference between the calculated and observed frequencies of the haplotype. Some of the more prominent gametic associations are listed in Table 35.3 (8).

Information about gametic association is important. For example, it can be used to predict the presence of antigens that are difficult to detect. The products of the HLA-C locus are very difficult to detect using conventional serological techniques even though biochemical techniques have been used to detect as many as 61 allelic variants at this locus. The antigens appear to be weak and the antisera produced in response to them often are multispecific. For this and other reasons, the gene products of the C locus have not been considered in some clinical situations such as solid organ transplantation. However, because of the strong association of some C locus antigens with certain B locus antigens, it often is possible to predict the occurrence of the former whenever the later antigen is detected.

The cause of gametic association is unknown (9). Usually, this condition is found in biologic systems with alleles that have undergone recent mutation and not yet had an opportunity to reach equilibrium. This does not appear to be the case with the HLA system. The gametic associations observed extend throughout widely divergent populations. Such a finding implies that simultaneous and identical mutations that have not yet reached equilibrium occurred in these widely separated populations. This is a highly unlikely event.

In the mouse system, there is a set of genes at the telomeric end of the MHC region that preferentially inhibit recombination in haplotypes that carry them, thereby favoring gametic association (10). A search for a similar set of genes in man has been underway but thus far has been without reward. Genes have been found telomeric to HLA-A that give rise to putative MHC molecules. However, their function does not appear to be equivalent to that found in rodents.

Another theory often proposed to explain such a finding is selective advantage. With respect to HLA, this is difficult to rationalize because many of the haplotypes in gametic association also are associated with certain diseases. As will be discussed subsequently in more detail, the common haplotypes HLA-A3,B7 and HLA-A1,B8 are strongly associated with multiple sclerosis and myesthenia gravis, respectively. Although this is far from a convincing argument for advantage, some proponents of the idea point out that there may be other very beneficial genes associated with these haplotypes that by far outweigh the occasional detrimental effects and therefore do offer some advantage to the organism.

As with many of the other alloantigens such as the blood groups of the human red cells, the alleles of the HLA system have gene frequencies that vary from one population to the next. The variability exists not only among the different races but also among various ethnic groups. Some of the more prominent differences in gene frequencies are shown in Table 35.4 (11). Key frequencies of antigens common to all three racial groups or peculiar to a given group are listed in bold type. Thus, some antigens may be present in high frequency in whites but altogether absent in Asians. Other antigens may be present in blacks in fairly high frequency and virtually absent from whites. Although not shown, such distinctions may be present among various ethnic populations within the same race. Thus there may be certain antigens that are characteristic of eastern Europeans such as the Poles and Russians that are rarely found among the English. Similar analogies in antigen distribution may be found among the various black tribes of Africa and among various Asian peoples such as the Japanese, Chinese, and Koreans.

TABLE 35.4. HLA GENE FREQUENCIES IN VARIOUS POPULATIONS[a]

HLA	White	Black	Oriental
A1	**.149**	.033	.005
A2	**.269**	**.147**	**.246**
A23	.025	**.108**	.005
A24	.066	.029	**.356**
A11	**.059**	.006	**.090**
A34	.006	**.065**	.001
B7	**.098**	**.089**	**.059**
B8	**.090**	.029	.001
B38	**.032**	.000	.002
B42	.003	**.077**	.006
B53	.009	**.065**	.001
B54	.000	.000	**.073**
DR3	**.118**	**.173**	.016
DR4	**.148**	.049	**.235**
DR5	**.103**	**.133**	.022
DR9	.011	.027	**.122**

[a]From Baur MP, Danilovs JA. Population analysis of HLA-A, B, C, DR and other genetic markers. In: Terasaki PI, ed. *Histocompatability Testing 1980*. Los Angeles: UCLA Tissue Typing Laboratory, 1980:955–993.

Although the majority of the HLA alleles present in whites have been discovered as evidenced by gene frequencies that total almost 100%, the same may not be true of the other races. In considering the HLA-A and -B loci, one would expect that an individual should have at least four antigens. Two HLA-A antigens, one from each parent, along with two HLA-B antigens inherited in the same fashion. In testing African-American families, for example, it is not uncommon to find individuals who appear to lack one or more HLA antigens. It usually is clear from the familial pattern of inheritance that the individuals in question are not homozygous for a given allele and therefore phenotypically express it only once but rather they have inherited

what appears to be an undetectable characteristic, a so-called “blank.” The obvious explanation for such a finding is that the individual has a gene that causes the expression of an antigen that cannot be detected with the currently available battery of anti-HLA antisera. Indeed, studies of such individuals among African-Americans have revealed the existence of unique HLA antigens. Similar race-specific antigens have been reported in the Asiatic populations as well as other ethnic groups. Many of these new alleles recently have been defined by DNA-based typing techniques (12, 13).

These antigens become an important consideration in transplant programs associated with large metropolitan hospitals. The random organ donor population available to such institutions as well as their patient populations will be composed of racial minorities. Therefore, knowledge of these special antigens and the ability to detect them are essential. In addition, a knowledge of the variation in the gene frequency of HLA is important because it can have an effect on the ability to find an HLA-matched random cadaveric donor from the general population in a given geographic area. It might be very difficult to locate a suitable donor for a black patient with one or more unique antigens in some Midwestern city populated primarily with whites of Scandinavian ancestry. Numerous other examples could be cited for other regions of the United States that have a blend of peoples with widely divergent racial and ethnic backgrounds. For this reason, a system of organ sharing on a national basis needs to be employed in order to give patients the maximum potential for obtaining a well-matched graft.

HLA Class I Null Alleles and Their Significance in Transplantation

The importance of HLA class I null alleles was reported during the 11th and 12th International Histocompatibility Workshops and Conferences (14). New HLA class I null alleles usually are formed by conversion or recombination between pairs of existing alleles or genes and only rarely by point substitution. New HLA-B alleles are formed at a higher rate than HLA-A or -C, with HLA-A evolving new alleles somewhat more rapidly than -C. The indigenous populations of Latin American countries have evolved many more HLA class I alleles than their counterparts in North America, a difference that primarily is a result of the HLA-B locus. In the populations of South America, older HLA-B alleles have been, to a considerable extent, replaced by newer recombinant alleles. This reveals the general trend for allele turnover and demonstrates that 10,000 to 35,000 years (the estimated time that human populations have lived in America) can be sufficient to replace a founding set of HLA-B alleles by a derived set of recombinants.

Although inactivated HLA null alleles (such as A*24 null) only differ from the common functional allele (A*2402) by a single nucleotide change, the potential effects of such changes on function and histocompatibility are great. For example, if an A*2402 solid organ was transplanted into an A*24 null recipient, the recipient's immune system would recognize it as a completely incompatible class I antigen against which potent alloreactivity and cytotoxic T-cell responses could be made.

The Complexity of HLA Serologic Cross-Reactivity

The majority of antigens in the HLA system can be detected by means of serologic methods employing specific antibodies. One of the characteristic phenomena that has been observed when using anti-HLA alloantibodies is the frequent appearance of serologic cross reactivity. This is a condition in which an antiserum previously shown to react with one particular HLA specificity, upon further testing is shown to react either wholly or partially with another specificity. An example of this is shown in Table 35.5.

TABLE 35.5. CROSS-REACTIVITY RESULTING FROM SHARED EPITOPES

				Reaction With:		
HLA Specificity	Epitopes	Possible Antibodies	Observed Specificity	B5	B35	B18
B5	1, 2, 3	Anti-1 Anti-2 Anti-3	Anti-B5	+++	+	+
B35	1, 4, 5	Anti-1 Anti-4 Anti-5	Anti-B35	+	+++	+
B18	3, 6, 5	Anti-3 Anti-6 Anti-5	Anti-B18	+	+	+++

The occurrence of serologic cross reactivity is most easily attributed to the presence of a mixture of antibodies having specificities directed against several but perhaps related antigens. However, this does not appear to be the cause of cross reactivity in the HLA system (15). Usually, the repeated absorption of the HLA antiserum in question with cells bearing one of the specificities totally removes all reactivity for the other specificity. Such a finding indicates that the antiserum is truly monospecific and establishes that the cross reactivity may be attributable to factors at the level of the antigenic determinant, a conclusion supported by the results of experiences with *in vivo* immunization. In some cases, an antigenic challenge with a given primary antigen may result in the development of antibodies that react with another antigen not present in the immunizing dose. Furthermore, an attempt to immunize an individual with an antigen that he/she is lacking but that is cross reactive with one of his/her own antigens may not result in an immune response.

The phenomenon of antigenic cross reactivity is from the sharing of specific epitopes on the antigenic molecule. It is a regular feature of the HLA system whose antigens may be arranged in “cross reactive groups” (CREGS). The most prominent CREGS of the HLA-A and -B loci are shown in Fig. 35.3.

CREGS are important in that they allow for greater latitude in donor-recipient HLA matching and at the same time may be used to predict potential problems with graft outcome. For example, transplantation or platelet transfusion within a CREG often can be performed without eliciting any great uncontrollable immune response. By the same token, if a patient is shown to

have an antibody of a particular specificity, it is possible that under the proper circumstances his or her serum will react as well with tissues bearing other antigens within in the same CREG.

As the techniques and reagents for clinical HLA typing became more sophisticated, CREGS became less important and greater emphasis was placed on typing and matching for discreet specificities. However, recently, matching strategies based on CREGS are resurfacing. A number of investigators are exploring the possibility of CREG matching as a way to find cross-match compatible donors for patients who have developed HLA antibodies to a large number of specificities as a result of transfusion or previous transplant. Some success has been achieved by means of this approach.

The so-called public and private specificities of the HLA system are a special type of CREG and constitute a prominent feature of the antigens in the HLA system. Public antigens are specific epitopes that are found associated with each of the different specificities of an entire family of HLA molecules. Two such public specificities are the HLA-Bw4 and Bw6, mentioned previously. As the HLA system unfolded, it became clear that many of the antisera belonging to a number of investigators in the field and supposedly detecting a variety of different antigens were in fact detecting a similar set of antigens that could be given an HLA designation. However, the supposedly allelic antigens 4^a and 4^b originally described by van Rood did not appear to fit the scheme. Further investigation revealed that 4^a and 4^b were not products of antithetical allelic genes but rather were found associated with all antigens of the HLA-B locus. Because these specificities cross the boundaries of all of the other specificities of this locus, they were envisioned as public specificities.

In contrast to public specificities, which tend to group antigens into large categories, private specificities tend to divide the groups. For example, the specificity HLA-A9 has two private, serologically detected specificities, HLA-A23 and A24. Cells expressing either of these antigens will react not only with anti-HLA-A23 or A24 antisera but also with anti-HLA-A9. Thus the private specificities are true specificities in their own right. In addition HLA-A23 and A24, each have a different gene frequency in whites and in blacks indicating that they probably arose from a common ancestral gene after the divergence of the races.

BIOCHEMISTRY OF HLA

Part of "35 - HLA: Structure, Function, and Methodologies"

In the previous section, the HLA chromosomal region was divided into sections that controlled the production of three different types of molecules that were called class I, class II and class III. Each of these is unique in terms of its tissue distribution, biochemical structure, and biologic function.

Class I molecules are expressed constitutively on the surface of all nucleated cells in the body with the possible exception of the cells of the central nervous system. They can be demonstrated readily on mature lymphocytes, granulocytes, monocytes, platelets, epithelial cells, endothelial cells and fibroblasts grown in culture. However, they are not found as normal components of the surface of mature human erythrocytes.

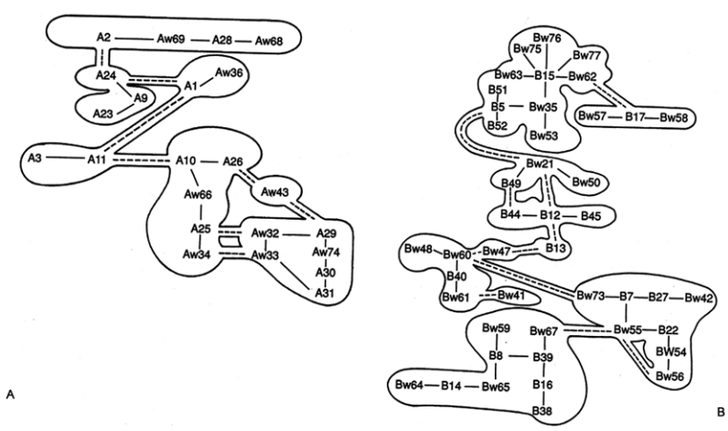


FIGURE 35.3. The prominent cross-reactive groups (CREGs) of the HLA antigens. The CREGs of the HLA-A antigens are shown in **A** and those of the HLA-B antigens in **B**. The antigens form tightly bound clusters that are strongly cross-reactive. In addition, some weaker cross-reactivity shown by the dotted lines exists between the clusters.

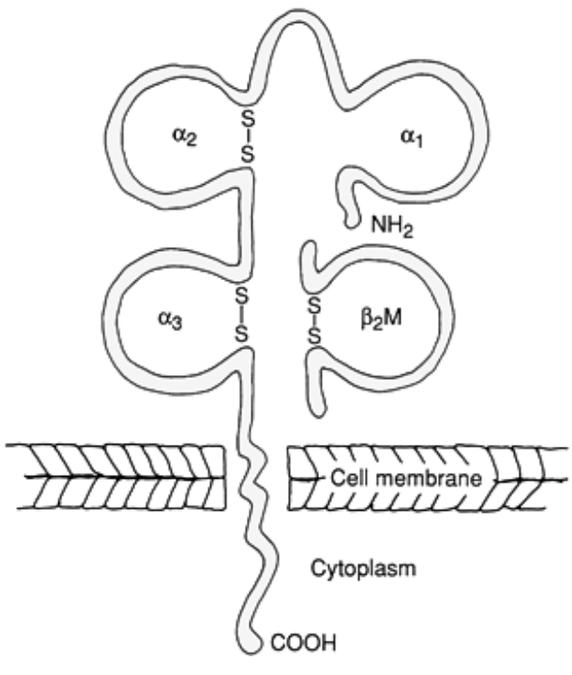


FIGURE 35.4. A diagrammatic representation of the HLA class I molecule. The α chain carries the specificity. It is always found in association with ß2-microglobulin.

Soluble HLA Antigens (sHLA)

There have been numerous reports that class I antigens may be found on red blood cells in humans just like their murine counterparts, the H-2 antigens. These observations are best explained by the finding that free HLA antigens are present as soluble substances in the plasma. Most likely they are shed from the surface of nucleated cells and maybe even from maturing erythroblasts. Occasionally, these soluble class I HLA antigens may become attached to the surface of mature red cells as they float freely in the plasma (16). It is highly unlikely that the antigens can be synthesized by anuclear mature red cells.

The amount of soluble HLA in blood transfusion products is significant. Red cells tested 30 days after collection and platelets from a single blood unit were shown to contain significant concentrations of HLA class I molecules. Such preparations completely inhibited mixed lymphocyte reactions and the cytotoxic activity of alloreactive T cells, suggesting an important biological role for sHLA molecules. However, the addition of anti-HLA class I monoclonal antibodies to the red blood cell preparations abolished their inhibitory effect. In addition, these molecules have a very practical value in the laboratory; forming the basis of a series of ELISA based techniques that can be used for antibody detection. These will be described in Chapter 36, Molecular HLA Typing.

HLA Class I and Class II Molecular Subunits and Three-Dimensional Structures

The class I MHC molecules, Fig. 35.4, consist of heterodimers composed of two, noncovalently linked polypeptide chains. The larger of the two, the α chain, is glycosylated and has a molecular weight of approximately 45kD. It is produced as a direct result of genetic information within the MHC and bears amino-acid sequences that are part of the allospecific epitopes. The α chain is anchored in the lipid bilayer of the cell membrane. This form of attachment allows the molecule to move freely on the surface of the cell and is loose enough to permit shedding.

The smaller, ß2-microglobulin chain does not penetrate the cell surface. It has a molecular weight of 12kD. Its production in man is controlled by nonpolymorphic genes found on chromosome 15. Thus, the two polypeptides that form class I HLA antigen are produced independently and associate posttranslationally to produce the final molecule. The presence of ß2-microglobulin is required to maintain the stability of the class I molecule and it has been found to be an integral part of such molecules from all species studied.

An analysis of the amino-acid sequence of the α chain has shown that it is divided into five domains consisting of three extracellular regions, α1, α2, and α3, a transmembrane domain, and a cytoplasmic domain. The α2 and α3 domains each have intrachain loops held by disulfide bonds having strong sequence homology to immunoglobulin domains. The α1 domain is not held in a loop and contains a glycosylation site at position 86. The amino-acid sequencing of α chains from MHC antigens of different specificities and even from different species shows a remarkable degree of conservation in some of the domains (17). Most of the variability occurs in the α1 and α2 domains, which are the regions expressing allospecificity. The transmembrane domain consists primarily of hydrophobic amino acids with some highly charged residues that may help anchor the molecule. The cytoplasmic domain contains serine residues that may be phosphorylated and could be involved in intracytoplasmic signaling. The associated ß2-microglobulin consists of a single domain formed by one disulfide loop. It is associated with the heavy chain through interaction with the α3 domain.

The class I molecule has been crystallized and its structure deduced by radiograph diffraction and computer-assisted analysis. (18) The resultant image shown diagrammatically in Fig. 35.5 is a globularlike structure containing an immunoglobulinlike groove that could act as a peptide-binding site. The α1 and α2 domains combine to form a structure composed of two a helices that form the walls of the groove and a ß-pleated sheet that forms the floor.

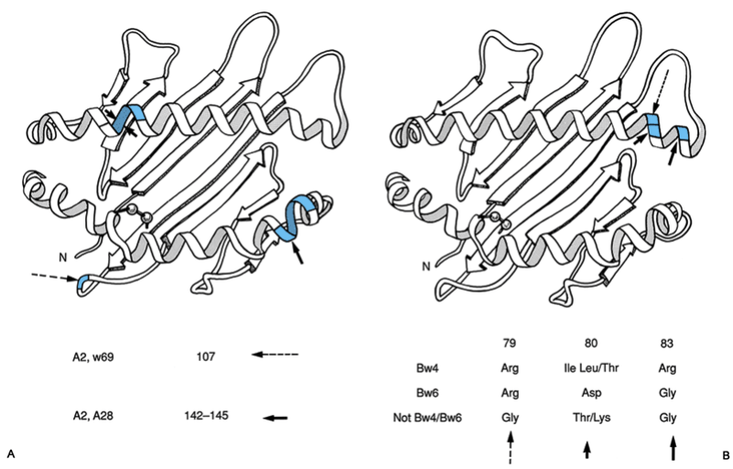


FIGURE 35.5. Diagrams of the structure of HLA molecules derived from the study of crystals showing the location of amino acid residues comprising HLA-specific epitopes. The epitopes of HLA-A2 are shown in **A**, and the amino acids forming HLA-Bw4 and Bw6 are shown in **B**. (Reprinted from Parham P, Lawlor DA, Slater RD, et al. HLA-A, B, C: patterns of polymorphism in peptide binding proteins. In: Dupont B, ed. *Immunobiology of HLA, II*, New York: Springer-Verlag, 1989:17.)

Amino-acid sequence analysis of the α chains of class I molecules of different HLA specificities has revealed the presence of at least 20 residues of high variability (17). Most of these differences are localized in exposed surfaces of the outer domain of the molecule, and in the peptide-binding groove in particular. The polymorphic residues that line the peptide-binding groove determine the peptide-binding specificity of the different HLA molecules.

Peptides bind to HLA class I molecules through specific anchor residues. The amino-acid side chains of these residues anchor the peptides by binding in pockets that line the peptide-binding groove. Polymorphism in HLA class I molecules affects the amino acids lining these pockets and thus their binding specificity. As a consequence, the anchor residues that bind to each allele are different. The set of anchor residues that allows binding to given antigen is called a sequence motif.

Linear sequences identifying residues critical to the expression

of allospecificities are found in the helices of the α1 and α2 domains as well as external portions of the ß-pleated sheet. Other polymorphic amino-acid residues that cannot be defined by alloantibodies can be identified by DNA-based techniques as discussed in Chapter 36, Molecular HLA Typing.

These findings explain certain phenomena that confounded serologists working in the HLA system. The public specificities Bw4 and Bw6 are the result of amino-acid substitutions in residues 79 to 83 of the α1 domain. Asparagine at position 80 is essential for the expression of the Bw6 epitope while arginine at position 83 is responsible for the Bw4 specificity. These amino-acid positions are different from those that confer the private specificities. Therefore, the so-called inclusion of private specificities within public ones can be explained because the HLA molecule has several epitopes in different locations.

The finding that within the HLA-A2, A28, Aw69 CREG a combination of three epitopes is required for the expression of HLA-A2, could be the basis for cross reactivity. While all three epitopes are required for the expression of A2, the expression of A28 depends only on residues at positions 142 to 145 and Aw69 at position 107.

In contrast to class I antigens, the class II antigens that include the products of the HLA-DR, DQ, and DP loci are limited in their distribution. In general, these antigens do not appear to be constitutive products of nucleated cells. Rather they are produced in conjunction with an immune response by means of lymphokines such as IL-2 and γ-interferon.

Most regularly, the class II antigens are found on B-lymphocytes, monocytes, macrophages, endothelial cells, Langerhans cells and activated T-cells. However, they also may be induced in other types of cells such as renal epithelial cells and pancreatic ß cells by means of lymphokines. The latter finding has led to the formation of an hypothesis regarding a possible mechanism for autoimmunization with subsequent tissue damage and disease.

Class II antigens, Fig. 35.6, are heterodimers consisting of two noncovalently linked polypeptide chains, each of which is a product of separate genes in the HLA region. The chains each have two external domains, a transmembrane portion anchored in the lipid bilayer of the cell wall and a cytoplasmic domain. One of the chains, the α chain, has a molecular weight of 34kD and only its α2 domain is arranged as disulfide loop. It bears structural homology to the constant region of immunoglobulin. The α chains of DR molecules are not polymorphic. Therefore, all of the DR specificities share identical α chains. However, the α chains of DQ molecules have some degree of limited polymorphism.

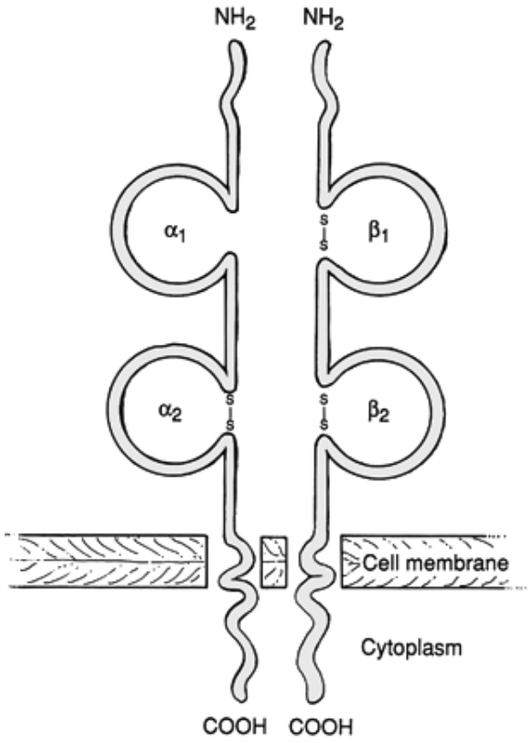


FIGURE 35.6. A diagrammatic representation of the HLA class II molecule. The carboxy-terminal ends of both the α and the ß chains are embedded in the lipid bilayer of the cell membrane.

The second polypeptide or ß chain has a molecular weight of 29kD. It is highly polymorphic in all class II molecules and carries sequences appropriate to the allospecificity in three hypervariable regions of its ß1 domain situated between residues 9–13, 26–33, and 67–74. In contrast to the α chain, both the ß1 and ß2 domains contain disulfide loops. Glycosylation sites are found in the α1, α2 and ß1 domains, which are not involved in the allospecificity of the intact molecule (19).

The crystal structure of the HLA-DR molecule shows that it is folded very much like the HLA class I molecule. The major differences lie at the ends of the peptide-binding cleft, which are more open in the HLA-DR molecules. The main consequence of this is that the ends of a peptide bound to any class I molecule are substantially buried within the molecule, wheras the ends of a peptide bound to HLA-DR are not (20). Different allelic variants of HLA class II molecules also bind different peptides; but the more open structure of the HLA class II peptide-binding groove, and the greater length of the peptides bound in it, allow a greater variety of peptide binding. It is therefore more difficult to predict specific peptides that might bind to HLA class II molecules.

Within the DR region, the DRB1 gene encodes for the DRß1chain, which determines the DR private specificities DR1, DR2, DR3, DR4, DR5, etc. The DRB3, B4, and B5 genes encode for the DRß3, ß4, and ß5 chains, which determine the DR52, DR53, and DR51 specificities, respectively. Originally, the DR52 and DR53 specificities were thought to be supertypic or public antigens similar to the class I antigens Bw4 and Bw6 because they were always found along with one of the other DR specificities. DR3, 5, 6, and 8 associate with DR52 and DR4, 7 and 9 with DR53. This is not the case. Rather, the observed phenotypic association occurs because of the tight linkage and almost absolute gametic association between the two ß chain loci involved (21).

The DQ antigens were likewise thought to be public specificities of DR because DR1, 2, and 6 are associated with DQ1; DR3, 7 with DQ2 and DR4, and 5 with DQ3. However, this too is the result of gametic association.

The posttranslational binding of the α and ß chains in the formation of the DQ heterodimers offers some interesting possibilities. In the heterozygous individual, there are two α chain genes for DQ, one from each parent. These may be different because the responsible genes exhibit at least limited polymorphism. In addition, there are two different DQ ß chain genes. It is possible for the products of genes in the *trans* position to associate posttranslationally as shown diagrammatically in Fig. 35.7. Therefore, a paternal DQ α chain could associate with a maternal DQ α chain, giving rise to a hybrid molecule. Such a molecule would display the maternal DQ allospecificity because this is controlled by the ß chain. However, the specificity of this molecule arising from the complementarity of the two chains could be entirely unique to antigen recognition cells and contribute to further polymorphism.

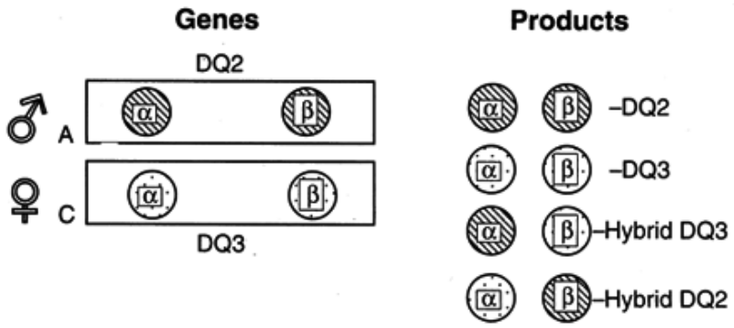


FIGURE 35.7. A scheme showing how the posttranslational combination of DQ α and ß chains in a DQ-heterozygous individual can result in four different molecules.

The early studies of the HLA system additionally defined a set of antigens that could only be detected in a biological assay called the mixed lymphocyte culture (MLC). These antigens correlated with polymorphic class II antigens when analyzed in families and were called HLA-D. A whole list of these specificities was identified and many are still found in the table of current antigens. It is now recognized that the HLA-D specificities are controlled by the DR alleles. However, they may be different than those recognized serologically.

It is clear that immune cells are more robust than antibodies in their ability to detect very small molecular differences. Thus two DR ß chains coming from two different individuals and having identical DR epitopes but differing by only a single amino-acid substitution in another part of the molecule, could not be distinguished by anti-DR antibodies. However, the cells expressing these chains could be recognized by other cells. Differences such as this and their impact on class II antigen detection will be considered in the section dealing with biochemical techniques.

BIOLOGICAL FUNCTION OF HLA GENES

Part of "35 - HLA: Structure, Function, and Methodologies"

The biologic function of the molecules of the MHC has been the subject of a great deal of basic research. In studying the biology

of the immune response in rodent systems, it became clear that the response to certain well-defined antigens was under genetic control (22). Furthermore, it could be demonstrated that the genes responsible for this function could be mapped to the MHC region of the chromosome.

In vitro experiments designed to study the immune response at the cellular level showed that both antigen presentation and antigen destruction by immunologically transformed cytotoxic cells was restricted by the MHC antigens (23). For example, cytotoxic lymphocytes induced in response to challenge with influenza virus can kill only other cells infected with such a virus if two conditions are met. First, the virus has to be antigenically identical with the one used for induction. Second, the MHC of the infected cell to be destroyed has to be identical to the MHC of the cell that harbored the immunizing virus. The primary biologic role of the class I antigens that make up the HLA-A, -B and -C repertoire described in this section is to function in such a capacity. Their behavior as alloantigens in transplantation by serving as the targets of the allograft response is secondary activity of iatrogenic origin.

Class II molecules function biologically in antigen presentation. *In vivo*, pathogens and other foreign substances are processed by antigen-presenting cells and delivered to specific T-cell receptors on immunologically competent lymphocytes in the context of class II antigen. The T-cell receptor must recognize both the foreign antigen and the self class II molecule in order to transmit a signal. Therefore, class II antigens play a very important role in the initiation of the immune response. Furthermore, in this context, they can act in the capacity of controlling the response.

Effects of HLA Polymorphism on Antigen Recognition by T-cells

HLA polymorphism affects antigen recognition by T cells, both indirectly, by controlling peptide binding, and directly, through contacts between the T-cell receptor and HLA molecule itself. Because the HLA genotype restricts the antigen specificity of T cells, this effect is called MHC restriction. MHC restriction is a critical feature of antigen recognition by most functional classes of T cells.

Some of the polymorphic amino acids on the HLA molecules are located on the a helices flanking the peptide-binding cleft in such a way that they are exposed on the outer surface of the peptide-HLA complex and can be contacted directly by the T-cell receptor. Therefore, when T cells are tested for their ability to recognize the same peptide bound to different HLA molecules, they readily distinguish the peptide bound by HLA^{a} from the same peptide bound to HLA^{b}. Thus specificity in a T-cell receptor is defined by both the peptide and by the HLA molecule binding it. This restricted recognition can be caused by differences in the conformation of the bound peptide imposed by the different HLA molecules rather than by direct recognition of polymorphic amino acids on the HLA molecule itself. Thus HLA restriction in antigen recognition reflects the combined effect of the peptide binding, and of direct contact between the HLA molecule and the T-cell receptor (24).

HLA Polymorphism Determines the Range of Antigens to Which the Immune System Can Respond

As different HLA molecules bind different peptides, the T cells responding to a given protein antigen presented by several different HLA molecules usually will be able to recognize different peptides. In rare cases, a protein will have no peptides with a suitable motif for binding to any of the MHC molecules expressed on the cells of an individual. When this happens, the individual fails to respond to the antigen.

As pointed out previously, such failures in responsiveness to simple antigens were first reported in inbred animals, where they were called immune response (Ir) gene defects. Rodent model systems clearly demonstrated that these genes can be mapped to the MHC and were the first clue to the antigen-presenting function of the MHC molecules. Subsequently, it was shown that the Ir genes encode MHC class II molecules. Ir genes and associated defects are easily recognized in inbred strains of mice because the mice are homozygous for all their MHC genes and thus express only one allelic variant from each locus. In humans, the polymorphism of HLA molecules guarantees a sufficient number of different HLA molecules in a single individual to make this type of nonresponsiveness unlikely.

The polymorphism at each HLA locus can potentially double the array of different HLA molecules expressed by an individual; because most individuals are heterozygotes. Polymorphism offers the additional advantage that different individuals in a population will differ in the combination of HLA molecules they express and will therefore present different sets of peptides from each pathogen. This makes it unlikely that all individuals will be equally susceptible to a given pathogen and its spread will therefore be limited. That exposure to pathogens over an evolutionary time scale can select for expression of particular HLA alleles is indicated by the strong association of the HLA-B53 allele with recovery from a potentially lethal form of malaria. This allele is very common in the people from West Africa, where malaria is endemic and rare in parts of the world where malaria is uncommon.

Recognition of Allogeneic HLA Molecules by T Cells

The discovery of MHC restriction has helped explain the phenomenon of nonself MHC recognition in graft rejection. Transplanted tissues or organs from donors bearing HLA molecules that differ from those of the recipient are reliably rejected. The rapid and very potent cell-mediated immune response to the transplanted tissues results from the presence in any individual of large numbers of T cells that are specifically reactive to particular nonself or allogeneic HLA molecules. *In vitro* MLC studies have shown that 1% to 10% of all T cells in an individual will respond to stimulation by cells from any allogeneic individual. This type of response is called alloreactivity because it represents recognition of allelic polymorphism on allogeneic HLA molecules.

Before the role of MHC molecules in antigen presentation was understood, it was a mystery why so many T cells should recognize

nonself MHC molecules. However, once it was learned that T-cell receptors have evolved to recognize foreign peptides in combination with polymorphic MHC molecules, alloreactivity became easier to explain. Experiments with T cells from animals lacking MHC class I or class II molecules have shown that the ability to recognize MHC molecules is inherent in the genes that encode the T-cell receptor, rather than being dependent on selection for MHC recognition during T-cell development. It is thought that the alloreactivity of mature T cells reflects the cross-reactivity of T-cell receptors normally specific for a variety of foreign peptides bound by self-MHC molecules.

The Class III MHC Gene Products

The class III MHC gene products include complement components: C2, Factor B, and C4 and the cytokines: TNF-α and -ß (lymphotoxin) that have important functions in the immune response. The function of these genes will be discussed in Chapter 37.

CLASSICAL TECHNIQUES FOR DETECTING HLA ANTIGENS

Part of "35 - HLA: Structure, Function, and Methodologies"

Serologic Techniques

The classical serologic techniques for the detection of HLA antigens of class I are based on the microlymphocytotoxicity method originally developed by Terasaki and McClellan (25, 26). This is a complement-mediated serologic assay in which peripheral blood lymphocytes are first mixed with an antiserum containing specific anti-HLA antibodies. After a period of incubation, complement is added and the mixture further incubated. If the lymphocytes carry the HLA antigen against which the antibodies in the antiserum are directed, an antigen-antibody reaction will occur, which activates the complement. The ensuing enzymatic cascade culminates in membrane damage that ultimately kills the lymphocytes. A positive reaction is ascertained by examining the cell suspension for cell death. This can be accomplished in a number of ways including phase contrast microscopy coupled with the use of supravital stains.

A number of variations of this general method have been developed to increase its sensitivity and/or simplify and improve the method used to distinguish live cells from dead ones. These will be described later.

The target cells of choice for HLA typing are the small lymphocytes of the peripheral blood. These cells fully express all of the class I antigens. They can be isolated and purified and easily remain stable during *in vitro* manipulation.

Separation is accomplished by layering diluted whole blood over a solution of Ficoll-Hypaque that has a specific gravity corresponding to the density of small lymphocytes. This is 1.077 to 1.080. The resultant gradient is centrifuged allowing the mononuclear cells consisting primarily of lymphocytes to settle at the Ficoll-blood interface. The erythrocytes and granulocytes being denser settle to the bottom of the tube while most of the platelets remain floating in the uppermost layer (Fig. 35.8).

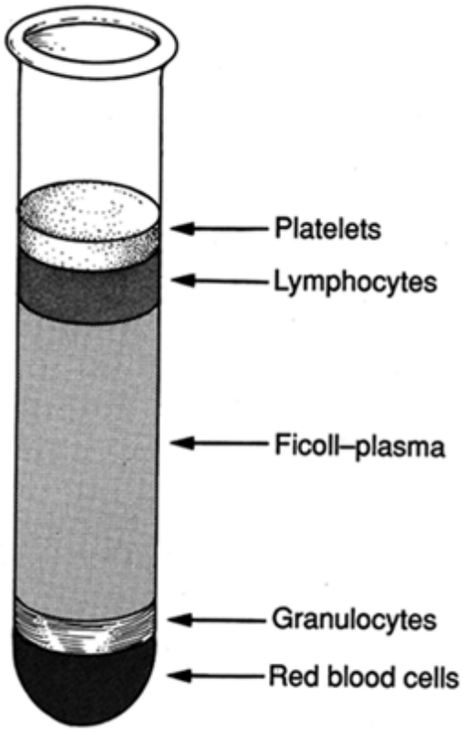


FIGURE 35.8. The gradient formed after the centrifugation of whole blood layered on a mixture of Ficoll-Hypaque. (From White TJ, Arnheim N, Erlich HA. The polymerase chain reaction:L the basic technique. *Trends Genet* 1989;5:6.)

Isolated small lymphocytes consist of a mixture of T cells and B cells. Such a suspension is perfectly adequate for the detection of class I HLA-A, -B, and -C antigens because these are well represented on all subsets of lymphocytes. However, in order to identify Class II HLA-DR and DQ antigens that are expressed primarily on B cells, an enrichment or purification of this subpopulation is required.

A number of techniques are available for the purification of B cells (27). All take advantage of one or more specific properties of T cells and B cells. These include the ability of T cells to form rosettes with sheep erythrocytes, and the adhesive properties of B cells. Monoclonal antibodies to T cell differentiation antigens together with complement also can be used to purify B cells.

Source of Typing Sera

The antisera used for HLA typing are available from a number of commercial sources both as bulk reagents and on prepared typing trays. The prepared trays are the most economical because they contain a wide range of antisera capable of detecting the majority of clinically relevant antigens. In addition, much of the initial quality control on the finished product has been performed by the manufacturer. Bulk sera, on the other hand, need to be selected according to some predetermined criteria based on their specificity and then dispensed into the typing trays. Finally, extensive primary quality-control procedures need to be carried out in order to ensure that the reagents are working and that the assortment of sera selected can accurately identify an individual's HLA phenotype.

The majority of typing sera used for the detection of class I and class II HLA antigens are derived from the sera of multiparous women. HLA antibodies are produced frequently as a result of alloimmunization during pregnancy (28, 29). Approximately 25% of para iii or greater have them in their serum. This event does not result in any harm to the fetus even though these antibodies are almost exclusively IgG and capable of crossing the placenta. Most

likely the protection is afforded by the placenta, which carries the fetal tissue antigens and acts like a sponge to absorb the antibodies. In fact, placentae are a rich source of HLA antibodies.

Anti-HLA antibodies may be produced as the result of other forms of exposure to antigenic stimuli including transplantation and blood transfusion. The method developed for the production of monoclonal antibodies has been employed to make some anti-HLA reagents (30). Although this procedure was initially fraught with difficulty, it currently shows very promising results. The clinically significant allospecific epitopes on the antigen molecules have been better defined and the procedure for monoclonal antibody production including antigen preparation and clonal selection has become more sophisticated. As a result, some of the newer reagents appear to be quite specific in their ability to distinguish clinically significant HLA types; yet they offer all of the advantages of a monoclonal reagent. For example, they are specific for either class I or for class II antigens. Thus, B-cell separation is not required to type for HLA-DR. In addition, they are free of many interfering substances found in human sera containing alloantibody; thereby allowing for the development of novel and perhaps more rapid techniques for HLA typing.

The antisera used for HLA typing regardless of their source are quite different than the alloantisera used for typing in other systems such as the red cell blood groups. One of the vagaries associated with them is illustrated by the observation that over the years very few sera directed apparently against the same specificity have been shown to be exactly identical when tested on a cell panel from a large number of donors. In other words, each serum is bound to give both false-negative and perhaps false-positive reactions with the lymphocytes of various individuals. Thus, because of the nature of the antisera, several examples of each must be used for the optimum definition of each specificity.

Over the years, various descriptive terms have been coined to describe these phenomena. Some of them are derived from the habit of most investigators to conveniently represent the reactivity of a serum by means of a histogram or serogram as shown in Fig. 35.9. The test cells are listed along the side, the sera across the top, and each block represents the reaction of a single serum with a single cell. The positive reactions are represented by a solid block while the negatives are assigned by a dash. The cells are arranged so that those expressing the key antigen, graphed on the left side of the diagram, are at the bottom of the chart. In this type of an illustration, a serum giving false-positive reactions will cause its serum bar to be longer than the occurrence of the antigen. Therefore, such sera were often referred to as “LONG.” On the other hand, sera that tend to give false-negative reactions result in reaction bars that are shorter than the occurrence of the antigen. These sera were called “SHORT.”

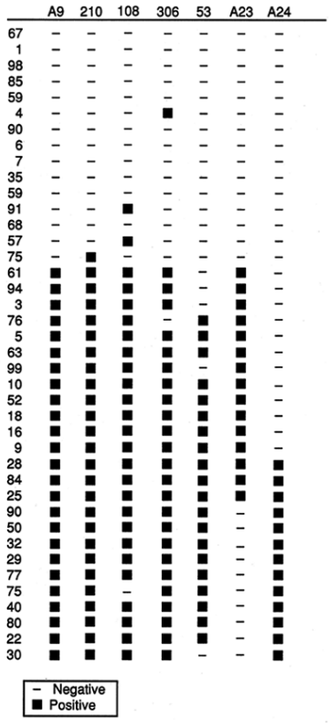


FIGURE 35.9. A serogram showing the reaction patterns of four sera reacting with cells positive for HLA-A9. The occurrence of the A23 and A24 antigens among the cells of the panel is shown in the last two columns. Serum 108 is “long” and serum 53 is “short” compared to the index specificity, HLA-A9.

Absorption studies performed using the cells giving the false-negative reactions with short sera often showed that they were able to fully absorb the specific antibody from such sera. Still, they failed to give a positive cytotoxicity reaction. A finding of this type has been called CYNAP (cytotoxicity negative, absorption positive). This phenomenon may be the result of a number of factors including poor antigen expression on the cell at the time of the test, low antibody affinity, or the presence of anticomplementary factors in the cell preparation or the antiserum containing the antibodies.

Fig. 35.9 also shows the reaction patterns of the public specificity HLA-A9 along with the private specificities HLA-A23 and A24. The pattern is such that the private specificities are “INCLUDED” in the public one. The practice of resolving the reaction patterns of anti-HLA sera into serograms has resulted in a

wealth of information relative to the relationship of the various specificities to each other. It formed the basis from which many of the features of the system could be defined and served as a template for further studies at the level of molecular biology and biochemistry.

The complement source used in the microlymphocytotoxicity test is normal rabbit serum. The sera from a large number of different animal species have been tested for their usefulness in cytotoxicity and the rabbit was found to be the best. This is because of the fact that normal rabbits naturally produce a small amount of antihuman antibody in response to environmental stimuli. These antibodies act in synergy with anti-HLA antibodies to promote rapid and efficient complement activation. Currently, a number of commercial sources supply excellent rabbit complement. As a matter of fact, some suppliers of reagent anti-HLA sera provide complement specifically tailored for use with their reagents. Still, it is considered good laboratory practice to quality control each new lot of complement. In this way, consistency of testing and therefore typing results can be assured. In addition, adequate positive and negative controls must be employed with each test.

The standard microlymphocytotoxicity test is performed in specially designed plastic trays containing wells with optically flat bottoms (Fig. 35.10). In practice, the wells first are filled with biologically inert mineral oil, which is used to retard the evaporation of the minute volumes of reactants. One microliter of antiserum-containing antibodies is added to the bottom of the well followed by 1 µL of cell suspension containing approximately 2,000 cells.

FIGURE 35.10. A standard 55 × 80 mm tray used for HLA typing. The reactions are carried out in the individual wells, whose contents can be observed from below by means of an inverted microscope. (Courtesy of Robbins Scientific Corporation, Sunnyvale, CA.)

This mixture is incubated at room temperature for 30 minutes to allow the antibodies in the serum to react with antigens on the cell surface. At the conclusion of the incubation, 5 µL of rabbit complement are added and the reaction further incubated for 60 minutes at room temperature. During this period, complement is activated by the antigen-antibody complexes at the cell surface resulting in cell death. At this point, a solution of Eosin-Y dye is added to stain the cells and make the test easier to read. Finally, a solution of neutralized formalin is added to stop the reaction and preserve the architecture of the cells.

The reaction is read using an inverted microscope equipped with phase contrast illumination. Under these conditions, live lymphocytes appear as clear, glistening, small, round cells not too different from red cells. For this reason, it is imperative to remove as many red cells from the final cell suspension as possible. The red dye used for staining penetrates the cells but can be excreted by live ones giving them a golden tone on a light red background. The dead cells cannot excrete the dye and therefore stain with an intense maroon color. They appear as dark, opaque, swollen cells with prominent nuclear detail (Fig. 35.11).

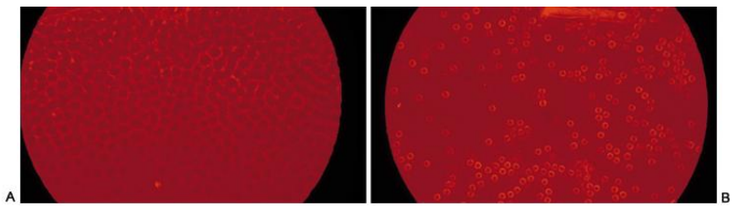


FIGURE 35.11. Microscopic appearance of the lymphocytotoxicity reaction. **A:** A positive reaction. The cells have taken up the eosin Y dye. They are swollen and have prominent nuclear detail. **B:** A negative reaction. Most of the cells appear clear, glistening, and refractive.

The reaction is scored by determining the percentage of cell death in each well. In practice, skilled workers estimate the percentage of cell death by trained optical-pattern analysis. The results are expressed according to the grading scheme shown in Table 35.6. A score of eight is maximum but the reaction is considered positive with a score of six. A negative is defined by a score of one.

TABLE 35.6. CYTOTOXICITY SCORING

Score	% Cell Death	Interpretation
1	0–10	Negative
2	10–25	Possibly negative
4	25–50	Possibly positive
6	50–80	Positive
8	80–100	Strong positive

In the performance and the reading of this type of test, which depends so heavily on a biological end point, there is always the possibility of some doubtful reactions. These are indicated by the scores of four and two. A score of four usually is interpreted to mean, “I don't know; but I think it's positive.” A score of two means, “I don't know; but I think it's negative.” Although this may seem to be somewhat subjective, it does have precision. Skilled workers are remarkably consistent in coding these reactions and very little if any variation is observed between duplicate readings performed by the same or by different individuals.

More Sensitive Microlymphocytotoxicity Tests

There are a number of variations to this basic test, which were introduced in attempts to make it more sensitive (31). The first of these, which is known as the Amos Mod technique, calls for the cells to be washed with tissue-culture fluid after their initial incubation with antiserum. This modification is based on the fact that the source of antiserum is frequently a patient whose serum may contain factors that bind complement nonspecifically. These could be aggregated gamma globulins, antigen-antibody complexes, or even the products of bacterial contaminants in stored sera. Any of these materials could deplete the system of complement and lead to false-negative reactions. Therefore, their elimination prior to the addition of complement effectively increases sensitivity.

Another modification that increases the sensitivity of the standard microlymphocytotoxicity test is based on the use of an anti-human globulin reagent (32). The reagent of choice in this case is directed against the light-chain of human IgG. An anti-κ containing serum is the most popular. The underlying principle of this modification is different from that of classical antihuman globulin or Coombs tests used in red-cell serology in which antiheavy-chain reagents are used. In this case, the antiglobulin serum serves as a source of additional Fc fragments to interact with those on bound anti-HLA to help activate complement rather than as an antibody bridge to promote agglutination. Using this approach, it is possible to achieve complement activation even with a relatively small number of anti-HLA antibody molecules bound per cell; thereby giving extreme sensitivity.

These procedures are most useful in those cases in which patients' sera are used as a source of potential anti-HLA antibody, such as in cross-matching, described in Chapter 37, HLA: The Major Histocompatibility Complex-Applications. However, this and other modifications should be avoided when using commercial typing sera that have been standardized according to routine methods. The direction circulars accompanying such products should be followed exactly.

Other Microlymphocytotoxicity Tests

A totally different approach to determining cell death in the microlymphocytotoxicity test is given by a technique called cytofluorochromasia (33). This technique is based on the fact that living lymphocytes possess esterases in their cytoplasm. One of these enzymes is capable of converting the colorless compound, fluoresceine diacetate, into fluoresceine, which glows with a green fluorescence under ultraviolet (UV) light. In the test procedure, the lymphocyte suspension is pretreated with fluoresceine diacetate before any exposure to antibody. Then the cells are incubated, first with antibody and then with complement, in the usual fashion. The reaction is read with an inverted microscope equipped with a UV light source and the appropriate excitation filters. Under these conditions, only the live cells can be seen glowing with an apple-green fluorescence. The fluorescent compound leaks from the dead cells and leaves them colorless. Therefore, the degree of fluorescence is proportional to the number of live cells remaining. A totally dark well indicates 100% cell death and a positive test. A negative reaction is given by a well filled with glowing cells. To improve the accuracy of reading and subsequent interpretation, the dead cells can be counterstained with ethidium bromide, a red fluorescing compound that can penetrate the damaged membranes of dead cells and stain their nuclei.

An interesting variation of this technique was developed by van Rood and his colleagues, which can be used to test for class II HLA antigens without the bother of preparing suspensions of pure B cells (34). The principle is based on the fact that B cells express immunoglobulins on their surface.

A cell suspension prepared from peripheral blood containing both B cells and T cells is treated with a predetermined dose of a reagent containing goat antihuman immunoglobulin antibodies, which have been conjugated with fluorescein diisothiocyanate. This reagent reacts with the immunoglobulins expressed on the B cells, effectively coating them with a fluorescent compound. In addition, the reagent cross-links the immunoglobulins expressed on the cell surface causing them to form a cap that migrates to one pole of the cell.

The microlymphocytotoxicity test is performed in the usual way with the exception that ethidium bromide is added as the final stain instead of eosin. This dye will penetrate the damaged walls of dead cells and will stain their nuclei. Live cells having intact membranes exclude this material leaving their nuclei unstained.

The reaction is read using a UV microscope equipped with filters capable of exciting both green and red fluorescence. Under these conditions, dead B cells will fluoresce red with green caps, live B cells will have only green caps, dead T cells will be red but with no caps, and live T cells will be colorless (Fig. 35.12).

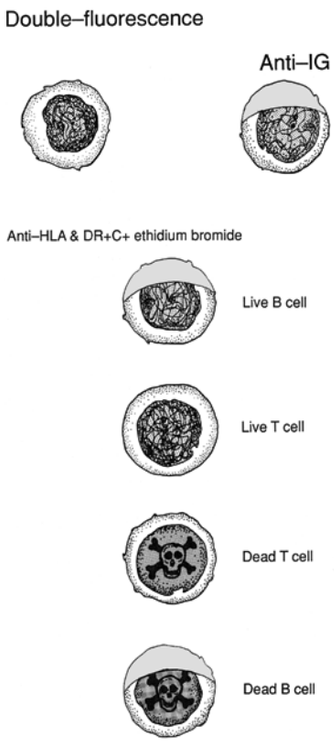


FIGURE 35.12. Double-fluorescence microlymphocytotoxicity.

The fluorescent methods offer no advantage over the standard cytotoxicity procedures with respect to sensitivity and carry the disadvantage of requiring specialized equipment. However they are somewhat easier to read by the untrained and form the basis for automated readers that employ photomultipliers for detecting fluorescence intensity. Using this type of instrumentation, the readings can be directed to computer programs that automatically assign the serologic HLA types.

Importance of Quality Control in HLA Typing

As with any laboratory procedure, meticulous attention must be paid to quality control. In the microlymphocytotoxicity system, the various components are especially fragile and susceptible to denaturation. Care must be excercised in making cell suspensions to ensure maximum viability at the outset. Collection procedures call for room-temperature storage and handling of blood specimens collected in preservative free heparin or ACD.

The anti-HLA antibodies found in the sera of patients and specific reagent donors are for the most part IgG immunoglobulins. Although examples of IgM antibodies have been described, they are usually the exception rather than the rule. IgG antibodies are the products of a well-developed, mature immune response that has completed undergoing an isotype switch.

They are remarkably stable molecules and under ideal experimental conditions have been shown to retain their activity even after storage at room temperature for several days. The storage conditions imposed on antisera are not in place to preserve the antibodies from denaturation, but rather to minimize the formation of “anticomplementary” factors consisting of substances such as aggregated gamma globulins, protein complexes, and microbial products that can interfere with the complement-dependent cytotoxicity reaction. Therefore, in order to assure maximum reactivity, storage temperatures need to be carefully monitored and maintained within the specified limits.

Controls need to be used with each test to assure a minimum viability of 90% at the start, and maximum viability with a negative control serum. A positive control must be included to assure that the complement is working and that the cell suspension being used is susceptible to its action.

Complement must be controlled carefully. Each new lot of complement should be compared with a previous lot to determine its optimum working titer. This assay should be performed using several antisera having different specificities and antibody titers to demonstrate that the new source of complement works with equal facility with both strong and weak sera. In addition, because complement has a natural denaturation time, which is related to temperature, storage temperatures for this reagent must be monitored carefully.

The anti-κ antiserum used in the antiglobulin modification of the standard procedure needs to be controlled in much the same manner as complement. Each new lot should be tested along with a previously used working lot to determine its optimum working dilution. It should be aliquoted in small volumes to avoid repeated freezing and thawing.

Finally, in performing quality control on the cytotoxicity procedure, the proficiency of the operator performing the test cannot be overlooked. This test is highly dependent upon individual skills, especially those required for reading the test. These skills must be monitored carefully by the periodic performance of blind repeats of HLA typings and the duplicate reading of tests by more than one individual. Readings should differ by no more than a single score value.

Limitations to Serologic HLA Typing in Non-Caucasian Populations

Since the 1972 International Histocompatibility Workshop, it has been appreciated that HLA alleles exist specific to particular ethnic groups. However, serologic typing with its reliance on white alloantisera has tended to underestimate ethnic differences and the number of such alleles. Thus, the serologic typing does not detect a large number of alleles found in human populations, and it is in nonwhite populations that the inaccuracies and insensitivities

TABLE 35.7. SOME HLA-DR AND HLA-D EQUIVALENTS AND THE RESPONSIBLE ALLELES[a]

HLA-DR Specificity	HLA-D Specificity	HLA-DR B Allele
DR1	Dw1	DRB1*0101
DR1	Dw20	DRB1*0102
DR15(2)	Dw2	DRB1*1501
DR15(2)	Dw12	DRB1*1502
DR16(2)	Dw21	DRB1*1601
DR16(2)	Dw22	DRB1*1602
DR17(3)	Dw3	DRB1*0301
DR4	Dw4	DRB1*0401
DR4	Dw10	DRB1*0402
DR4	Dw13	DRB1*0403
DR4	Dw14	DRB1*0404
DR4	Dw15	DRB1*0405
DR11(5)	DW5	DRB1*1101
DR11(5)	Dw “FS”	DRB1*1104
DR12(5)	Dw “DB6”	DRB1*1201
DR13(6)	Dw18	DRB1*1301
DR13(6)	Dw19	DRB1*1302
DR14(6)	Dw9	DRB1*1401
DR14(6)	Dw16	DRB1*1402
DR7	Dw17	DRB1*0701
DR8	Dw8.1	DRB1*0801
DR9	Dw23	DRB1*0901
DR52	Dw24	DRB3*0101
DR52	Dw25	DRB3*0201
DR52	Dw26	DRB3*0301
DR53	Dw4, Dw10, Dw14, Dw15, Dw17, Dw23	DRB4*0101
DR51	Dw2	DRB5*0101
DR51	Dw12	DRB5*0101
DR51	Dw21	DRB5*0201
DR51	Dw22	DRB5*0202

[a]From Bodmer JG, Marsh SGE, Albert ED, et al. Nomenclature for factors of the HLA system, 1991. *Tissue Antigens* 1992;39:161–173

in serologic typing are the greatest. The extent of the ethnic diversification of HLA molecules will continue to be unveiled as population studies and molecular typing venture to various ethnic groups.

Several HLA specificities are difficult to define unambiguously with serology. The antigens in the B15 serologic group are an example of antigens inconsistently defined. In practice, an individual carrying the HLA-B*1508 allele might be assigned an HLA-B62 serologic type in one laboratory and an HLA-B75 serologic type in an another. Furthermore, defined antigens such as HLA-A74, -A80,-B81 and -B82 may lack serologic reagents for definition and may be missed.

In addition to alleles that are undetected, patterns of serologic cross-reactivity do not always reflect the underlying structural relationships between alleles. One example is HLA-B45, a common antigen in blacks. By serology, this antigen is closely related to HLA-B44 and the pair can be distinguished by their association with the Bw4 and Bw6 epitopes. In contrast, nucleotide sequences show that B44 and B45 are only distantly related, with B45 being much closer in structure to HLA-B50.

Serological Typing for HLA Null Class I Alleles

Although the difference between null and normal subtypes of an allele has the potential to produce major histoincompatibility in certain transplant situations, the overall relevance of HLA class I null alleles will depend upon their frequency within the population of patients and donors. The number of HLA null alleles that have emerged from DNA-based testing is impressive and seems to be preferentially found in particular populations (e.g., Native Americans and natives of Northern Ireland).

Comparison of serological and DNA typing is the best approach for the identification of HLA class I null alleles, and the need to define the null alleles provides a compelling reason for maintaining serological typing while the DNA-based typing methods are implemented and refined. The nucleic acid sequences that characterize some of these null alleles may not be in the genetic region normally targeted by DNA typing; thus, without serologic information, these null alleles may go undetected. However, once a null allele is identified by serology and characterized by nucleic acid sequencing to identify the source of the lack of expression, the null allele then can be readily detected in subsequent individuals by DNA testing procedures alone. The frequency of yet undetected null alleles is not known. It is expected to be low and will likely vary in different populations.

HLA class II Antigens Identified by the Mixed Lymphocyte Culture

The mixed lymphocyte culture (MLC) is a method that was used to investigate Class II HLA antigens. It is based on the observation that when populations of lymphocytes from two different individuals are mixed in culture, the cells from one individual are able to recognize foreign antigens on the cells of the other and to respond to these antigens by undergoing blastogenic transformation (35).

The antigens recognized in this reaction are Class II HLA antigens. Originally, they were called HLA-D and were believed to be recognizable only by lymphocytes. When the HLA-DR antigens were detected serologically on the surface of B lymphocytes, it became clear that they were the source of antigenic stimulation of MLC reactivity. Nevertheless, as will be described later, the serologically defined specificities alone could not fully explain cellular reactivity. This had to await until the true molecular diversity of the class II region was revealed by studying it at the DNA level.

In order for the MLC reaction to occur, there must be a difference between the HLA DR antigens of the two cell populations. Discrepancies in the HLA-A and/or -B antigens alone will not result in stimulation. However, cytotoxic T lymphocytes (CTL), which are directed against class I antigens, are generated as a by-product of this reaction.

In the normal MLC, 5 days of culture are required for the cells from a given individual to respond to the antigens of another individual and undergo blastogenesis. However, if this reaction is allowed to continue until all of the stimulating cells are expended, the multiplying responding cells return to a mature steady state. A rechallenge with fresh stimulating cells from the original donor will result in a brisk blastogenic response within the first 24 hours. This is a secondary response and takes place because the responding cells have been primed to recognize the stimulating antigen.

TECHNIQUES OF MOLECULAR BIOLOGY APPLIED TO HLA TYPING

Part of "35 - HLA: Structure, Function, and Methodologies"

Many new techniques, originally developed for use in molecular biology and described in detail in Chapter 36, have been applied recently to the study and elucidation of the HLA system. Some of these techniques have resulted in data that contribute significantly to a better understanding of the system. Others are in the process of being used for special applications in diagnostic and forensic laboratories.

One biochemical method shown to be very useful is two-dimensional gel electrophoresis, a technique that has been applied extensively to the investigation of the ß chains of DR molecules. It is possible to test a group of individuals whose cells express HLA-DR4; but who stimulate and respond to each other in the MLC. Electrophoretic and isoelectric focusing studies have shown that cells from individuals belonging to HLA-DR4 can be sorted by this method into groups corresponding to their reaction patterns in the MLC. In this method, lymphocytes are cultured in the presence of 35S methionine, which biosynthetically labels the HLA gene products. The harvested cells are lysed and their HLA molecules immunoprecipitated with monoclonal antibody directed against epitopes present on the constant region of the DR ß chain.

The immunoprecipitated molecules are first separated on a gel containing a pH gradient according to their isoelectric point. The focusing gel then is transferred to an electrophoretic gel and further separation according to size and electrical charge takes place in a second dimension that is 90° to the first. The resulting radioactive spots are developed by exposure to photographic

TABLE 35.8. ISOELECTRIC FOCUSING VARIANTS OF HLA-A AND HLA-B[a]

HLA Specificity	10th W IEF[b] Variants
A2	A2.1
A2	A2.2
A2	A2.3
A2	A2.4
B7	B7.1
B7	B7.2
B27	B27.1
B27	B27.2
B27	B27.3
B27	B27.4
B27	B27.5
B27	B27.6

[a]From Bodmer JG, Marsh SGE, Albert ED, et al. Nomenclature for factors of the HLA system, 1991. *Tissue Antigens* 1992;39:161–173.
[b] IEF = Isoelectric focusing.

film. A typical result is shown in Fig. 35.13. Each DR molecule separates into several spots. The invariant α chain forms a single large spot. The ß chains, however, usually separate into three spots that reflect different levels of the glycosylation of individual metabolically labeled molecules. This technique, coupled with the DNA methods described below have been applied to the study of DRß chains from other DR specificities. The relationship between HLA-DR detected serologically and class II specificities detected by means of lymphocyte interactions are now becoming clearer. The various HLA-D types are a result of specific DRB1 alleles, some of which are shown in Table 35.7. Several of these alleles may result in the production of ß chains with identical epitopes recognizable by anti-DR antibody and at the same time distinctly different epitopes recognizable by lymphocytes.

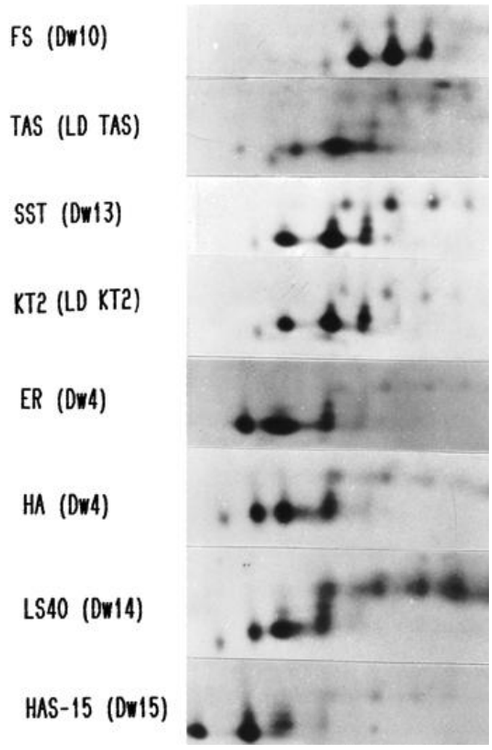


FIGURE 35.13. Two-dimensional gel electrophoresis of immunoprecipitated DR4 molecules. The ß chains appear as three separate spots as a result of differences in glycosylation. The chains migrate in different areas depending on their individual amino-acid composition. Even though these examples all behave serologically as DR4, they have different HLA-D specificities. (Reprinted by permission of Elsevier Science Publishing Co., Inc. from Monos DS, Mickelson E, Hansen JA, et al. Analysis of DR and DQ gene products of the DR4 haplotype in patients with IDDM: possible involvement of more than one locus. *Hum Immunol* 1988;23:289–299. Copyright 1988 by the American Society for Histocompatibility and Immunogenetics.)

Isoelectric focusing alone has been used to study the HLA-A and -B molecules. For the most part, this technique has been able to detect all of the common class I variants currently detected serologically. These are listed in Table 35.8.

HLA typing also can be accomplished at the DNA level and by direct sequencing. Such molecular methods are an important area that has revolutionized HLA typing and presented a great deal of new data about HLA polymorphism. It will be discussed in detail in Chapter 36.

These DNA-based methods are particularly useful in the clinical setting for the HLA class II typing of patients whose B cells are present in low numbers or difficult to isolate for other reasons. Such situations may be encountered in very young children, patients with hematologic disorders such as aplastic anemia or various forms of leukemia, and in patients undergoing chemo- or immunosuppressive therapy. Frequently, patients such as these are being considered for bone-marrow transplantation in which class II identity is essential. Prior to the development of this approach, many repeat serologic and MLC tests had to be performed, which more often than not yielded poor results.

Biochemical methods have contributed a great deal to a better understanding of the HLA system. In particular, they have helped to explain some of the inconsistencies that were experienced by earlier workers. For example, the original discovery of antibodies reactive with HLA antigens on B cells heralded a serologic solution to the lengthy and troublesome MLC. With the greatest confidence, the new HLA-DR antigens were assigned numbers corresponding to the then-defined HLA-D types because they were considered equivalent. Yet this enthusiasm was soon dampened when it became apparent that quite often lymphocytes from unrelated individuals who had identical HLA-DR types stimulated and responded to each other in the MLC. Information obtained using two-dimensional gel electrophoresis and RFLP analysis (Fig. 35.14 and Table 35.8) has revealed that the ß chains conferring a given DR specificity may differ slightly in other areas along their length. Lymphocytes can detect these differences resulting in MLC reactivity among cells with the same DR type but differing in their DR ß molecules.

The isolation, crystallization, and DNA sequencing of the class I molecules has afforded an explanation for the way in which the HLA-Bw4 and Bw6 antigens are related to the rest of the HLA-B series. These studies have shown that the amino-acid sequences responsible for the Bw4 and Bw6 specificities are located at positions 79 and 83, whereas the private specificities are controlled by amino acids at positions 107 and 142 to 145 (see Fig. 35.5).

Studies at the genomic level have extended our knowledge of the relationship between the public antigens and the private specificities. In addition, it has explained why cells from two individuals bearing identical serologically defined DR specificities might stimulate each other in the MLC. Thus, the investigation of HLA at the molecular level has revealed many secrets of this system including the concepts of cross reactivity, inclusions, multispecific antisera and antibody production to antigens presumably not present in the challenging dose. These are now clear when viewed from the standpoint of shared epitopes.

The DNA-Based Molecular Methods Compared to Serologic and Culture Techniques

The DNA-based molecular methods are certainly more accurate than the routine serologic and culture techniques that were the standard since the inception of HLA typing. For example, as mentioned earlier, typing based on DNA sequencing has revealed the true polymorphism of HLA-C that escaped detection by cytotoxicity techniques. In one study, as many as 19 previously undetected alleles were described (36).

The use of molecular methods for the determination of class II HLA types in donor selection for bone marrow transplantation has become routine. In addition, this method is being applied to class I typing as well. It would appear that the molecular techniques may replace serologic methods in all forms of transplantation. Indeed, current investigative efforts are concerned with making the methods more efficient, easier, and quicker to perform so that they can be applied readily to the clinical laboratory. A key issue to be resolved before this can become routine is the fact that molecular methods are able to distinguish a large number of polymorphisms in the HLA. At this time, little is known about the true clinical significance of these minor antigenic differences that may be represented by only one or two amino-acid substitutions. However, this is an area currently under investigation in the area of bone-marrow transplantation. For kidney transplantation in which such high-resolution typing may not be necessary, it is possible to employ a low-resolution typing from which cosmic HLA types such as HLA-A10 or HLA-DR4 can be inferred. This will be discussed in more detail in Chapter 37.

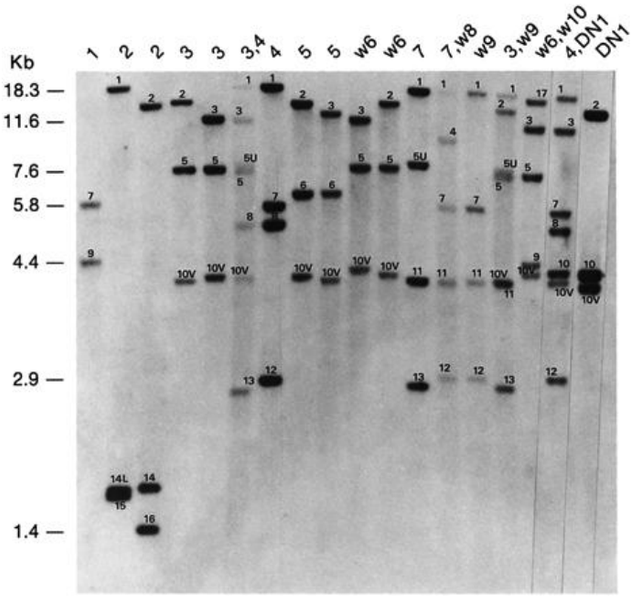


FIGURE 35.14. The RFLP patterns of the different DR specificities. The bands are numbered to indicate the pattern found within each DR antigen. The DNA was cut using the restriction enzyme TaqI, and the patterns were developed using a DR [backwards E] probe. (Reprinted with permission of The University of Chicago Press and The American Society of Human Genetics. From Cox NJ, Mela PA, Zmijewski CM, et al. HLA-DR typing "at the DNA level:" RFLP, and subtypes detected with a DRB cDNA probe. *Am J Hum Genetc* 1988;43:956.)

REFERENCES

1. Medawar PB. Behavior and fate of skin autografts and homografts in rabbits. *J Anat* 1944;78:176–200.

2. Snell GD, Smith P, Babrielson F. Analysis of the histocompatibility-2 locus in the mouse. *J Natl Cancer Inst* 1953;14:457–480.

3. Bach FH, van Rood JJ. The major histocompatibility complex Part 1. *N Engl J Med* 1976;295:806–813.

4. Albert E, Amos DB, Bodmer WF, et al. Nomenclature for factors of the HLA system. *Tissue Antigens* 1977;11:81–86.

5. Volz A, Boyle JM, Cann HM, et al. Report of the second international workshop on human chromosome 6. *Genomics* 1994;21:464–472.

6. Bodmer JG, Marsh SGE, Albert ED, et al. Nomenclature for factors of the HLA system. *Tissue Antigens* 1999;53:407–446.

7. Hurley CK, Schreuder GMTh, Marsh SGE, et al. The search for HLA-matched donors: a summary of HLA-A*, -B*, -DRB1/3/4/5* alleles and their association with serologically defined HLA-A,-B,-DR antigens. *Tissue Antigens* 1997;50:401–418.

8. Baur MP, Neugebauer M, Albert ED. Reference tables of two-locus haplotype frequencies for all MHC marker loci. In: Albert ED, Baur MP, Mayr WR, eds. *Histocompatibility Testing 1984*, New York: Springer-Verlag, 1984:677–755.

9. Thomson G. The effect of a selected locus on linked neutral loci. *Genetics* 1987;85:753–788.

10. Hammerberg C, Klein J. Linkage disquilibrium between H-2 and t complexes in chromosome 17 of the mouse. *Nature* 1975;258:296–299.

11. Baur MP, Danilovs JA. Population analysis of HLA-A, B, C, DR and other genetic markers. In: Terasaki PI, ed. *Histocompatibility Testing 1980*, Los Angeles: UCLA Tissue Typing Laboratory, 1980:955–993.

12. Beatty PG, Mori M, Milford E. Impact of racial genetic polymorphism on the probability of finding an HLA-matched donor. *Transplantation* 1995;60:778–783.

13. Mori M, Beatty, PG, Graves M, et al. HLA gene and haplotype frequencies in the North American population: The National Marrow Donor Program donor registry. ASHI Web Site: http://www.ashi-hla.org/

14. Parham P. Definition and importance of HLA class I null alleles: a new role for serological HLA class I typing. *Proceedings of the Twelfth International Histocompatibility Workshop and Conference*, Vol. II. Paris: EDK, 1997:28–34.

15. Schwartz BD, Luehrman LK, Lee J, et al. A public antigenic determinant in the HLA-B5 cross-reacting group-a basis for cross-reactivity and a possible link with Behcet's disease. *Hum Immunol* 1980;1:37–54.

16. Krangel MS. Two forms of HLA class I molecules in human plasma. *Hum Immunol* 1987;20:155–165.

17. Parham P, Lawlor DA, Salter RD, et al. HLA-A, B, C: patterns of polymorphism in peptide binding proteins. In: Dupont B, ed. *Immunobiology of HLA*, Vol. II. New York: Springer-Verlag, 1989:10–33.

18. Bjorkman PJ, Saper MA, Samraoui B, et al. Structure of human class I histocompatibility antigen, HLA-A2. *Nature* 1987;329:506–512.

19. Kaufman JE, Auffray C, Korman AJ, et al. The class II molecules of the human and murine major histocompatibility complex. *Cell* 1984;36:1.

20. Madden DR. The three-dimensional structure of peptide-MHC complexes. *Ann Rev Immunol* 1995;13:587–622.

21. Carroll MC, Katzman P, Alicot EM, et al. Linkage map of the human major histocompatibility complex including the tumor necrosis factor gene. *Proc Natl Acad Sci USA* 1987;84:8535–8539.

22. Benacerraf B. Role of major histocompatibility complex in genetic regulation of immunologic responsiveness. *Transplant Proc* 1977;9:825–831.

23. McMichael AJ, Ting A, Sweerink HJ, et al. HLA restriction of cell-mediated lysis of influenza virus-infected human cells. *Nature* 1977;270:524.

24. Garcia KC, Degano M, Stanfield RL, et al. An α:ß T cell receptor structure at 2.5A and its orientation in the TCR-MHC complex. *Science* 1996;274:209–219.

25. Terasaki PI, McClelland JD. Microdroplet assay of human serum cytotoxins. *Nature* 1964;204:998–1000.

26. Hopkins KA. Basic microlymphocytotoxicity test. In: Zachary AA and Teresi GA, eds. *ASHI Laboratory Manual*. 2nd ed. Lenexa, KS: American Society for Histocompatibility and Immunogenetics, 1990:195–201.

27. Falk JA, Goeken NE, eds. Cell isolation. In: Zachary AA and Teresi GA, eds. *ASHI Laboratory Manual*. 2nd ed. Lenexa, KS: American Society for Histocompatibility and Immunogenetics, 1990:49–80.

28. Payne R, Rolfs MR. Fetomaternal leukocyte incompatibility. *J Clin Invest* 1958;37:1756–1763.

29. Zmijewski CM, Zmijewski HE, Honeycutt HC. The relationship of the frequencies of white cell antibodies and red cell antibodies in the sera of multiparous women. *Int Arch Allergy* 1967;32:574–582.

30. Brodsky FM, Parham P, Barnstabel CJ, et al. Monoclonal antibodies for analysis of the HLA system. *Immunol Rev* 1979;47:3–61.

31. Amos DB, Baskin H, Boyle W, et al. A simple microcytotoxicity test. *Transplantation* 1969;7:220–222.

32. Johnson AH, Rossen RD, Butler WT. Detection of allo-antibodies using a sensitive antiglobulin microcytotoxicity test: identification of low levels of preformed antibodies in accelerated allograft rejection. *Tissue Antigens* 1972;2:215-255.

33. Bodmer WF, Tripp M, Bodmer JG. Application of a fluorochromatic cytotoxicity assay to human leukocyte typing. In: Curtoni ES, Mattiuz PL, Tose RM, eds. *Histocompatibility Testing 1967*. Copenhagen: Munksgaard, 1967;341–350.

34. van Rood JJ, van Leeuwen A, Ploem JS. Simultaneous detection of two cell populations by two color fluorescence and application to the recognition of B cell determinants. *Nature* 1986;262:795–797.

35. Hirschhorn K, Bach F, Kolodny RL, et al. Immune response and mitosis in human peripheral blood lymphocytes in vitro. *Science* 1963;142:1185–1187.

36. Turner S, Ellexson ME, Hickman HD, et al. Sequence-based typing provides a new look at HLA-C diversity. *J Immun* 1998;161:1406–1413.

36
Molecular HLA Typing

Marcelo Fernández-Viña

Kai Cao

Derek Middleton

Class I and class II genes of the major histocompatibility complex (MHC) encode cell surface glycoproteins whose main biological function is binding and presentation of antigenic peptides to T-lymphocytes. These molecules play an important role in self—nonself recognition as well as in tolerance and rejection of allografts. The class I MHC region spans approximately 1,600 to 2,000 kb of DNA and contains a number of highly homologous genes including those encoding classical HLA class I antigens: HLA-A, HLA-B, and HLA-C. These molecules are highly polymorphic and are found on the surface of nearly all nucleated cells (1, 2, 3 and 4).

Conventionally, identification of HLA class I polymorphism and typing of these allo-antigens has been performed by serological methods. A hallmark in HLA typing is the high degree of serological cross-reactivity between specificities that in some instances may hinder the assignment of some alleles. Inaccuracy in the typing of some HLA phenotypes may arise because of poor quality and a limited supply of or lack of some serological reagents (5, 6, 7 and 8). Similarly, some particular amino-acid variations that introduce HLA microheterogeneity can be detected by T-lymphocytes and not by serological reagents.

The biochemical structure of class I and class II alleles has been elucidated. The class I molecules are heterodimers composed of a polymorphic heavy-chain subunit encoded by genes in the MHC complex on chromosome 6 with a nonvariable light chain subunit (ß-2-microglobulin) that is encoded by a gene located outside the MHC (chromosome 15) (1, 3). Class II molecules are heterodimers as well. Both α and ß subunits are encoded in the MHC. There are differences in the level of polymorphism of the genes encoding these subunits. For example, the gene encoding the DR α subunit is virtually nonpolymorphic while the genes encoding the subunits of DQ and DP molecules present a considerable degree of polymorphism (1, 3).

The heavy chain of the class I molecules comprises three extracellular domains and a transmembrane/cytoplasmic domain; in contrast, each of the class II subunits comprises two extracellular domains and a transmembrane/cytoplasmic domain. In general, most of the differences in the amino-acid sequences of different HLA alleles are found in the distal membrane domains (α-1 and α-2 domains of the class I molecules and α-1 and ß-1 domains of the class II molecules) (1, 9). The genes encoding the subunits of the HLA molecules are discontinuous, containing coding (exons) and noncoding (introns) segments. The class I genes contain 8 exons while the class II genes contain 6 or 7 exons (1, 10).

With the advent of gene cloning nucleotide sequencing, most of the alleles of HLA class I and class II loci have been characterized and the nucleotide and amino-acid differences between HLA alleles have been identified (11). The vast majority of the nucleotide polymorphism occurs in exon-2 of the class II genes and exons 2 and 3 of the class I genes. These exons encode for the distal membrane domains of the class I and class II molecules. Polymorphisms in these exons correlate with the HLA specificites defined by serologic and cellular methods.

Analysis of the differences between alleles of various homologous loci showed that alleles of the same locus share nucleotide sequences that are absent in alleles of other loci. Locus-specific sequences can be found in coding (exons) and noncoding regions (introns). The comparison of alleles of the same locus shows that the distinguishing nucleotide differences between alleles are restricted to some segments or regions of the gene (variable regions) (1, 12). Most of the allelic variation of the class I genes resides in exons-2 and 3 (encoding for the α-1 and α-2 domains), while most of the class II variation occurs in exon-2. DNA based methods to type for HLA alleles have therefore focused in the analysis of nucleotide variation occurring in both exons 2 and 3 of the class I genes and exon-2 of the class II genes. The nucleotide sequences in exon 2 of selected DRB1 alleles is shown in Fig. 36.1 (13, 14).

The application of DNA-based procedures has increased the accuracy of HLA typing and led to the identification of serologically undetected alleles and of many subtypes of many serological specificities. The methods used include DNA amplification and hybridization with oligonucleotide probes (SSOP), amplification with sequence-specific primers (SSP, ARMS) and sequence-based typing. Typing of HLA class II alleles by molecular methods are firmly established and have been incorporated in the routine; more recently, molecular approaches of PCR-based DNA typing have been applied to HLA class I typing including PCR-based DNA sequencing (15, 16, 17, 18, 19 and 20), PCR using sequence-specific primers (SSP) (21, 22, 23 and 24), and sequence-specific oligonucleotide probes (SSOP) (25, 26, 27, 28, 29 and 30).

MOLECULAR METHODS APPLIED TO TYPING OF HLA ALLELES

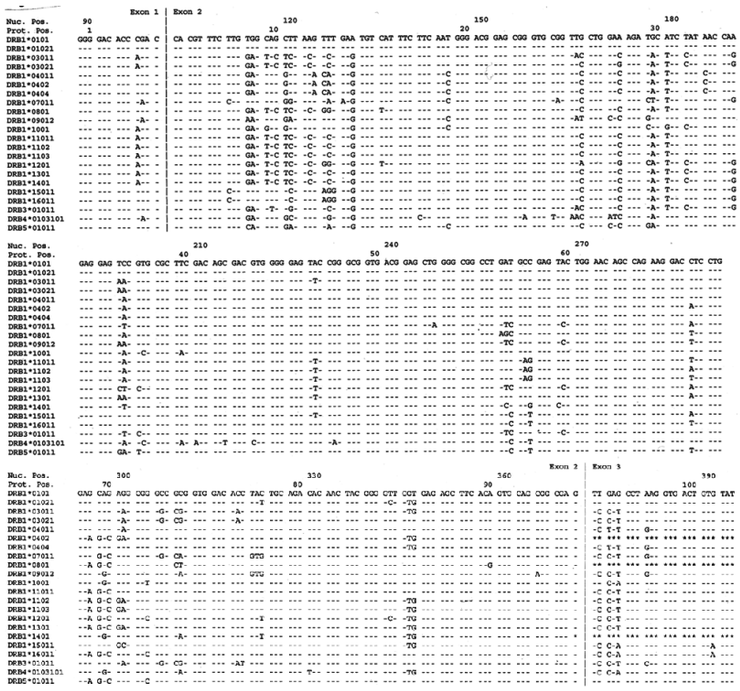


FIGURE 36.1. Exon 2 nucleotide sequence alignment of selected DRB alleles. Complete exon-2 and flanking segments of exons 1 and 3 of representative DRB alleles are shown. Nucleotides are grouped in triplets corresponding codons (protein position). The nucleotide sequence shown corresponds to the cDNA sequence. In the genomic organization of DRB genes, exons are separated by noncoding DNA stretches (introns). The first nucleotide of codon 5 is encoded in exon 1, while the second and subsequent nucleotides are encoded in exon 2. This pattern corresponds to the mRNA splicing pattern observed for other exons from eukaryotes. Polymerase chain reaction primers used for amplification of genomic DNA may span partially both intron and exon sequences. The top sequence corresponds to DRB1*0101. Dashes in other alleles indicate identical nucleotides in the corresponding positions; nucleotide differences are shown by the letter of the corresponding nucleotides. Asterisks (*) indicate that the nucleotide sequence has not been determined (Codons 95–102). In naming HLA alleles in general, the first two digits correlate with the serologic equivalency (e.g., DRB1*0101 and DRB1*01021 type serologically as DR1) and the next two digits indicate differences in nonconservative substitutions. A fifth digit is used to distinguish alleles with identical amino-acid sequence and different only by silent nucleotide substitutions. The nucleotide positions in which alleles of the DRB loci differ from each other are clustered in regions (hypervariable regions) separated by long DNA stretches in which the nucleotide sequences are identical in almost all alleles. The most noticeable hypervariable regions span codons 9–13, 26–38, 57–60, 67–77 and 85–86. Notice that many alleles with similar serotype (e.g., DRB1*0101 and DRB1*01021) differ by a few nucleotide substitutions; DNA stretches containing the distinguishing nucleotide substitutions are found in other alleles (see DRB1*01021 and DRB1*1201 in the stretch spanning codons 78–86) as well.

TABLE 36.1. GENERIC AND GROUP-SPECIFIC (ALLELE GROUP DRB1*01) PRIMERS USED FOR AMPLIFICATION OF EXON 2 OF DRB ALLELES

Amplification Type	Intron-Exon	Nucleotide Sequence[a]	Orientation
DRB Generic			
		Intron 1 NUC/Exon-2 CODONS[c]	Forward
2DRBAMP-A[b]	**Intron 1**	-8 -1 5 6 7 8	Mixture
	TO	5′ CCCCACAG CA CGT TTC TTG 3′	
	Exon2	5′ ---------- --- ---- ---- C--- 3′	
2DRBAMP-B	Exon2	93 92 91 90 89 88 87	Reverse
		5′ CCG CTG CAC TGT GAA GCT CT 3′	
DRB1*01			
Group Specific			
2DRBAMP-1[d]	**Exon2**	7 8 9 10 11 12 13	Forward
		5′ TTC **T**TG **T**GG **CAG CT**T A**AG TT** 3′	
2DRBAMP-B	Exon2	93 92 91 90 89 88 87	Reverse
		5′ CCG CTG CAC TGT GAA GCT CT 3′	

[a] Nucleotide sequences of oligonucleotide primers used for amplification of DRB alleles. Primer pairs were selected to match the sequences of alleles of the locus (generic primers) or groups of alleles (group-specific primers).

[b] A mixture of primers is required to amplify all the DRB alleles. Note Codon 8 in Figure 36.1 which shows similarity of DRB1*07011, *15011, and *16011 (and their difference from *0101) necessitating the use of a different primer in the mixture.

[c] Relative positions are shown as codons for nucleotides from exons. The sequences shown in Figure 36.1 can be used for comparison. Please notice that the sequence alignment in Figure 36.1 shows only coding DNA (exon sequences). Therefore, for comparison with Figure 36.1, the exon segments of the primer should be used. The nucleotide sequence of the forward primers can be visualized directly; the sequence of the reverse primers can be matched by converting the sequence to the complementary nucleotide and by inverting the orientation of the complementary bases.

[d] For amplification of the DRB1*01 alleles, group-specificity is obtained with the forward primer 2-DRBAMP-1; the nucleotide sequence of this primer matches only DRB1*01 alleles; in bold and underlined are shown polymorphic nucleotide positions.

The development of novel molecular techniques has allowed the direct detection of nucleotide substitutions encoding amino-acid differences that determine the differential immunologic characteristics of the HLA antigens. There are many molecular methods available to define the HLA alleles. The ability to amplify DNA segments by the polymerase chain reaction (PCR) greatly facilitates the application of molecular techniques in the routine laboratory tests. PCR-based methods may be broadly classified into three categories: (i)those that generate a product containing internally located polymorphisms that can be identified by a second technique, e.g., PCR sequence-specific oligonucleotide (SSO) probing, PCR-RFLP, PCR followed by sequencing (SBT), (ii)those in which the polymorphism is identified directly as part of the PCR process, although there are postamplification steps, e.g., PCR sequence-specific primer (SSP), and (iii) conformational analysis in which different mutations generate specific conformational changes in PCR products. These are identified by electrophoretic analysis e.g., heteroduplex analysis, single-strand conformation polymorphism (SSCP), denaturing-gradient gel electrophoresis (DGGE) and temperature-gradient gel electrophoresis (TGGE).

The use of a specific technique will depend on the laboratory's requirements. Clinical urgency and requirement, sample numbers, availability of equipment, staff skills, and budget will influence the choice. Some laboratories, depending on their needs, may use a combination of methods. All techniques need to be updated continually to allow for the detection of newly discovered alleles. According to the clinical application, high- or low-resolution typing may be required. Kidney transplant candidates, for instance, do not necessarily have to be typed at a high-resolution level, because only the “broad” serological specificities (DR1-DR10) usually are taken into consideration for organ allocation. For bone marrow transplantation purposes, high-resolution typing is required.

STEPS FOR MOLECULAR TYPING

Part of "36 - Molecular HLA Typing"

In order to achieve molecular typing by any of the most commonly used procedures (SSP, SSOP, and SBT) three general steps are followed: (i) extraction of genomic DNA, (ii) amplification of the genes of interest, and (iii) detection of the sequence polymorphisms that define the alleles.

DNA Extraction

Generally, HLA typing is performed from genomic DNA extracted from nucleated cells, usually blood specimens. A few micrograms of genomic DNA are sufficient to perform complete molecular typing of all HLA loci at the highest resolution level. The purity of the DNA extracted may be an important factor for successful results. Usually, short PCR fragments (less than 300 base pairs) are amplified easily, and a simple salting-out method may be sufficient to achieve successful amplification; however, longer fragments may require higher purity. In general, amplification of exon-2 (approximately 270 bp) of HLA class II is sufficient to achieve the highest resolution level. For typing HLA class I, both exons 2 and 3 and the intervening intron (fragment longer than 900 bp) are amplified by a single pair of primers.

DNA Amplification

DNA is amplified by repeating thermal cycling of the PCR mixture. This mixture contains genomic DNA, dNTPs, primers, Taq DNA polymerase, and the corresponding buffer. This mixture is subject to repeated cycles in which the sample is subject to the following steps: (a) heating to 94°C to 96°C for denaturation of double-stranded DNA, (b) cooling down to the corresponding annealing temperature of the primers used, and (c) warming up again to 72°C in which the Taq DNA polymerase primed by the annealed primers has its optimal activity to incorporate the complementary nucleotides to the single-stranded DNA.

The DNA copies obtained after one cycle of amplification serve as template to be copied in the following cycles. This process allows one to obtain an exponential growth of the PCR products. In order to achieve exponential growth, both DNA strands of the same fragment need to be amplified. Therefore, primers for both strands should be designed and the fragment amplified will include the intervening segment between both primers. The number of PCR cycles and the incubation times at each temperature depend on the length and GC content from both the fragment to be amplified and the primers used. The dNTPs are added in excess and should not be a limiting factor for the efficiency of the PCR reaction.

PCR Primers

The PCR primers are selected to be complementary to sequences of the fragment to be amplified. These usually are 18 to 25 bp and the orientation of the sequence from 5′ to 3′ because the DNA polymerase adds the complementary nucleotides in this orientation. The selection of primers is crucial to the success of the procedure; the nucleotides at the 3′ end of the primer are crucial to achieve the desired specificity and efficiency. Usually, nucleotide mismatches with the target sequence in the middle or at the 5′ end are tolerated and DNA amplification with primers mismatched with the target sequence in those areas can be achieved. On the contrary, nucleotide mismatches at the 3′ end of the primer do not allow the DNA polymerase to extend the complementary sequence. This feature of the DNA polymerase is used to select the primers with the correct specificity.

When amplification of all alleles of a locus is desired, then a common sequence to all of them is selected and this must be different from other homologous loci. In some cases, the primers are selected to amplify a single allele or a group of alleles; because both primers are required to achieve exponential amplification, then specificity of a particular segment can be achieved by either of the two primers. Table 36.1 shows the nucleotide sequences of forward and reverse primers designed for amplification of all DRB alleles (DRB generic) and for amplification of selected groups of alleles [group-specific primers (GS)]. The DRB forward and reverse generic primers match the sequences of all

DRB alleles in the 3′ end of the primer while one or both GS primers match the sequences of only a group of alleles.

Detection of the Sequence Polymorphisms That Define the Alleles

Identification of the HLA alleles carried by a particular individual is accomplished by the detection of nucleotides present in one allele and absent in others. Because most of the distinguishing nucleotide substitutions are not unique to a particular allele, alleles usually are typed by analyzing several polymorphic positions, and the allele assignment is made by the reactivity pattern.

The most common methodologies used to detect nucleotide sequence polymorphism of HLA alleles include SSP (or amplification with group/allele specific primers), sequence specific oligonucleotide probes (SSOP) hybridization and sequenced based typing (SBT).

Sequence-Specific Primers

In this procedure, sets of primer pairs are used to amplify genomic DNA of samples. The primer pairs (both reverse and forward primers) included in each set match sequences of HLA alleles or groups of alleles of the HLA locus to be typed. In addition to the allele/GS primer pair, each tube contains an amplification control (AC) primer pair. The primers of AC match conserved sequences of a selected gene and are used to control the efficiency of the amplification reaction. The PCR products obtained with each primer pair are run in an agarose gel and stained with ethidium bromide. The gel then is visualized under ultraviolet (UV) light. The presence or absence of bands corresponding to PCR products obtained with each allele/GS primer pair is scored. Valid negative scores are those in which the band corresponding to the AC is visualized and that of the allele/GS primer pair is not observed. In these reactions, the molecular weight of the band visualized also is used to assign the scores. The expected size of each band corresponds to the distance (number of nucleotides) between both primers. Usually, the AC band has a larger size than that of the allele/GS primer pair. The AC band is a required control because longer DNA fragments are more prone to fail. Therefore, the presence of the AC band and the absence of the allele/GS band allow confident determination of the negative reaction scores. Patterns of amplifications with each set of primer pairs are analyzed to make the assignment of the corresponding genotype.

Usually, molecular typing by SSP methods is performed with a limited number of primer pairs that identify groups of alleles of a particular locus. Most of the approaches developed identify groups of alleles that belong to the same serologic group. Once these groups are identified, additional SSP reactions are selected to distinguish the alleles within each group.

The SSP methodology is relatively easy to perform and results are obtained within a few hours. However, with increasing numbers of alleles, high resolution typing of a large number of samples may become unwieldy because it involves the performance of many PCR reactions.

HYBRIDIZATION WITH SSOP

Part of "36 - Molecular HLA Typing"

These methods involve the specific amplification of HLA alleles of a particular locus in the HLA region by the PCR, immobilization of PCR products onto nylon membranes, and the subsequent

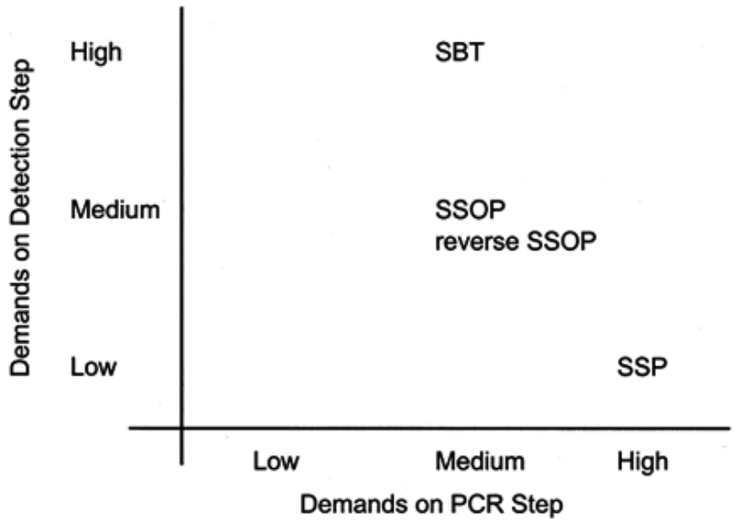


FIGURE 36.2. Comparison of the technical, equipment, supply, and skill demands of various molecular methodologies applied in typing of HLA alleles. Abbreviations: SBT, sequence based typing; SSOP, sequence-specific oligonucleotide probes; SSP, sequence-specific primers.

TABLE 36.2. COMPARISON OF SOME CLASS I DNA TYPING METHODS

Method	Resolution	Advantages	Disadvantages
SSP[a]	Medium	Rapid	Controls
	(High)	single samples	Large number samples
Reverse	Medium	Rapid	Multitiered for high resolution
SSOP[b]	(High)	Single samples	
SSOP[2]	Medium	Robust	Slower
	High	Flexible	Batch mode
		Inexpensive in large volume	Expensive in small volume
SBT[c]	High	Highest resolution	Slower
		Definitive standard	Labor intensive
			Capital costs

[a] SSP: Sequence Specific Primers.
[b] SSOP: Sequence Specific Oligonucleotide Probes.
[c] SBT: Sequence Based Typing.

probing of this product with sequence SSOP. The vast polymorphism of the HLA system results from conversion events whereby small nucleotide sections of one allele (usually no more than 100 bases long) are transferred to another allele. Thus, many of the sequences tend to be shared by alleles and not to be allele-specific. Therefore, several sequence-specific probes must be used to identify a particular allele. In order to differentiate the alleles, a battery of probes is required and it is the pattern of reaction with these probes, which distinguishes the HLA alleles.

The so-called “PCR SSOP forward format” involves PCR amplification of several samples with the same primers. Then, the PCR products of different samples are blotted onto replicate nylon membranes. Each replicate membrane then is hybridized individually with each SSOP. Each SSOP contains radioactive or nonradioactive labels. The excess SSOP then is rinsed off the membranes followed by stringent wash steps in which the SSOP is removed from dots that contain PCR products that do not match its nucleotide sequence. Finally, either an autoradiogram or a colorimetric reaction detects each SSOP.

The SSOP are designed to match the sequence of a particular allele or groups of alleles. Under controlled hybridization and washing conditions, SSOP that match completely a target sequence are detected and SSOP with a single nucleotide mismatch do not sensibly hybridize. Usually, the SSOP are 6 to 19 nucleotides long and their sequence is chosen in a manner that the polymorphic nucleotides are centered. The specificity of the hybridization is achieved by adjusting the hybridization and washing conditions. Usually, SSOP that tend to display cross-hybridization give more specific hybridization if the salt concentration in the wash step is lower or if the stringent wash is performed at a higher temperature.

Several detection systems may be used. One of them consists of labeling the probes with digoxigenin (DIG) and detecting the presence of hybridization of these probes to an identical sequence present in the PCR-amplified HLA allele of an individual by adding an antidigoxigenin antibody conjugated with alkaline phosphates (ALP). The ALP then acts on a chemiluminescent substrate and the light emitted is detected by autoradiography. Other procedures involve the direct conjugation of the SSOP with alkaline phosphatase or horseradish peroxidase. Either a chemiluminescent substrate or a colorimetric substrate can be used for detection. An alternative labeling procedure involves the incorporation of a radioisotope in the SSOP.

The “forward dot blot” SSOP procedure is advantageous because many samples can be tested simultaneously. However, it is time consuming and may be unwieldy in typing samples that require a rapid turn around time. Other hybridization methods used to test single samples in a short period of time have been developed. One of them includes a “reverse format” SSOP procedure in which many SSOP are immobilized on a single solid support (nylon strips) and exposed to hybridization with the PCR products that are in a liquid phase. The latter procedure involves the labeling of the PCR product for detection of hybridization. Another oligonucleotide hybridization-based procedure for testing simultaneously multiple SSOP with a single sample involves hybridization in 96-well plates. In this procedure, PCR products produced with biotin-labeled primers are immobilized on avidin-coated trays and hybridized to a different SSOP in each well. The SSOP are tagged with digoxigenin or labeled with an enzyme. Detection of hybridization is performed with colorimetric substrates.

As a consequence of the large degree of sharing of sequences between alleles of the same and homologous loci, PCR SSOP methods may be performed in sequential tiers to accomplish the highest level of typing resolution.

COMPARISON BETWEEN CLASS I AND CLASS II TYPING SYSTEMS

Part of "36 - Molecular HLA Typing"

The slight differences between typing systems for HLA class I and class II alleles reside in the nature of their diversity. HLA class II genes, in particular DRB loci, present a large number of homologous genes carried by the same chromosome; as a result of the multiloci sequence homology, the selected primers for PCR amplification of all alleles of one locus coamplify alleles of other loci as well. So far, no primer pairs that amplify exclusively exon-2 of all DRB1 alleles have been described. In DRB, many coamplified alleles of various loci share nucleotide sequences in polymorphic areas that require probing. The extensive sharing of sequences very often results in patterns that correspond to several genotypes (ambiguity).

A widely used approach for typing HLA DRB alleles involves a multiple-step procedure. This approach includes an initial step in which the loci and or groups of alleles carried by a sample are identified. If the resolution level required is not achieved in the first step, then locus or group-specific amplifications are performed and the PCR products are probed with the corresponding sets of SSOP to achieve the desired resolution. In general, the highest level of nucleotide polymorphism in HLA class II loci resides in exon-2; in this exon are encoded the distal membrane, α-1 or ß-1, domains of the α and ß subunits, respectively, of the class II molecules. In general, analysis of the nucleotide variations in this exon is sufficient to achieve allele-level typing or to identify polymorphisms that are likely to be relevant in clinical histocompatibility.

In contrast, typing alleles at HLA class I loci requires a different approach. In spite of the many homologous genes and pseudogenes, it has been possible to identify locus-specific sequences. These served as the basis for design of PCR primers that amplify all alleles of the locus to be tested without the co-amplification of alleles of other homologous loci.

The class I genes present high levels of nucleotide variation in two exons (2 and 3). Generally, the molecules corresponding to alleles of the same class I loci present differences in either the α-1 or α-2 domains or in both domains. The exons-2 and -3 of HLA genes encode these domains, and both exons must be analyzed to achieve HLA class I typing. Therefore, the PCR primers selected for HLA class I typing must encompass segments spanning both exons. Locus-specific primers are used to obtain PCR products spanning exons-1, -2, and -3, and the intervening introns 1 and 2 for HLA-A and -B. A shorter segment spanning introns 1 to 3 is obtained with the primers selected for HLA-C amplification. The primers selected are consistently efficient in amplifying all alleles of a particular locus.

The “forward dot blot” approach uses PCR products that are immobilized to membranes; the double-stranded PCR products

then are denatured by alkali and no reconstitution of the double strands occurs after this treatment. Therefore, reannealing of the long PCR fragments obtained with the selected primers does not prevent optimal hybridization with SSOP. In other dot-blot formats, like the reverse-dot-blot assays, the secondary structures of long PCR fragments may cause hindrance to the annealing to some oligonucleotide probes.

For high resolution typing of HLA class I alleles, oligonucleotide probes designed to match almost all sequence segments encoding polymorphisms in exons-2 and -3 of HLA-A, -B, and -C are used. The chosen sequences are 19 nucleotides or less, and the variable nucleotides are centered. These probes were tested under various hybridization and washing conditions to select the optimal conditions for specificity. For SSOP in which it was not possible to determine consistently specific and robust hybridization conditions, they were redesigned or shortened.

TYPING OF HLA ALLELES BY NUCLEOTIDE SEQUENCING (SBT)

Part of "36 - Molecular HLA Typing"

The direct identification of the complete nucleotide sequence of the HLA alleles carried by a sample provides the most accurate procedure for typing. Currently, HLA typing using automatic nucleotide sequencers has a widespread application. Usually, the dideoxy chain termination method is the most widely used approach. Detection of the sequencing fragments usually is performed utilizing fluorescent dyes. Different laboratories use fluorescent primers and/or tagged nucleotides. The use of multiple dyes usually facilitates the performance of electrophoresis. For SBT typing, the same primers used for SSOP-based methods can be used to amplify the HLA genes; the PCR products obtained with these primers then can be sequenced. In order to obtain a more accurate typing, both DNA strands (forward and reverse) are sequenced.

Typing of HLA-A, -B, and -C usually involves the simultaneous amplification and sequencing of both alleles carried by a particular sample, and heterozygous nucleotide assignments are made for each particular position. For loci that do not have unique specific sequences (e.g., DRB loci coamplified with generic primers) the PCR products sequenced are those obtained with allele/GS primer. Otherwise, multiple-base assignments could be made at certain polymorphic positions. As observed in SSOP typing, some heterozygous genotypes result in the same sequencing pattern; these combinations usually are resolved by performance of GS amplification that enables physical separation of the alleles of a sample.

Even though nucleotide sequencing may be technically more demanding than other procedures, SBT presents the advantages over the other procedures because of the relatively fast (24 to 48 hours) turn-around time. In addition, SBT generally achieves allele-level resolution. However, this procedure may be unwieldy for typing large numbers of samples.

COMPARISON OF DNA TYPING METHODS USED FOR TYPING HLA ALLELES

Part of "36 - Molecular HLA Typing"

In principle, the methodologies mentioned above can be used for typing alleles at any locus of the HLA region. Therefore, the choice of methodology used in a particular laboratory will be made based on work volume, equipment, and reagent costs as well as on requirements of personnel training and skills. Usually, most of the technical demands are in the performance of various steps of the procedure and in the interpretation of the data. Fig. 36.2 compares the levels of complexity in the performance of the DNA amplification and detection/interpretation steps. As previously described, the SSP-based tests require the use of several primer mixes and therefore the number of thermocyclers in a laboratory may present a limitation for testing large number of samples. On the other hand, the SSP results are easy to interpret. The other methodologies in contrast require one or a few DNA amplification reactions for typing of the alleles of a single locus, and are highly demanding in the detection (SSOP) and interpretation steps (SBT).

ADVANTAGES AND LIMITATIONS OF THE DNA-BASED TYPING METHODS

Part of "36 - Molecular HLA Typing"

Molecular typing techniques such as PCR-SSOP, SSP, and SBT provide a strong typing tool for typing the highly polymorphic HLA genes, especially those of class I and class II. These methods have advantages because of greater accuracy and a higher level of resolution compared with serological typing. For high volume typing, SSOP methods seem to be the most convenient and powerful methods to use because a large number of specimens can be tested simultaneously. Various levels of resolution for alleles of HLA class I and class II loci can be achieved by using selected sets of SSOP or SSP. A summary of the advantages and resolution offered by each procedure is presented in Table 36.2. This table also presents the disadvantages and limitations inherent to each procedure.

DELINEATION OF THE TYPING SYSTEMS, TRENDS AND FUTURE STRATEGIES

Part of "36 - Molecular HLA Typing"

It is likely that as more alleles are discovered, more ambiguities will be found and the performance of group-specific amplification

(GSA) will have to be applied more often. In order to make the HLA class I typing process more efficient, the procedure may have to be redesigned to a higher frequency of ambiguities. This process may include the performance of an initial screening for heterozygosity at a few polymorphic segments; once heterozygosity is identified, then GSA with selected primers may allow separation of the alleles. Either PCR-SSOP or SSP may be used for the screening procedure for heterozygosity. If the SSP approach is chosen as the initial step, the PCR-SSP products obtained could be further analyzed by hybridization with SSOP or by nucleotide sequencing.

APPROACHES FOR TYPING OF NOVEL POLYMORPHISMS

Part of "36 - Molecular HLA Typing"

The method for typing HLA class I alleles described here is based on the analysis of exons-2 and -3 while for typing class II alleles only exon-2 is analyzed routinely. Recently, new polymorphisms outside these exons have been described and several pairs of alleles have been found to differ only by substitutions in the new variable segments. In fact, some of the nucleotide substitutions in the novel polymorphic site result in a termination codon that results in nonexpression of the alleles carrying them (31, 32, 33, 34, 35 and 36). Therefore, in spite of identical exon-2 and exon-3 nucleotide sequences, some pairs of HLA class I alleles will differ in their expression pattern. Other mutations that alter HLA expression have been described in intronic splicing sites (36) and in the promoter regions. Typing strategies for mutations outside exons-2 and -3 will require further development. The methodologies described here may be supplemented by the performance of additional amplifications to probe for the expression variants by either SSP or SSOP when ambiguities including these alleles are identified. Alternatively, new PCR primers that amplify DNA segments including exons-2 and -3 and spanning the areas containing the mutations that affect expression may be developed.

The current trend in HLA typing is to replace the serological methods that detect the expressed products by genomic testing procedures. However, not many extensive population studies in which serologic and molecular testing have been performed in parallel have taken place. Therefore, it is possible that novel null-alleles remain to be identified. Until such studies have been completed, serological testing should not be completely omitted from clinical histocompatibility testing because alleles typed only at the genomic level may contain undetected expression defects. It has been shown that HLA mismatches influence the outcome of allogeneic transplants and this effect is highly relevant in bone marrow transplantation. In clinical histocompatibility testing, assumptions made about the expression of some alleles, when in fact the corresponding are not expressed, may result in the introduction of undetected HLA mismatches.

The aim of typing at the intermediate-resolution level is to identify groups of alleles at a level that yields at least the same information as that obtained by serologic methods. The dot-blot procedure is well suited for typing large numbers of samples. This is a highly accurate procedure and therefore may replace serologic typing of volunteers of bone-marrow donor registries. It is likely that still newer methodologies will enhance the performance of molecular testing by improving turn-around time, ease of performance, and level of resolution. Some of the currently used PCR primers and the target sequences chosen for analysis also may prove useful for development of new approaches for HLA typing.

REFERENCES

1. Lawlor DA, Zemmour J, Ennis PD, et al. Evolution of Class-I MHC Genes and Proteins: from natural selection to thymic selection. *Annu Rev Immunol* 1990;8:23–63.

2. Yunis EJ, Dupont B. The HLA system. In: Nathan and Oski, eds. *Hematology of Infancy and Childhood: Genetics*. 4th ed. Philadelphia: WB Saunders, 1992:1692–1728.

3. So A. Genetics, polymorphism and regulation of expression of HLA region genes. In: Lecher A. *HLA and Disease*. San Diego: Academic Press, Inc., 1994:1–34.

4. Clay TM, Bidwell JL, Howard MR, et al. PCR-fingerprinting for selection of HLA matched unrelated marrow donors. Collaborating centers in the IMUST study. *Lancet* 1991;337:1049–1052.

5. Bozon MV, et al. Error rate for HLA-B antigen assignment by serology: implications for proficiency testing and utilization of DNA-based typing methods. *Tissue Antigens* 1997;50(4):387–394.

6. Bozon MV, et al. Comparison of HLA-A antigen typing by serology with two polymerase chain reaction based DNA typing methods: implications for proficiency testing. *Tissue Antigens* 1996;47(6):512–8.

7. Sintasath DM, et al. Analysis of HLA-A and -B serologic typing of bone marrow registry donors using polymerase chain reaction with sequence-specific oligonucleotide probes and DNA sequencing. *Tissue Antigens* 1997;50(4):366–71.

8. Winden T, Hurley CK, Maiers M, et al. Standardization of large scale SSOP typing of HLA-DRB1, DRB3, and DRB5 for unrelated bone marrow donor registries. *Human Immunology* 1999;59(suppl 1):103.

9. Giles RC and Capra JD. Structure, function, and genetics of human class II molecules. *Adv Immunol* 1985;37:1–71.

10. Mach B, Gorski J, Rollini P, et al. Polymorphism and regulation of HLA class II genes of the major histocompatibility complex. *Cold Spring Harb Symp Quant Biol* 1986;51(part 1):67–74.

11. Bodmer JG, Marsh SG, Albert ED, et al. Nomenclature for factors of the HLA System, 1998. *Hum Immunol* 1999;60:361–95.

12. Erlich HA and Gyllensten UB. Shared epitopes among HLA class II alleles: gene conversion, common ancestry and balancing selection. *Immunol Today* 1991;12:411–414.

13. Kimura A, Dong RP, Harada H, et al. DNA typing of HLA class II genes in B-lymphoblastoid cell lines homozygous for HLA. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991*. New York: Oxford University Press, 1992:419–425.

14. Fernandez-Vina MA and Bignon JD. Primers and oligonucleotide probes (SSOP) used for DNA typing of HLA class II alleles. In: Charron D, ed. *Genetic Diversity of HLA: Functional and Medical Implication*. Paris: EDK, 1997:596–632.

15. Bignon JD and Fernandez-Vina MA. Protocols of the 12th International Histocompatibility Workshop for typing of HLA class II alleles by DNA amplification by the polymerase chain reaction (PCR) and hybridization with sequence specific oligonucleotide probes (SSOP). In: Charron D, ed. *Genetic Diversity of HLA: Functional and Medical Implication*. Paris: EDK, 1997:584–595.

16. Holmes N, Ennis P, Wan AM, et al. Multiple genetic mechanisms have contributed to the generation of the HLA-A2/A28 family of class I MHC molecules. *J Immunology* 1987;139:936–941.

17. Ennis PD, Zemmour Z, Salter RD, et al. Rapid cloning of HLA-A, B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in ampification. *Proc. Natl Acad. Sci. USA* 1990;87:2833–2837.

18. Santamaria P, et al. HLA class I sequence-based typing. *Hum Immunol* 1993;37:39–50.

19. Kotsch K, Wehling J, Blasczyk R. Sequencing of HLA class II genes based on the conserved diversity of the non-coding regions: sequencing based typing of HLA-DRB genes. *Tissue Antigens* 1999;53:486–497.

20. Kotsch K, Wehling J, Kohler S, et al. Sequencing of HLA class I genes based on the conserved diversity of the noncoding regions: sequencing-based typing of the HLA-A gene. *Tissue Antigens* 1997;50:178–191.

21. Krausa P, Bodmer JG, Browning MJ. Defining the common subtypes of HLA-A9, A10, A28 and A19 by use of ARMS/PCR. *Tissue Antigens* 1993;42:91–99.

22. Krausa P, Moses J, Bodmer W, et al. HLA-A locus alleles identified by sequence specific PCR. *Lancet* 1993;341:121–122.

23. Krausa P, Browning MJ, Moses JH, et al. A locus DNA typing of the 4AOH cell panel by ARMS PCR. *Euro J Immunogenetics* 1994;21:199–204.

24. Bunce M, O'Neill CM, Barnardo MC, et al. Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). *Tissue Antigens* 1995;46:355–367.

25. Oh SH, Jin ZW, Khan R, et al. HLA class I typing by sequence-specific oligonucleotide probes after genomic DNA amplification. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991*. New York: Oxford University Press Inc., 1992;2:210–213.

26. Fernández-Viña MA, Falco M, Sun YP, et al. DNA typing for HLA class I alleles: I. Subsets of HLA-A2 and of -A28. *Hum Immunol* 1992;33:163–173.

27. Oh S, Fleischhauer K, Yang SY. Isoelectric focusing subtypes of HLA-A can be defined by oligonucleotide typing. *Tissue Antigens* 1993;41:135–142.

28. Allen M, Liu LM, Gyllensten U. A comprehension polymerase chain reaction-oligonucleotide typing system for the HLA class I A locus. *Hum Immunol* 1994;40:25–32.

29. Williams F, Mallon E, Middleton D. Development of PCR-SSOP for HLA-A typing of bone marrow registry donors. *Tissue Antigens* 1997;49:61–66.

30. Williams F, Mawhinney H, Middleton D. Application of an HLA-B PCR-SSOP typing method to a bone marrow donor registry. *Bone Marrow Transplant* 1997;19:205–208.

31. Williams F, Curran MD, Middleton D. Detection of HLA-A*24 null alleles by DNA typing methods. *Tissue Antigens* 1998;52:208–212.

32. Mine H, et al. An HLA-B null allele (B*1526N) with a stop codon in exon 3 generated by a point mutation. *Tissue Antigens* 1997;50:351–354.

33. Magor KE, et al. Natural inactivation of a common HLA allele (A*2402) has occurred on at least three separate occasions. *J Immunol* 1997;158:5242–5250.

34. Ishikawa Y, et al. HLA-A null allele with a stop codon, HLA-A*0215N, identified in a homozygous state in a healthy adult. *Immunogenetics* 1996;43:1–5.

35. Laforet M, et al. A nucleotide insertion in exon 4 is responsible for the absence of expression of an HLA-A*01 allele. *Tissue Antigens* 1997;50:347–350.

36. Laforet M, Froelich N, Parissiadis A, et al. An intronic mutation responsible for a low level of expression of an HLA-A*24 allele. *Tissue Antigens* 1997;50:340–346.

37
HLA: Applications

Malek Kamoun

Chester M. Zmijewski

- TRANSPLANTATION: SOLID ORGANS
- TRANSPLANTATION: BONE MARROW AND STEM CELLS
- PLATELET TRANSFUSION
- DISEASE ASSOCIATION
- PARENTAGE TESTING

TRANSPLANTATION: SOLID ORGANS

Part of "37 - HLA: Applications"

The critical need for allogeneic renal transplantation as a therapeutic modality in end-stage renal disease served as the impetus for the investigations that led to the rapid discovery and understanding of the HLA system.

The concept of donor-recipient matching for histocompatibility antigens prior to transplantation is founded on the same principles underlying those used for the selection of blood for transfusions. An allograft is rejected as a result of an immune response directed against alloantigens expressed on the graft that are absent from the host. Matching for the most prominent antigens should result in some abrogation of the antigenic stimulus, resulting in a response that is more amenable to control by immunosuppressive drugs. It is recognized that exact matching for all possible antigens may be a practical impossibility. However, closer matches for major antigens will result in lesser degrees of antigenic challenge and therefore will require lower levels of immunosuppressive drugs, resulting in better overall outcomes.

Renal Transplantation

Renal transplantation is carried out successfully using two groups of donors as sources for allografts. The first of these are living donors, usually relatives of the patient, who have two functioning kidneys, are in good health, and consent to donate. This practice has been extended to include living nonrelated donors such as spouses. The second source is from cadaveric donors who were in reasonable health before death and whose kidneys were not affected by the fatal disease or circumstances of death. The most suitable donors in this category are victims of head trauma or other situations that result in irreversible brain damage.

There are advantages and disadvantages to the use of each of these types of donors, and each presents a different set of problems with respect to matching. Although the HLA system automatically comes to mind in the consideration of donor-recipient matching for transplantation, its importance is not exclusive. The strongest and probably most significant transplantation antigens in humans are the antigens of the ABO blood group system (1). These antigens, in addition to being present on erythrocytes, are fully expressed in all tissues of the human body with the exception of those of the central nervous system. In addition, all people possess antibodies against the antigens (of the ABO system) they lack, developed as a result of natural environmental stimuli. These antibodies have the potential of reacting with antigens expressed on the vascular endothelium of the donor organ and can result in a hyperacute rejection similar to the white graft reaction described earlier. Therefore, matching for ABO within the same general guidelines used for blood donor selection is an essential prerequisite for organ transplantation.

The HLA matching of living related donors is performed with respect to the sharing of haplotypes containing the HLA-A, HLA-B, and HLA-DR antigens. Data excerpted from the United Network for Organ Sharing (UNOS) transplant registry are shown in Table 37.1 (2). The overall 1-year survival among 1,013 patients with grafts matched for both haplotypes was 96.6% and the survival at 5 years was 85.9%. Among 1,872 patients with grafts matched for only one haplotype, the 1-year survival was slightly lower (95.3%) and the 5-year survival was only 73.4%. The overall graft survival in blacks was lower than in whites. UNOS data indicate an overall 1-year graft survival of 93.7% in whites versus 91% in blacks. At 3 years, the difference is greater with a graft survival of 86.2% in whites and 76.2% in blacks. From these data, it is clear that HLA typing has a significant influence on the survival of kidneys from living, related donors.

TABLE 37.1. LIVING RELATED KIDNEY GRAFT SURVIVAL ACCORDING TO HAPLOTYPE MATCHES[a]

Level of HLA Mismatch	N	1 Year Survival %	N	3 Year Survival %	5 Year Survival %
0	1013	96.6	4179	91.2	85.9
3	1872	93.0	6605	83.1	73.4

[a]Data from United Network for Organ Sharing 1998 (2).

The striking difference in the overall outcome of transplantation between whites and blacks is noteworthy. These differences, which have been observed not only with living, related donors but with cadaveric donors as well, cannot be readily explained. The inability to type for certain antigens prevalent only in blacks plays no part in the cases of living, related donors because they are matched for haplotypes derived from family studies. However, so-called center effects, physiology, and lack of patient compliance have been cited as possible contributory causes.

The center effect must be considered whenever analyzing data of this type, which have been collected in a central registry from a large number of individual transplant centers. The effect originates from a collection of events that may exert a significant influence on outcome but that are not attributable to the condition being measured, for example, HLA matching. These influences

would include such factors as surgical technique, standards for organ harvesting, immunosuppressive protocols, clinical condition of the patient prior to transplantation, nursing care, socioeconomic status of the patient, patient compliance, postgraft monitoring, and assessment of clinical condition. Interestingly enough, this collection of factors, which contribute to the center effect and therefore may obfuscate the interpretation of collated data to some extent, may have a marked effect on the results from a single center. Thus, any single center might appear to show either a very high or a very low correlation of graft survival with typing. Therefore, the analysis of pooled data from a large number of patients gives the best picture because the center effect is minimized.

HLA matching appears to have an effect on long-term graft survival of kidneys from cadaveric donors. The data from the UNOS registry (2), shown in Table 37.2, indicate a 1-year graft survival, which is not very different in cases involving a complete match from those that are completely mismatched. However, after 5 years, the gap in percent survival widens. In an earlier analysis of long-term survivors of kidney transplantation from cadaveric donors, the average half-life of grafts was 10.1 years for those with no HLA-A or -B mismatches and 6 years for those with no HLA antigens in common. When both HLA-A, -B and HLA-DR matching are considered, the results are clearly superior, with patients receiving organs with zero mismatches for the entire constellation of HLA antigens and achieving a graft survival half-life of 11.3 years compared to a half-life of 6.3 years for grafts that were completely mismatched for HLA-A, -B and DR (3).

TABLE 37.2. CADAVER KIDNEY GRAFT SURVIVAL ACCORDING TO HAPLOTYPE MATCHES[a]

	1 Year Survival	5 Year Survival
Level of HLA Mismatch	%	%
0	95.2	82.1
1	96.1	82.2
2	94.8	82.0
3	95.0	81.8
4	94.9	81.0
5	94.4	80.1
6	93.8	78.8

[a]Data from United Network for Organ Sharing 1998 (2).

HLA matching is especially beneficial in second transplants and in patients with preformed antibodies (4). Patients matched with their donors for both DR antigens (disregarding any HLA-A, -B matching) have a 1-year survival of 83% versus 76% for those with one match and 74% for those with no matches. At 5 years, 64% of the fully matched grafts are still functioning while only 55% of those with one or no matches survive. Matching for the HLA-A and -B antigens in addition to HLA-DR has an enhancing effect. In the study cited, the patients matched for both HLA-DR antigens achieve a 1-year graft survival of 87% if they also are matched for a single HLA-A antigen and a single HLA-B antigen versus 74% for those with no HLA-A or -B match. Additional analyses have confirmed these findings and have shown that HLA-A, -B matching has a significant effect on long-term graft survival (5). Interestingly, the same study revealed that HLA-DR matching alone does not improve long-term graft survival even though it appears to have an effect during the early period after transplantation.

A donor-recipient matching scheme is being evaluated that is based on the matching for the relevant immunogenicity of shared epitopes in AB and DR (6). The examination of the antibodies from 50,000 pregnancies revealed that most of the specificities were directed against public and private specificities defined by molecules with shared epitopes (7). Furthermore, this analysis allowed the calculation of an immunogenicity score for each specificity, based on the chance of immunization. This immunogenicity score varied by more than 10 times between different specificities. In addition, the shared epitopes based on the cross-reactivity between different DR types could be defined. This approach to matching is receiving a great deal of attention and is being supported by the ever-increasing understanding of the HLA antigens at the amino-acid level (8). Hopefully, this should lead to a better definition of allowable mismatches like those that are routinely applied to the choice of blood for transfusion.

Still another sophisticated procedure for donor-recipient matching is being explored at the peptide level. Using a computer-generated table of peptides unique to each HLA specificity, Takemoto and Terasaki (9) used this system for matching. They claim that this approach was more effective in identifying patients with extremely poor outcomes than conventional techniques. Although such a tool is far from practical in its present state of development, it may offer new insights to the problem.

Under ideal circumstances, it would be most desirable to obtain donors who are perfectly matched with respect to all six HLA antigens of the recipient. In spite of the polymorphism of the HLA system, some haplotypes are more prevalent in the population than others. Therefore, in certain cases, it is possible to find exact matches, even among random donors. Unfortunately, this can take time, and for many patients with end-stage renal disease, the time needed to find an exact match may not be available. For this reason, each case must be evaluated individually, and compromises based on both scientific and clinical judgment must be made.

The analysis of data such as those presented earlier has led to the development of a hierarchy of matching: A,B,DR >B,DR >A,DR >DR >B >A. A convenient way to select a recipient for any given random donor is based on the awarding of points. The points assigned by UNOS for the various degrees of matching are given in Table 37.3. In addition to HLA matching, this scheme also allows for the awarding of points for other factors related to the quantity of preformed anti-HLA antibody in the recipient serum and the length of time on the waiting list.

TABLE 37.3. UNOS POINT SYSTEM FOR ORGAN SHARING

Degree of HLA Matching	Points
0 B or DR mismatch	7
1 B or DR mismatch	5
2 B or DR mismatches	2
PRA-> 80	4

Graft Rejection

Solid-organ graft rejection can take place as one of three different types or a combination. The most common types can be defined as follows.

Hyperacute rejection. The improved new crossmatching techniques greatly diminished the incidence of hyperacute rejection. Hyperacute rejection occurs when a kidney is transplanted into a recipient with preformed cytotoxic antibodies against donor class I or more rarely, class II HLA antigens. It also occurs when engraftment is attempted between ABO-incompatible individuals.

Hyperacute rejection happens so rapidly that transplanted organs may become visibly devitalized while the patient is still on the operating table. There is no effective treatment for this type of rejection. A hyperacutely rejected kidney should always be removed

Accelerated rejection. Accelerated rejection occurs within the first two to five posttransplant days. This form of rejection is thought to be a result of an amnestic response to donor tissue resulting from prior sensitization to donor alloantigens. Accelerated rejection may be mediated by cytotoxic T cells, anti-HLA alloantibodies, or both and often is associated with vasculitis or T-lymphocyte infiltration of renal vessels. The prognosis of these early rejection episodes is poorer than that seen with acute cellular rejection. Accelerated rejection rarely occurs in patients receiving antilymphocytic antibodies for induction immunosuppression. This fact emphasizes the central role of lymphocytes in its occurrence.

Acute cellular rejection. Acute cellular rejection occurs in 20% to 60% of renal transplant recipients. It occurs usually in the first 3 months following engraftment and is most common in the 2nd and 3rd weeks posttransplant. Mild acute cellular rejection is marked by diffuse mononuclear cellular infiltrates and lymphocytic infiltration of tubular structures. It usually is treated by the administration of large doses of immunosuppressive drugs.

Other Solid Organs

In contrast to the kidney, donors of nonpaired organs such as the heart and liver must be of cadaveric origin, although single hepatic lobes from living donors have been used. Furthermore, in these cases, clinical considerations play a very prominent role because, unlike the kidney patient who can be supported on dialysis, end-stage coronary and liver diseases are not amenable to long-term maintenance therapy. In the case of these organs, ABO compatibility and size are the two most important considerations. Size is especially important because these grafts must fit when placed orthotopically and also must be capable of handling the physiologic burden of the recipient. Speed is essential, as preservation techniques either are not developed or are very critical. For these reasons, there is little time for prospective donor-recipient matching based on HLA.

Pancreas and Kidney-Pancreas Transplantation

Pancreas is transplanted both as a solid organ and as dissociated islet cells. Both of these tissues are exceptionally susceptible to the action of HLA antibodies, isolated islet cells somewhat more. The greatest success is obtained when pancreas and kidney from the same donor are transplanted simultaneously. Therefore, the same HLA matching criteria used for the selection of kidneys are employed.

Vascularized pancreas grafts have assumed an increasingly important role in the treatment of insulin-dependent diabetes mellitus (IDDM). Simultaneous kidney-pancreas is gaining acceptance as a viable alternative to kidney transplant alone in diabetic transplant recipients because of its ability to provide superior glycemic control and an improved quality of life. Although morbidity is still higher after simultaneous kidney-pancreas transplant compared to kidney transplant alone, the addition of the pancreas does not appear to jeopardize the kidney transplant. The most common complication, rejection, does not appear to adversely influence long-term kidney allograft survival compared to solitary kidney transplantation.

Cardiac Transplantation

Retrospective studies on the significance of HLA matching in cardiac transplantation are controversial. Some data appear to support a positive effect from DR matching, while others indicate no apparent effects from such matching (10, 11). Recent data suggest that pretransplant antibodies to HLA antigens appear to have an effect on outcome. It appears that patients with positive cross-matches to HLA antigens do more poorly than patients whose cross-matches are negative. In addition, positive B-cell

cross-matches have been reported as the cause of early rejection. Experience gained from the evaluation of renal transplantation suggests that the myriad of factors associated with cardiac patients is even more complicated. This contributes significantly to the center effect, requiring the study of many more cases before any meaningful conclusions can be reached. The relationship of the donor HLA status to that of the recipient generally is not determined prospectively, and therefore, allocations schemes for heart are not based upon HLA matching criteria as they are for kidney transplant.

The constraints of donor-ischemic time and unpredictability of donor availability do not permit prospective matching. However, several studies have shown correlation between the degree of HLA mismatch and acute rejection, allograft vasculopathy, and survival. Indeed, HLA mismatch and other potential risk factors for rejection were analyzed using multivariate analysis of data from 27 institutions participating in the Cardiac Transplant Database Research between 1990 and 1992. The few patients with zero, one, or two mismatches had a 54% freedom from rejection at 1 year versus only a 36% freedom from rejection in those patients with three or more HLA mismatches.

Screening for anti-HLA antibodies is very important in heart transplant candidates. Some studies have indicated the need to use sensitive techniques such as flow cytometry or ELISA for the detection of anti-HLA antibodies. If a heart transplant candidate has preformed panel-reactive anti-HLA antibodies, a T-cell cross-match should be performed using a sensitive technique. In general, if the panel-reactive antibody (PRA) is less than 10%, donor allocation can be accomplished without a prospective cross-match. The presence of PRA prior to sensitization is associated with the poorer outcome and might dictate a more aggressive immunosuppressive regimen.

Liver Transplantation

In liver transplantation, clinical considerations are even more important than in heart transplantation; as a result, a careful prospective study has never been attempted. However, data from retrospective studies performed at the largest liver transplant center in the United States are showing some interesting, although somewhat ironic, results. They seem to suggest that HLA matching for DR has a reverse effect; that is, patients with grafts poorly matched for class II antigens appear to do better than those with well-matched grafts (12).

This finding seems to parallel results from rodent models of endocrine organ transplantation. In these systems, it has been suggested that the graft failure stems from the ability of the cells in the matched tissue to act as antigen-presenting cells to the host immune system. Class II antigen expression by the nonlymphoid cells being transplanted may be induced by IL-2 and interferon gamma released during the normal inflammatory stages of the transplant procedure. Class II identity with the host allows these cells to act as antigen-presenting cells. The antigens being presented are thought to be minor histocompatibility antigens expressed on the donor cells that are absent from the host. If this is indeed the cause of the failure of DR-matched livers in humans, it could lead to a better understanding of human minor histocompatibility antigens (13).

Some of the more recent data on liver transplantation seem to suggest that this organ appears to be quite resistant to the action of anti-HLA antibodies. Patients who demonstrate the presence of preformed antibody against donor antigens fail to exhibit hyperacute rejections. They do undergo rejection episodes. However, these can be controlled and reversed with aggressive immunosuppressive therapy (14).

Lung Transplantation

Lung transplantation is one of the most recent therapeutic modalities currently being practiced. It is performed in a relatively few centers and as a result data are lacking relative to the significance of HLA matching. As in heart transplants and liver transplants, clinical considerations are paramount. Nevertheless, as in all solid organ transplants, matching for the ABO blood groups is essential. The very best pretransplant strategy is to obtain an HLA antibody profile of the patient that is as complete as possible. It has been noted that antibodies to class II major histocompatibility complex (MHC) antigens can be particularly dangerous to this type of graft. Therefore, it is extremely important to screen for such antibodies in these patients, determine their specificity, and select donors accordingly, prior to organ harvest.

Special Immunologic Considerations in Pediatric Transplantation

A significantly greater percentage of children under the age of six suffer graft failure after their first rejection episode than do older recipients. There is evidence that young children have heightened immunologic reactivity; because children under the age of 15 years are more likely to develop anti-HLA antibodies in response to blood transfusion. Thus, presensitization is a significant predictor of poor graft outcome. Highly sensitized children can wait inordinately long periods for a crossmatch-negative transplant, and this wait can represent a problem for small children who often tolerate dialysis poorly.

Significance of Alloantibodies

A very important consideration in the selection of a donor of any solid organ is the antibody status of the recipient. Preformed antibodies in the serum of the recipient directed against donor antigens can have a deleterious effect on the outcome of transplantation. Consequently, presence of such antibodies must be ascertained prior to donor selection. Antibodies to HLA antigens are produced by recipients awaiting transplantation in response to a number of different stimuli, including blood transfusion and previous transplantation. The dose of challenging antigen is presented differently in each of these situations. This results in a variety of different immune responses that may vary with respect to the quantity, biochemical affinity, and in some cases the predominant immunoglobulin class of the antibodies produced as well as their specificity and persistence. In addition to developing alloantibodies, some patients awaiting renal, heart, or pancreas transplants may have produced a variety of autoantibodies as a result of their primary disease or pregraft therapeutic modalities.

In marked contrast to the well-developed and sharply defined response obtained in alloimmunization of pregnancy, the immunologic response to a clinical allograft is somewhat impaired. Under these circumstances, the antibody producer is an unhealthy individual whose immune system has been compromised as a result of disease; this unhealthy immune system has been further compromised deliberately by immunosuppressive drugs. This doubly impaired immune system is then presented with an antigenic challenge as a single bolus containing a large variety of antigenic specificities. The result is a broad polyclonal response directed against a host of cross-reactive antigens, some of which may not even be represented on the graft.

The quality of the antibody response to blood transfusion may be classified as falling somewhere between that of pregnancy and that of transplantation. In this case, the subject is an unhealthy individual whose immune system may be somewhat depressed as a result of uremia, but the patient has not been pharmacologically immunosuppressed. The antigen is introduced intravenously, a route that is optimal for tolerization by soluble HLA antigens shed from the cellular components of the transfusion. On the other hand, the number of antigen-bearing cells, leukocytes, and platelets in a unit of packed red blood cells is relatively small and decreases even further with time after the unit is drawn. As a result, multiple transfusions of units of blood are required to constitute an adequate antigenic dose. These units come from different blood donors whose cells display a variety of private specificities. The antigens that are most readily recognized by the immune system, therefore, are those directed against public specificities shared by the donors and those directed against private specificities that occur with relatively high frequencies in the general population. Patients undergoing transfusion routinely produce a spectrum of antibodies, including anti-HLA-A, -B, -C, and -DR.

The antibodies produced as a result of either of these challenges are IgG and IgM, though the IgG antibodies are frequently of low affinity. They are directed against a variety of HLA antigens and tend to react with the public specificities of the HLA cross-reactive groups. In addition, IgM antibodies with very broad specificities may be produced as well. Current data seem to indicate that, with some exceptions, these antibodies tend to disappear with time, and, in the case of transfusion-induced antibodies, fail to reappear upon secondary challenge (15).

Antibodies to class I HLA antigens can be directed specifically against private specificities or public specificities. Antibodies with private specificities are clinically relevant and are directed against epitopes formed by a unique amino-acid sequence that forms a single specificity. Antibodies to public specificities, on the other hand, define various cross-reactive groups. They have been the subject of a great deal of study to establish their importance in clinical transplantation.

Antibodies to the class II HLA antigens DR and DQ appear to be clinically significant with respect to adversely affecting graft survival. On the other hand, other antibodies that preferentially react with B cells but that have no apparent specificity for class II or class I MHC antigens may have little or no clinical significance. Many of these antibodies are composed of IgM immunoglobulin and react preferentially at 20°C or below. They are not very stable and can be inactivated by treatment with dithiothreitol (DTT) or other similar reducing agents (16).

Autoantibodies occasionally may be found in the sera of pretransplant patients who are suffering from diseases with a heightened immune response such as lupus nephritis. They tend to react with a patient's own cells in *in vitro* serologic tests and may sometimes have HLA specificity (17). In addition, autoantibodylike substances may be produced in response to certain drugs such as hydralazine or procainamide. These may not be true antibodies at all; they may be aggregated γ globulins or other macromolecules.

Autoantibodies appear to play no role in transplantation rejection, but they serve as an indicator of a patient's abnormal immune status and frequently interfere with the interpretation of serologic tests for alloantibody detection. Most often these antibodies are IgM and react preferentially at room temperature or below. However, their thermal amplitude is occasionally higher and they will react at 37°C. For the most part, such antibodies can be inactivated with DTT or removed by absorption with the patient's own cells.

Alloantibody Screening

Testing for donor-reactive antibody has been a routine part of histocompatibility evaluation for transplantation. Clinical benefits include: (i) the identification of unacceptable antigens; (ii) the detection of unreported or unrecognized sensitizing events; (iii)the selection of appropriate crossmatch techniques and interpretation of results; and (iv)the provision of an estimation of the likelihood of finding a crossmatch-compatible donor. Federal standards under Clinical Laboratory Improvement Amendments (CLIA) regulation, *Code of Federal Regulation* (CFR) 4933.1265, stipulate that sera from potential transplant recipients must be screened for preformed anti-HLA antibodies at the time of initial typing and, thereafter, following sensitization events. Laboratories must have a program to periodically screen serum samples from each patient.

Until recently, the federal standards required the use of lymphocyte targets for the detection of anti-HLA antibodies. However, in 1998 the Health Care Financing Administration (HCFA) ruled that solid phase immunoassays could be used for antibody screening, provided that each laboratory had documented validation of the assay's equivalent or superior sensitivity to cytotoxicity assays.

Screening procedures used in the pretransplant patient for antibody detection and identification must be designed to take all of the aforementioned characteristics into consideration. The approach is designed to uncover the total antibody content of a patient's serum and its characteristic reaction patterns. This information then is used to render a judgment regarding the interpretation of the pretransplant test between the serum of the patient and the cells of the donor with respect to the suitability of a given donor-patient combination. Therefore, it is essential to fully ascertain all of the properties of any antibodies involved. Are they public, private, or auto? Are they class I or class II? Do they react by the standard testing method, or do they require enhanced antiglobulin procedures? Finally, what is their expected frequency of positive reactions with the general population?

From these data, much important information may be derived, including an estimate of the likelihood of finding a cross-match–compatible donor among the random cadaver population in a given region. This, in turn, serves as the basis for a prediction of the waiting time for a given patient on the list. Additionally, from these data it is possible to predict the HLA types of donors who are likely to cause a positive cross-match and thus avoid unnecessary testing. Finally, the data derived permit the selection of optimal serologic cross-matching methods to ensure maximum sensitivity for antibody detection as well as the interpretation of any unexpected positive or equivocal results

Antibody Screening by Cytotoxicity (CDC) Methods

The patient's serum to be screened is tested serologically against a panel of cells whose antigenic composition has been predetermined according to rigorous standard typing criteria. The distribution of positive reactions is compared with the distribution of each of the known HLA antigens in the cell panel. The specificity of the antibody in the serum is equivalent to the antigenic specificity in the panel that most closely matches the distribution of positive reactions. The match will hardly ever be exact, because most frequently the sera contain a mixture of antibodies. In addition, the sera may contain a heterogeneous population of antibody molecules that have a spectrum of affinities for antigen. As a result, although the observed pattern may be similar to that for a particular antigen distribution, discrepancies will occur and a certain degree of uncertainty will be experienced.

The serologic methods used in screening patient sera should be the most sensitive available. One of the purposes is to learn as much as possible about the reactivity of a given patient's serum during a time of relative leisure and to predict how it will behave in the cross-match procedure performed under more stressful conditions and time constraints.

Ordinarily, the screening is performed using panels of purified T cells for the detection of class I antigens and purified B cells for the detection of class II antigens, as well as the patient's own cells to detect autoantibodies.

The practice of treating all sera prior to screening and cross-matching with dithiothreitol (DTT) to eliminate bothersome and clinically insignificant reactions from IgM antibodies has been advocated. Although this practice saves some time, it masks the presence of IgM antibodies and occasionally may dilute weak IgG antibodies, thus inhibiting their activity beyond their threshold of detectability. It is preferable to recognize that the serum of a particular patient may contain an IgM antibody so that it can be dealt with accordingly and the serum's reactivity interpreted intelligently.

Antibody Screening by Solid-Phase Immunoassays

Solid-phase immunoassays employ solubilized HLA antigens coupled to a solid matrix. The solubilized antigens are obtained after detergent solubilization of membrane antigens from pools of cells or from established cell lines. The solid matrix may be either the polystyrene wells of microtiter plates or latex beads. Coupling of the antigens to the matrix is accomplished by a “Capture antibody” consisting of an anti-HLA, class-specific antibody or by a chemical bond. Patient serum samples are incubated with the antigen-matrix and then washed to remove unbound antibodies. Specific antibody binding then is assessed by the addition of antihuman immunoglobulin conjugates. Conjugates used as reporter molecules for these assays usually are either enzymes that are used for enzyme-linked immunoassays (ELISA) or fluorochromes for immunofluorescent assays using flow cytometry as described below (18, 19).

In the ELISA assays, enzymes such as peroxidase or alkaline phosphatase permit visual detection of specific binding when their appropriate substrate is added. Cleavage of the substrate yields a chromogenic reaction product that can be visualized. In addition to the initial evaluation performed by laboratories to ensure the compatibility, equivalence and reproducibility of ELISA assays, each test run must include proper controls including those for nonspecific binding to verify the accuracy. In addition, patient sera must be tested at dilutions optimal for the assay in use. Results are dependent upon the optical density (OD) observed in the test wells. Cut-off values for reactivity above that of known negative controls must be established for each assay.

With some commercially available ELISA assays, computer software programs are available that allow the derivation of PRA values from the pattern of raw scores of ODs observed in the test wells. Assignment of antibody specificity for a given HLA antigen is performed by statistical programs in the software that determine the degree of concordance between the reactivity of a serum sample and the presence or absence of antigen. However, final results should still be reviewed by an experienced technologist familiar with cross-reactive epitope groups (CREGs) and linkage disequilibrium.

The complement-dependent cytotoxicity (CDC) assays have a number of disadvantages including the requirement for large panels of viable cells to cover all HLA specificities, the subjective reading of the tests, and the difficulty in differentiating HLA specificities in the presence of non-HLA cytotoxic antibodies. The solid-phase immunoassay offers several advantages in that it detects antibodies specific to HLA antigens, provides increased objectivity and reproducibility, it eliminates the need for a viable lymphocyte cell panel and can be performed in the face of immunosuppressive drugs, cytotoxic pharmaceuticals, and monoclonal antibodies. In addition, ELISA assays can be automated and thus, can provide significant cost savings and achieve faster turn around time. The clinical significance of the antibodies detected by solid-phase assays recently was investigated and showed a strong correlation with transplant rejection and clinical outcome. Some reports suggested that ELISA-detected anti-HLA is more clinically informative than CDC-PRA and correlates with the delineation of pre- and posttransplant immunologically high-risk allograft recipients.

Statistical Analysis of Antibody Screening Data

To interpret the results, the reaction patterns are normally analyzed using statistical methods that allow for a measure of this

uncertainty. If the level of uncertainty is small, being a result of chance alone, the similarity observed between the reactions of the serum and the distribution of an antigen in the cell panel probably is real. On the other hand, if the level of uncertainty is large, then the differences may be real and the plausibility of similarity diminishes. The most frequently used statistic for this purpose is the 2 × 2 contingency chi-square test.

The following is an example of a 2 × 2 analysis showing the comparison of the reactions of serum S.J. with the distribution of HLA-B7 in the test cell panel.

		HLA-B7 +	HLA-B7 −	
Serum S.J.	+	43	2	45
	−	3	27	30
		46	29	75

The theoretical distribution of X^2 against which the computed value will be compared is a continuous distribution. The data, on the other hand, lead to a discontinuous distribution of computed X^2 values. In practice, it is customary to correct for continuity by employing the Yates correction factor (20). For the example given above, the corrected X^2 is computed as follows:

$$X^2_{\text{CORRECTED}} = \frac{([(43\cdot27) - (3\cdot2)] - [0.5\cdot75])^2\, 75}{(45\cdot30\cdot46\cdot29)} = 52.01$$

If the two events — namely the occurrence of the antigen and the positive reactions of the serum — are independent, then the X^2 value will be low. On the other hand, if the occurrence of a positive reaction depends on the presence of the antigen in a given cell (or, stated another way, the antibody defines the specific antigen in the panel), then the X^2 value will be high and will approach the number of cells tested (75 in this example), but will never quite reach it. In addition, the customary use of the Yates correction factor to correct for discontinuity further lessens this chance. The hypothesis that the serum is detecting the given antigen in the panel is accepted or rejected depending on the probability corresponding to the calculated X^2 value obtained from a table of the X^2 distribution for 1 df (degree of freedom). When the X^2 is low, the probability of independence between reaction patterns and specificity is high. When the X^2 is high, the probability is low. A probability in the range of 0.01 to 0.001 or less indicates that the serum reaction patterns most likely correspond to the antigen distribution pattern. In the example given above, the probability is less than 0.001. Therefore, the hypothesis that the serum S.J. most likely contains antibodies with HLA-B7 specificity may be accepted.

In practice, these calculations are performed with the aid of computer programs. The X^2 values are then ranked. Frequently, sera will give high values with more than one antigen because they are multispecific. A crude approximation of the range of specificities that is adequate for clinical work can be obtained by performing a careful examination of the antigen distribution among the cells in the panel and the resultant X^2 values that are significant. Extensive screening against highly selected cell panels to confirm the specificities implied by the initial studies is usually not performed for diagnostic purposes in pretransplant patients. A careful consideration of the preliminary screening data along with the HLA type of the patient and the typing data of the previous grafts, if any, usually will give a good indication of the HLA antigens to be avoided in future grafts.

The knowledge of the specificity of antibodies in patients' sera is essential in making a sound decision regarding the suitability of a given donor for a particular patient. In addition, it serves as the basis for computing a predictive measure of the possibility of finding a suitable donor among the cadaveric organs available from the general population. The predictive value is given by the percent or PRA. It is computed from the frequency in the local population of the antigens defined by the patient's serum when it was tested and gives a reasonable reflection of how the sera might be expected to react in cross-matches with donors from the random population. Over the years, this value has become a standard parameter for judging a patient's transplantability and is used as one of the criteria that govern a patient's clinical management.

In summary, antibody screening can be defined as the systematic examination, under controlled conditions, of a serum for its content and immunoglobulin type of antibodies directed against surface antigens of nucleated cells, and the analysis of the resultant data. The information sought from these data determines the methods and extent of the systematic examination and the resultant data analysis.

The useful information required for patient antibody screening includes an exhaustive characterization of the antibody reactivity under a variety of very sensitive serologic methods, a general feeling for the scope and range of its reactivity with an array of different specificities, and a measure of the frequency of its reactivity with cells from the random population expressed as PRA.

Donor-Specific Cross-Matching

Donor-specific HLA antibodies in the serum of a recipient prior to transplant are especially significant. Such antibodies can readily combine with antigens on the grafted organ immediately after a blood supply is established and result in hyperacute or accelerated graft rejection. This can be avoided by performing a pretransplant cross-match in which a serum sample from the recipient is tested against lymphocytes obtained from the intended donor. A negative cross-match is an essential prerequisite for the performance of a transplant.

The serum sample from the patient must be one that best represents a sampling of the patient's anti-HLA antibody repertoire at the time the transplant will be performed. Ideally, this should be one collected just prior to surgery. Quite frequently, however, this may not be possible because of practical considerations. For example, the test itself is time consuming: if the intended recipient is not readily available to provide a serum sample, donor organ ischemia time might be prolonged to the point where irreversible damage could occur. Similarly, the recipient is usually one of several candidates; therefore the identity of the actual recipient might not be known. For these reasons, serum samples

are collected from potential recipients of cadaveric organs at monthly or more frequent intervals, especially those coinciding with the peak of an immunizing event such as a blood transfusion. These samples are carefully preserved in the frozen condition to be used for cross-matching with a potential donor.

Standard immunologic dogma dictates that once a patient formed an HLA antibody it would be of significance even after it had disappeared from the serum. Such antibodies might remain at undetectable levels or could be rapidly produced in a secondary response upon reexposure to antigen. Consequently, historical serum samples containing the maximum levels of antibody reactivity should be used routinely in cross-matches along with current samples. A positive historic cross-match would be considered to be a contraindication to transplantation with that particular donor. There is some controversy regarding the point that this does not hold true in cases in which HLA antibodies disappear prior to a first transplant (15). Nevertheless, it is worthwhile to consider that in such cases, the antibodies developed before the graft are most likely the result of exposure to HLA antigens present in blood transfusions. Apparently, the response produced to such an immunization does not result in long-term immunologic memory. Therefore, positive cross-matches with historical sera are frequently disregarded in such cases. On the other hand, in cases involving transplant immunization such as found in cases awaiting a second or third transplant, the reactions of historically positive sera with the potential donor are most important (16, 21). It seems that patients immunized or in conjunction with a previous graft develop strong immunologic memory and can produce a brisk secondary response when rechallenged with the offending antigen.

Traditionally, the cross-match is performed using a sensitive modification of the microlymphocytotoxicity test. Many centers employ the antiglobulin technique; others opt for an extended incubation of the antibody-lymphocyte mixture; still others favor some modification incorporating one or more wash steps prior to the addition of complement. However, any method selected should be capable of detecting the varieties of antibody found in the patient's serum during the pretransplant antibody screening. Thus the importance of pretransplant screening cannot be overemphasized. This procedure allows a detailed and accurate assessment of the antibody repertoire carried by the patient in a relatively leisurely manner rather than under the stresses associated with a cadaveric donor organ harvest and the overzealous anticipation of a possible transplant.

When using a cadaveric donor, the lymphocytes of choice are prepared from excised lymph nodes or a section of spleen obtained in conjunction with organ harvesting. The nodes should be kept moist after their removal and during delivery to the laboratory. Some workers advocate the use of peripheral blood from the potential donor in order to save time. This is a debatable point. The peripheral blood cells of patients who have been kept on life support systems may be unsuitable for use in the cytotoxicity test for a number of reasons, including the effects of antemortem steroids administered to reduce cerebral edema. Consequently, uninterpretable results may often be obtained, requiring the tests to be repeated. Thus, little time is saved and the expense of duplicate work is incurred.

It is best to perform the cross-match using pure T cells as targets (22). This is good practice, because anti-HLA class I antibodies are the most frequently encountered and the most clinically significant. In addition, many types of patients such as those with lupus erythematosus, certain diabetics, and others often demonstrate nonspecific autoreactivity. This is most frequently directed against B cells and seems to have no clinical significance. If mixtures of T and B cells are used in the cross-match with the sera of such patients, they would give a positive result. Therefore, a transplant would be denied because of a false-positive reaction.

In contrast to nonspecific B-cell antibodies, those directed against class II HLA antigens are of definite clinical relevance; associated with increased graft rejection and graft loss. In addition, positive B-cell cross-matches have been implicated in early rejection of both cardiac and lung transplants. Pretransplant HLA class II antibody screens should be designed for their detection and identification. In addition, cross-matches with the serum of patients containing such antibodies should be performed using suspensions of B cells.

Application of Flow Cytometry Techniques in Potential Solid Allograft Recipients

The application of flow cytometric techniques in the pretransplant work-up of potential solid allograft recipients has grown rapidly during the past decade. One of the most important applications is the flow cytometric crossmatch (23, 24). CDC is adequate for routine crossmatching in that it detects most antibodies that are responsible for hyperacute graft rejection. However, this method does not detect low concentrations of antibodies that may not be the cause of hyperacute rejection; but nevertheless may be responsible for an accelerated rejection of the graft.

Several reports have shown that low levels of alloantibodies, undetectable by CDC, could be detected by flow cytometry. This technique is more sensitive since it is not dependent on complement binding. In these studies, renal transplant recipients exhibiting a negative CDC assay but a positive flow cytometric cross-match were more likely to experience early accelerated rejection episodes, an increased number of such episodes and early graft failure. The patient groups more likely to experience this type of rejection were those individuals who were previously alloimmunized from a rejected transplant, multiple pregnancies, or multiple blood transfusions. Thus the flow cytometric cross-match may be important in identifying a subgroup of patients who would be at increased risk for graft failure or for a more complicated posttransplant clinical course.

The flow cytomeric cross-match offers additional advantages over CDC. It can identify IgG antibodies specifically without the additional treatment of patient's serum with DTT. Furthermore, cell viability is not as critical for the flow cytometric cross-match because the dead cells can be gated-out by light scatter and/or viability stains. Furthermore, it is unnecessary to prepare separate suspensions of T cells and B cells; because these can be identified by specific monoclonal antibodies. Finally, the interpretation of the flow cytometric cross-match tends to be more objective in that CDC assays rely on the visual scoring of stained cells that can be influenced to a significant degree by reader bias.

Reagents

The most critical reagent for the flow cytometric cross-match is the secondary antibody. This reagent is used to identify the human

immunoglobulin that is specifically bound to the target cell. Thus, the reagent selected should provide optimum specificity with low background staining. In addition, the secondary antibody must not cross-react with mouse or horse immunoglobulins. Importantly, each lot of secondary antibody should be tested using a checkerboard titration to determine the optimum working dilution. Depending upon the cell population to be evaluated, appropriate monoclonal secondary antibodies are selected based on their specificity for the desired cluster antigen and their fluorochrome tags. Most laboratories utilize a fluorescein-conjugated secondary antibody.

The flow cytometric cross-match requires both a positive and a negative control serum. The most common sources of target cells are mononuclear cells derived from peripheral blood, lymph nodes, or spleen. Only a small number of cells is needed. One of the important variables is the ratio of patient's cells to the volume of patient's serum used in the test. Increasing the volume of serum used per tube or decreasing the number of cells per tube will increase the sensitivity. Thus, it is important that the laboratory establish its staining protocol and strictly adhere to methods that provide the accuracy and consistency of cell and serum concentrations and volumes.

Critical Points in the Flow Cytometric Crossmatch Procedure

As indicated above, the ratio of serum volume to number of cells is important because it can affect the sensitivity of the assay. A second point relates to the number of wash steps used following the primary antibody incubation. Washing is done in order to ensure that residual immunoglobulins are removed prior to the addition of the secondary antibody. Insufficient washes may result in a weak or even a false-negative cross-match. A third point relates to the nature of the secondary antibody used. It is important that the anti-human immunoglobulin reagent does not cross-react with mouse immunoglobulins because the directly conjugated monoclonal antibodies used to define T- and B-cell populations are of mouse origin. Cross-reactivity with these immunoglobulins will render the crossmatch uninterpretable. Finally, in preparing patient and control sera it is useful to airfuge (100,000 g for 15 min.) all sera prior to testing. This type of ultracentrifugation removes immunoglobulin aggregates that may produce nonspecific background staining, particularly with B cells, because of Fc receptor binding.

Standards and Controls

Each laboratory must establish its own set of standards and controls for the performance and the interpretation of the flow crossmatch. The choice of adequate controls is very important. Care must be taken to ensure that the normal human serum selected for use as a negative control is free of any anti-HLA antibodies. This reagent can be either a pooled or a single-donor serum. However, it must be tested thoroughly with a number of different normal cell donors to ensure that the reagent has no reactivity with human lymphocytes.

The positive control is usually a pool of sera from patients who are known to have high levels of anti-HLA antibodies. Such a pool can be diluted several fold and stored for use. This reagent should be used at a dilution that yields a reasonable shift in fluorescence (displacement between 100 and 300 channels) that is below a level of saturation.

Despite the potential of flow cytometry to provide quantitative data, the flow cytometry cross-match is used mostly as a semiquantitative test; i.e., negative, weak, or strong positive. Fluorescent intensity results may be expressed as either channel shift or fluorescence ratio and can show a high degree of variability among different laboratories. Using a cutoff value of two standard deviations, one can define a channel value that can be used as a discriminator for positivity.

Alternative methods also can be used including the utilization of semi-quantitative methods for determining the relative immunofluorescence of the attached molecules. Using a mixture of beads, each with a different level of fluorescence, one can perform a relative calibration of the flow cytometer. Subsequent channel values then can be converted to the number of Molecules of Equivalent Soluble Fluorochrome (MESFs). Results then are reported in MESF units and the determination of positive is made by indicating a minimum MESF value above which a cross-match is considered positive. At present, it is uncertain whether results expressed as MESF units are more advantageous and whether they can reduce the variability among laboratories than results expressed as channel shift. The American Society for Histocompatibility and Immunogenetics has published standards for the performance of flow cytometric cross-match for transplantation.

Results and Interpretation

Data acquisition is performed usually after setting gates around lymphocyte populations. Post acquisition analysis is performed next by gating on a selected fluorescence specific for T- or B-cell populations. In a three-color cross-match, fluoresceine isothiocyanate (FITC) can be used as the conjugate for the anti-IgG reagent; in conjunction with phycoerythryn-conjugated CD20 (specific for B cells) and PerCP (Peridinin Chlorophyll protein)-conjugated CD3 (specific for T cells). Following acquisition of lymphocyte-gated data, individual regions are set around CD3- and CD20-positive cells. This is performed by displaying a dot plot of CD3 versus CD20 (Fig. 37.2). Subsequently, analysis of the FITC staining can be made by displaying a single parameter histogram for the CD3 region or from the CD20 region (Fig. 2). Following completion, the median channel value for each cell population is recorded. This three-color combination allows for the simultaneous detection of alloantibodies reacting with T cells and B cells and eliminates background staining from NK cells and monocytes. Each laboratory must establish its own cutoff for determining the magnitude of channel displacement to be considered a positive result.

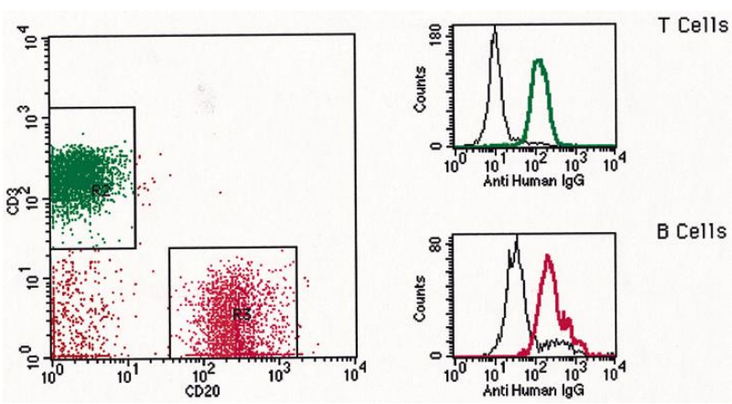


FIGURE 37.2. Three-color flow cytometric crossmatch. **Left:** Dot-plot displaying CD-20 PE staining of B cells (x axis) versus CD3-PerCP staining of T cells (y axis). **Right:** single parameter histogram of T and B cells. Staining fluorescence with the normal human serum control is shown in black; staining fluorescence observed with a positive serum sample is shown in green and red color for T and B cells respectively (Becton Dickinson FACSCalibur Cytometer; 4-decade log scale; 1056-channel resolution; CellQuest analysis).

One of the general problems in evaluating the significance of crossmatching by any technique is in the establishment of anti-HLA antibody specificity. Thus, in addition to crossmatching, routine antibody screening also should be performed by sensitive techniques that can detect antibodies specific to HLA class I and class II antigens. This knowledge is key to interpreting cross-match results. The second problem is the determination of the clinical relevance of a positive flow cytometric crossmatch.

Knowledge of the patient's clinical history is most important. There certain individuals who, under appropriate immunosuppression, could achieve a good transplant outcome despite a positive-flow cytometric cross-match. Thus a positive-flow cytometric cross-match will not always constitute a contraindication to transplantation. The identification of this type of patient depends heavily on the patient's sensitization history, as well as biological and clinical parameters.

In addition to the cross-match, flow cytometry also can provide very useful information on routine antibody screening. Significant levels of anti-HLA antibodies can be detected by flow cytometry in patients whose cytotoxic PRA was 0%. Flow cytometric PRA can be performed using pooled cells that are selected based on their CREG antigens. More recently, a microparticle assay utilizing purified HLA class I and HLA class II proteins coated onto beads has been used successfully to detect antibodies to HLA class I and class II antigens. Class I and class II beads can be admixed so that patient serum can be simultaneously assessed for the presence of class I and/or class II. Newer configurations of the microparticles have been designed recently, which can identify the specificity of the antibodies.

Evaluation of Tolerance to Donor HLA Antigens and Post Transplant Monitoring

Recipient's T cells can recognize allogeneic HLA molecules by two distinct routes: either directly as intact molecules or indirectly as peptides produced as a result of antigen processing. The direct recognition pathway is thought to be the primary mediator of acute allograft rejection and can be observed as antidonor mixed-lymphocyte culture and cell-mediated lympholysis reactivity exhibited by the recipient's T cells. Experimental work to date has shown that indirect presentation of allogeneic MHC molecules can occur not only *in vitro* but also *in vivo* where it can influence graft rejection. The precise role of this route of antigen presentation is yet to be established, although there is strong circumstantial evidence that this pathway is important in the initiation of chronic or late graft rejection and the induction of unresponsiveness following allogeneic transfusion. Late in the transplanted organ's course, donor MHC class II positive-passenger lymphocytes will have been depleted from the graft, but cells expressing MHC class I and class II alloantigens are likely to be shedding these antigens into the circulation. Recipient antigen presenting cells (APC) would then be able to process these allo-MHC molecules and present them as peptides to recipient CD4+ T cells. This would result in the activation of cellular and humoral effector mechanisms. That such a scenario highly likely is evidenced by the fact that tolerizing strategies targeting T cells that recognize MHC antigen indirectly are being explored as possible therapeutic and/or prophylactic modalities.

The development of useful immune or surrogate markers of tolerance to donor HLA antigens that are measurable by standardized and reliable biological assays is very desirable. Such markers should be capable of measuring the effect of immunomodulating intervention on clinical outcome in transplantation. The measure of immune responses should include the

quantitation of lymphocyte responsiveness to donor alloantigens using methods such as proliferation assays, cytotoxicity activity, *de novo* antibody production, and cytokine secretion.

Several assays are available that can be used in monitoring the posttransplant immune status of patients receiving solid organ grafts. These include the monitoring of the levels of anti-HLA antibodies directed against graft antigens, the monitoring of the appearance of various forms of soluble HLA antigens discussed below, as well as several cellular assays.

Although antibody screening is a mandatory pretransplant procedure, antibody screening also is useful for posttransplant monitoring. The advent of antibody detection using soluble HLA antigens and ELISA or flow cytometric techniques has allowed for the reliable detection of all classes of anti-HLA antibodies in the posttransplant period. These methods circumvent the constituents found in post-graft sera that make the cytotoxicity procedure unreliable. These include such elements as immunosuppressive drugs and anticomplementary proteins.

The data being accumulated thus far using ELISA techniques indicate that screening for newly formed HLA antibody may be a useful procedure for graft monitoring. It has been demonstrated that the production of HLA antibody directed against graft antigens is correlated with the incidence of severe graft rejection episodes that can lead to decreased graft survival. This allows for the identification of patients at risk and allows for the more rapid employment of therapeutic steps to prevent excessive damage to the graft.

Cellular assays might include the monitoring of graft-infiltrating cells obtained from biopsy material and can be accomplished by evaluating *in vitro* lymphocyte growth in response to IL-2. Another assay might consist of measuring the number of T cells participating in the indirect pathway of allorecognition. This can be measured by limiting dilution analysis of the precursor frequency of $CD4^+$ cells recognizing synthetic allopeptides that correspond to polymorphic regions of the mismatched donor MHC class II antigens.

Currently, the use of assays such as the ones described is experimental. Data collected from retrospective studies strongly suggest that this type of monitoring should be highly useful in the management of the posttransplant patient. Such assays would be a more direct indication of the patient's response to the allograft at the level of the immune response. Current monitoring strategies consist of biochemical assays to measure graft physiology and in most cases, graft biopsies. Such biopsies are invasive and may be subjective and prone to error. In addition, once the biopsy is positive the damage to the graft already has occurred. Therefore, immunologic assays would be superior. However, rigorous prospective studies are needed to evaluate the clinical utility of these assays.

Soluble HLA Antigens and Their Peptides

The relationship of soluble HLA (sHLA) proteins to the graft-host relationship is a double-edged sword. On one hand, several laboratories have demonstrated tolerance to allograft following the pretreatment of experimental animals with soluble forms of MHC class I and class II antigens, or derivative donor allopeptides. Soluble specific HLA complexes cause allospecific $CD8^+$ T cells to undergo apoptosis *in vitro*. It was recently suggested that this could be relevant for tolerance induction after liver transplantation. Hepatocytes are an important source of sHLA class I molecules, which *in vivo* may induce apoptosis of alloreactive T cells in a similar way. It now appears that sHLA may be part of an elaborate system for regulation of cytotoxic T lymphocyte (CTL) by soluble forms of membrane bound HLA proteins that may be involved in the maintenance of peripheral tolerance (26). Furthermore, soluble HLA-C molecules can form complexes with the NK receptor. This interaction leads to inhibition of the NK cells (25).

On the other hand, there is evidence for donor-specific HLA class I antigen release in organ transplant patients as an indicator of transplant rejection. ELISA measurements of serum for sHLA class I mismatched antigens in patients with acute or chronic rejection showed increased serum levels of donor sHLA-A and -B in heart, lung, and kidney recipients. No detectable donor sHLA (over pretransplant background) could be found in the sera of nonrejecting recipients. However, not all rejections in all patients lead to detectable soluble donor HLA release (27). The reason for this remains unknown.

Recent studies have suggested a possible explanation for the “good” versus “bad” soluble HLA proteins. Biochemical studies using the western blot technique to define the molecular weights of various forms of sHLA class I released in liver transplant recipients, indicated that a proteolytic cleavage mechanism might be operative in acute rejection. Metalloenzymes on the membrane of the graft parenchymal cells might be responsible for the release of rejection-associated soluble HLA molecules. These soluble β_{2m}-free truncated forms of HLA class I molecules are highly susceptible to proteolysis and, because they contain polymorphic donor peptide sequences, would be a readily available substrate for spreading the T helper cell alloresponse. In contrast, the β_{2m}-associated forms that are released by well functioning organ allografts, may promote allograft tolerance by inhibiting “direct pathway” CTL responses, while activating a suppressive “indirect pathway” T-cell response to a few immunodominant epitopes.

Clinical Immunosuppression

A discussion of laboratory testing in organ transplantation would not be complete without describing some of the pharmacological approaches to immunosuppression, their sequelae, mode of action, and possible laboratory monitoring. In transplantation, there are three main uses of immunosuppressants: to (i) induce initial acceptance of a graft (induction therapy) using heavy doses to prevent rejection and favor the tolerance response; (ii) to reverse a rejection episode (acute-rejection therapy) using heavy doses to reverse an established rejection and; (iii) for maintenance, using moderate doses to prevent rejection but preserve host defense mechanisms.

Several immunosuppressive drugs are used in transplantation, all of which can be toxic by virtue of the induction of immunodeficiency. This may result in infection as well as certain malignancies such as skin cancer, and B-cell lymphoma. In addition, many of these drugs may have a direct deleterious effect on nonlymphoid cells such as kidney or bone-marrow stem cells. The

list of the currently approved immunosuppressants includes azathioprine (AZA), corticosteroid, Cyclosporine (CsA), Tacrolimus (FK506), Atgam (Upjohn, Kalamazoo, Michigan), anti-CD3 (OKT3), and mycophenolate mofetil (Table 37.4).

TABLE 37.4. CLASSIFICATION OF IMMUNOSUPPRESSIVE AGENTS

Pharmacological Agents	
1. Immunophilin binding drugs	a) calcineurin inhibitors: cyclosporine[a], FK-506[a]
	b) calcineurin independent actions: rapamycin
2. Inhibitors of cell division/nucleotide metabolism	a)nonselective antiproliferative and cytotoxic drugs: azathioprine[a], cyclophosphamide[a]
	b)lymphocyte-selective specific drugs: mycophenolate mofetil[a]
3. Corticosteroids[a]	
Biological Agents	
1. Antibodies	a) polyclonal antilymphocyte: horse anti-thymocyte globulin[a], rabbit antilymphocyte or antithymocyte globulin
	b) murine monoclonal: anti-CD3 (OKT3)[a], anti-CD4 (OKT4), anti-LFA, anti-ICAM
	c) humanized monoclonals: anti-IL-2R α chain
2. Fusion proteins	CTLA4Ig, IL-2 toxin
3. Cytokine and cytokine receptors	
4. Peptide therapy	

[a] Agents released in the United States of America

AZA is a purine analog and an antiproliferative agent that suppresses the proliferation of activated B and T lymphocytes. It now is used most often as a supplement to CsA, but was used widely along with steroids for maintenance immunosuppression before the introduction of CsA. The major nonimmune toxic side effect of AZA is marrow suppression. Other important side effects include hepatotoxicity and the usual immunodeficient toxicity.

Corticosteroids are used for maintenance immunosuppression as well as for treatment of acute rejection episodes. The pharmacological effects of steroids are complex and their mechanism of action is not fully understood. The side effects of corticosteroids are numerous, dose-dependent, and contribute significantly to the morbidity of transplant patients.

CsA is a natural cyclic amino acid peptide, isolated from fungi. It has been an essential component in most immunosuppressive protocols in transplantation. FK506 can now be substituted. CsA acts by binding to cyclophilins and the resultant CsA-cyclophilin complex is the active drug. This complex inhibits calcineurin (CN), a serine protein phosphatase. Inhibition of CN prevents or reduces the activation of several CN-dependent transcription factors such as IL-2 and other cytokine genes, as a result, CsA inhibits the transcription of several T-cell cytokines and the proliferation of T cells. The most important nonimmune side effect of CsA is nephrotoxicity, which may be confused with rejection in renal transplant recipients. Levels of this drug in patients' sera must be carefully monitored biochemically to ensure that optimum but not excessive blood levels are maintained.

Tacrolimus (FK506) is a macrolide antibiotic. Its mode of action is analogous to CsA. FK binds to an immunophilin, FK-binding protein. The resultant complex then interacts with CN and inhibits its action in the same manner as CsA. FK is much more potent and toxic than CsA. It has been used most extensively in liver transplantation as an alternative to CsA. There is currently no definitive evidence that FK is more effective than CsA in renal transplants for patient and graft survival, but clinical trials are in progress.

Atgam is a horse serum containing polyclonal antilymphocyte antibodies. The immunosuppressive activity of Atgam is mediated through the depletion of circulating T lymphocytes. Atgam also blocks the function of T cells and inhibits their proliferative activity. Because Atgam severely impairs cellular immunity, patients treated with Atgam are prone to develop opportunistic infections and malignancies. Other important side effects include skin rash, serum sickness, leukopenia, and thrombocytopenia.

OKT3 is a murine monoclonal antibody against the CD3 complex of molecules on the surface of T cells. OKT3 was released for the treatment of acute rejection, where it is very effective. It also has been widely used for induction immunosuppression in sensitized patients and in patients with delayed graft function. However, concern over the emergence of early proliferative disease in patients treated with OKT3 has somewhat limited its prophylactic use. The immunosuppressive effects of OKT3 are mediated mainly through the depletion of circulating T lymphocytes and modulation and removal of CD3 molecules from the cell surface, thereby making the T cells nonfunctional. The most common side effect of OKT3 is the cytokine-release syndrome. This is characterized by fever, chills, nausea, diarrhea, and myalgia. As with Atgam, OKT3 therapy significantly impairs cell-mediated immunity. As a result, patients receiving OKT3 also are predisposed to opportunistic infections and malignancies. Patients undergoing Atgam and/or OKT3 therapy should subjected to the regular monitoring to assess the number of T cells in their circulation.

Mycophenolate mofetil is a new immunosuppressive agent that has recently been introduced in transplantation. Its active metabolite is mycophenolic acid (MPA). It selectively inhibits lymphocyte proliferation and has been developed as a replacement for AZA. *In vitro*, MPA has been shown to inhibit antibody formation and inhibit cytotoxic T-cell generation. In clinical trials, the incidence of acute rejection in patients treated with

mycophenolate mofetil is reduced by 40% to 50% as compared to placebo or AZA. It is hoped that by reducing the incidence of early graft rejection, mycophenolate mofetil also may improve the long-term outcome of the renal transplant recipients. The major side effects are gastrointestinal. In contrast to AZA, mycophenolate mofetil does not induce nephrotoxicity, hepatotoxicity, or significant bone-marrow suppression.

There has been some controversy regarding the benefit of HLA matching in the era of newer modalities of immunosuppression. This occurred first when cyclosporine, a remarkable immunosuppressive drug that made liver transplantation a reality was introduced. However, even then studies showed that HLA matching could improve kidney graft survival by 10% to 15%. Therefore, even with the use of powerful new immunosuppressives graft survival could be improved even more by donor selection based on HLA matching (Table 37.5).

TABLE 37.5. SELECTIVE MATCH GRADES FOR PLATELET TRANSFUSION[a]

Match Grade	Extent of HLA Matching	Example		
A	Four HLA-A,B antigen match	A2,30,B12,7	→	A2,30,B12,7
B1U	One HLA antigen is unknown	A2,?,B12,7	→	A2,30,B12,7
B2U	Two HLA antigens are unknown	A2,?,B12,?	→	A2,30,B12,7
B1X	One antigen is cross-reactive	A2,30,B12,**27**	→	A2,30,B12,7
B2X	Two antigens are cross-reactive	A2,**33**,B12,**27**	→	A2,30,B12,7
BUX	One antigen unknown, one cross-reactive	A2,**?**,B12,**27**	→	A2,30,B12,7
C	One major antigen mismatch	A2,30,B12,**5**	→	A2,30,B12,7
D	Two or more major mismatches	A2,**10**,B12,**5**	→	A2,30,B12,7

[a]Adapted from Dahlke MB, Weiss KL. Platelet transfusion from donors mismatched for crossreactive HLA antigens. *Transfusion* 1984;24:299–302.

TRANSPLANTATION: BONE MARROW AND STEM CELLS

Part of "37 - HLA: Applications"

The HLA system is of vital importance in bone marrow transplantation. This subject will be covered in detail in Chapter 38.

PLATELET TRANSFUSION

Part of "37 - HLA: Applications"

Patients receiving repeated infusions of platelet concentrate over long periods of time frequently become refractory. They fail to achieve an incremental increase in their platelet count proportional to the number of platelets transfused. It has been shown that this condition results from the development of antibodies frequently produced in response to class I HLA antigens expressed on the surface of platelets. HLA class II antigens play no role in this phenomenon.

The use of HLA-matched platelets is beneficial in many of these cases for achieving normal platelet survival in the immunized patients and the prevention of immunization in the immunologically virgin patient (28). However, exact HLA matching for platelet transfusion may not be possible for practical reasons. Fortunately, not all HLA antigens are expressed on platelets with equal strength. Furthermore, it has been shown that matching for some of the antigens as outlined in Table 37.5, together with matching for CREGs is adequate to achieve successful clinical results (29).

The selection of the type of matched product — whether matched for private antigens or matched for CREGs — depends entirely upon the patient. Platelet recipients can be divided into three groups. The first consists of patients who are nonresponsive and never become refractory. The second group are those who become refractory to random donor platelets but tolerate HLA-matched products. These may be further subdivided into those who require HLA-identical products from siblings and those in whom platelets matched according to the schemes outlined above result in adequate increments and provide adequate hemostatic function. The third group are those who become refractory very quickly after a minimal number of random units and cannot tolerate even very well-matched platelets.

Another strategy for obtaining good platelet survival in refractory patients is based on good antibody screening and identification. Extensive antibody screening using either cytotoxicity or ELISA techniques is carried out on patient serum samples collected over a period of time. All of the antibody specificities are identified and listed. Platelet donors are selected who carry none of the specificities against which the patient has antibodies. Using this scheme, the data indicate a good *in vivo* response for HLA nonmatched but antibody-compatible platelets as opposed to antibody-incompatible platelets (30). This system appears to achieve the best results except in those patients that are refractory for reasons unknown.

DISEASE ASSOCIATION

Part of "37 - HLA: Applications"

The association of HLA types with various diseases has been the subject of intense investigation (31). Originally discovered as a chance observation and later enforced by the finding of a strong association between ankylosing spondylitis and the presence of the infrequent antigen HLA-B27, this line of investigation has led to a clearer understanding of some of the disease processes themselves. For example, diabetes mellitus has been reclassified into two distinct types. Juvenile-onset, insulin-dependent diabetes, or type I diabetes mellitus, is strongly associated with HLA; whereas type II or maturity-onset diabetes shows no apparent association. As a result of these findings it is now recognized that each of these types is a clinically distinct disease entity with its own pathogenesis.

In discussing the concept of HLA and disease, it is important to recognize the distinction between association and linkage. It is tempting to speculate that the reason for the observed relationships

between the occurrence of a given disease and the presence of a particular HLA antigen is linkage. However, linkage implies a formal genetic analysis of family inheritance patterns and indicates that two loci are situated closely enough on the chromosome to allow crossing over to occur less than 50% of the time. With the exception of diseases that exhibit familial tendencies, most of the studies carried out with respect to HLA and disease have been population studies that do not permit formal genetic analysis.

Indeed, such analyses have shown that genes responsible for hemochromatosis, for deficiencies in 21-hydroxylase (resulting in congenital adrenal hyperplasia), and for C2 are linked to HLA. Further studies resulting from these observations have revealed that the loci for 21-OH and C2 are not only linked to the MHC but are within the MHC itself. However, the majority of diseases studied in this manner exhibit association rather than linkage. Association of a particular HLA antigen with a certain disease is a statistical event established by comparing the frequency of the antigen in the disease population with its frequency in the normal population. The normal population in this case should consist of random individuals matched for racial composition with the disease population. A comparison is made using the X^2 statistic in a 2 × 2 contingency table and corrected for the number of antigens tested.

The strength of the association can be obtained by calculating the relative risk (RR) or the statistical chance that a patient with a certain disease will have the given HLA antigen. This may be computed either by the method of Woolf (32) or that of Haldane (33). Both of these are shown in the example given below. This example demonstrates the steps in the analysis of data obtained in a hypothetical study of the incidence of the antigen HLA-B27 among 40 patients suffering from ankylosing spondylitis and 904 normal control subjects.

	HLA-B27 +	HLA-B27 −	
DISEASE	[a] 35	[b] 5	[a+b] 40
CONTROL	[c] 67	[d] 837	[c+d] 904
	[a+c] 102	[b+d] 842	[n] 944

$$X^2 = \frac{(ad - bc)^2 n}{(a + b)(c + d)(a + c)(b + d)} =$$

$$\frac{[(35)(827) - (5)(67)]^2(944)}{(40)(904)(102)(842)} = 254.9$$

$$RR = \chi = \frac{ad}{bc} = \frac{(35)(873)}{(67)(5)} = 87.4$$

according to the method of Woolf; and

$$RR = X = \frac{[2(a + 1)][2(d + 1)]}{(2c + 1)(2b + 1)}$$

$$= \frac{(2)(35 + 1)(2)(837 + 1)}{(2)(67 + 1)(2)(5 + 1)} = 73.9$$

according to the method of Haldane. The X^2 value of 254.9 is highly significant and indicates that the higher incidence of HLA-B27 among the patient population is not the result of chance. In this example, both populations were tested only for the presence or absence of HLA-B27; consequently no corrections are required. However, when studying a disease for the first time, it is customary to test for a battery of HLA antigens. Because one would expect that some deviation from the normal distribution of antigen frequencies could occur from chance, the X^2 is corrected by dividing it by the number of antigens that were tested.

The RR of 87.4 given by the method of Woolf agrees with that published. The method of Haldane is designed to correct for small sample sizes and gives a value of 73.9. This is slightly less than the published value but is nevertheless quite significant in terms of its order of magnitude.

Associations have been shown to be of two types. Those in which the disease is associated with class I antigens are shown in Table 37.6. Those having strong associations with class II antigens are given in Table 37.7. Interestingly, the diseases have a strong immunologic flavor, being either diseases of the immune system itself or having autoimmunity as their underlying theme.

TABLE 37.6. SOME PROMINENT ASSOCIATIONS OF DISEASES WITH CLASS I ANTIGENS[a]

Disease	HLA	RR
Idiopathic hemochromatosis	A3	6.77
Vitiligo (Yemenites)	B35	13.9
Acute anterior uveitis	B27	10.4
Ankylosing spondylitis	B27	87.4
Reiter's disease	B27	37.0
Duodenal ulcer	B35	2.7
Subacute thyroiditis (de Quervain)	B35	13.7
Psoriasis vulgaris	Cw6	13.3

[a]From Zmijewski CM. HLA and disease. *CRC Crit Rev Clin Lab Sci* 1984;20:285–370.

TABLE 37.7. SOME PROMINENT ASSOCIATIONS OF DISEASES WITH CLASS II ANTIGENS[a]

Disease	HLA	RR
Allergy (Ra 5 response)	DR2	19.0
Multiple sclerosis	DR2, DQ6	4.1
Narcolepsy	DQ6	>38
Celiac disease	DQ2	>250
Goodpasture's disease	DR3	15.0
Insulin dependent diabetes	DQ8	14
Systemic lupus erythematosus	DR3	2.6
Rheumatoid arthritis	DR4	9
Insulin dependent diabetes	DR4	15.4
Pemphigus vulgaris	DR4	14.4
Juvenile rheumatoid arthritis	DR8	8
Pernicious anemia	DR5	5.4
Nephrotic syndrome (steroid responsive)	DR7	5.9

[a]From Zmijewski CM. HLA and disease. *CRC Crit Rev Clin Lab Sci* 1984; 20:285–370.

Of particular note are rheumatoid arthritis (RA) and IDDM or type 1 diabetes. These diseases have been studied extensively at the molecular level and have yielded a great deal of innovative information. Both RA and IDDM are strongly associated with DR4. However, in the case of RA the association appears to be with the

DR molecules themselves; whereas in IDDM the association may be because of interactive effects of both DR and DQ loci (34).

HLA-associated diseases are caused by an interplay of different genes and environmental factors, in which HLA complex genes most often confer the strongest genetic predisposition. This can be illustrated by taking type 1 diabetes as an example. The concordance rate for this disease among monozygotic twins is between 35% to 50%. However, the concordance rate among HLA identical siblings is approximately 15% to 25%, compared to approximately 1% for siblings differing at both HLA haplotypes. This indicates that genes in the HLA complex are the most important for genetic predisposition, an idea that is consistent with the results of a genomewide search for susceptibility genes (34, 35).

Several autoimmune diseases seem to be associated primarily with peptide-presenting HLA molecules. As an example, susceptibility to develop RA is strongly associated with the DRB1*0401, *0404, *0405, and *0408 alleles (see Fig. 35.1). On the other hand, DRB1*0402 is not associated with RA and may instead confer protection. Interestingly, these genes all encode a very similar stretch of amino acids at DRß chain residues 67 to 74, with one important exception. At residue DRB71, the former genes encode the positively charged amino acids lysine or arginine. In contrast, DRB1*0402 encodes negatively charged glutamic acid at this position. Residue 71 points inward towards pocket four of the peptide-binding cleft of DR molecules. It may thus play a critical role binding peptides to the DR molecule. This suggests a direct involvement of given peptide-presenting DR molecules in the pathogenesis of RA (34 and references therein).

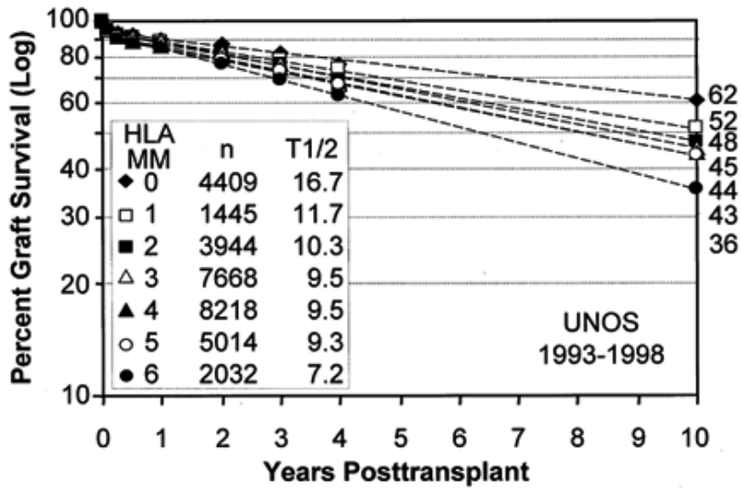


FIGURE 37.1. Survival of cadaveric kidneys according to HLA matching assuming a constant failure rate after the first year for all cadaver transplants performed between 1993 and 1998. This is based on UNOS Scientific Registry Data as of March 1999. Provided by M Cecka.

Similarly, in type 1 diabetes it appears that both susceptibility and protection are associated primarily with some class II HLA molecules. Because of the strong linkage disequilibrium that exists between genes in the HLA complex, it often is difficult to dissect the genes primarily involved in the disease from those genes that are secondarily involved due to gametic association. Susceptibility and protection are thought to be the result of the occurrence of certain HLA-DQ heterodimers. In most populations a stronger susceptibility to type 1 diabetes is found in DR3,DQ2/DR4,DQ8 heterozygotes, than in DR3,DQ2/DR3,DQ2 or DR4,DQ8/DR4,DQ8 homozygotes. Furthermore, disease susceptibility can be observed in patients from different populations who have combinations of the same DQA1 and DQB1 genes in either the *cis* or in the trans position. The susceptibility associated with the DQ heterodimer is strongly influenced by the accompanying DR4 antigen in the haplotype. Susceptibility is usually seen when the DRB*0405, *0402 or *0401 code for DR4. In contrast, little susceptibility but rather protection against the development of the disease is observed when either DRB*0406, *0403 or *0404 code for the accompanying DR4 antigen (34 and references therein).

It currently is believed that HLA-associated predisposition to develop type 1 diabetes is mainly determined by the combination of DQ and/or DR molecules of a particular individual in whom certain DQ and DR4 molecules appear to be most characteristic (34, 35). This may explain differences between populations with respect to susceptibility associated with particular DQ or DR molecules. For example, in most white populations, DQ8 is associated with susceptibility because it frequently occurs together with susceptibility-associated DR4 subtypes. In contrast, in the Chinese susceptibility may not be associated with DQ8 because in this population DQ8 often occurs together with the protective DRB*0406 or *0403 molecules.

Many of the HLA-associated diseases are associated with HLA molecules that share certain pockets in their peptide-binding clefts, which may play a critical role in determining which autoantigen-derived peptides can be bound. As discussed previously, the susceptibility-associated DR molecules in RA may preferentially bind an arthrithogenic peptide at pocket four that contains a negatively charged amino-acid residue. In diabetes, the DQ molecules associated with susceptibility often share an amino acid other than aspartic acid at position 57 on the DQß chain that is located in pocket nine of the peptide-binding cleft (36, 37). On the other hand, those DQ molecules associated with protection carry aspartic acid at this position. Although there are exceptions to this rule, the importance of positively or negatively charged amino-acid residues at position 57 suggests an important role for pocket nine of DQ molecules in the binding of pancreatic ß-cell–derived peptides; thereby conveying either susceptibility or resistance.

The biological mechanisms underlying the associations between HLA antigens and various diseases is not understood. A number of possibilities have been suggested, including antigenic mimicry, faulty immune response, or other susceptibility genes in gametic association with HLA. Each of these can be supported with a good deal of positive evidence. However, in each case there is an equally impressive body of contradictory evidence. Some of the newer approaches at the molecular level are producing an explanation for this most intriguing mystery.

PARENTAGE TESTING

Part of "37 - HLA: Applications"

The extensive polymorphism of the HLA system makes it a powerful tool in the resolution of litigation resulting from disputed parentage (38). The method is based on the logical genetic premise that cell membrane antigens or any other readily discernible phenotypic markers cannot appear in an offspring unless they are present in either or both of the parents. Applying this premise to the resolution of paternity, it can be concluded that if the putative father lacks an antigen expressed by the child that could not have been inherited from the mother (the obligatory gene), he is excluded from consideration as the biological father of that child. This is referred to as a first-order exclusion and is positive evidence that the accused man is not the father.

If the accused man expresses the obligatory characteristic he cannot be excluded. However, this finding does not offer proof that he is indeed the true father; it merely offers evidence against an exclusion. Nevertheless, based on the findings, a probability of the likelihood of his being the true father can be calculated. It is in this arena that the HLA system demonstrates its great potential.

The calculations used to compute the probability of true paternity are based on the frequency of the obligatory gene or combination of genes in a population of random individuals of the same race as the accused. Because the HLA system is so polymorphic, the frequency of ‘each combination usually is very low. Thus, if an accused does indeed have the obligatory gene or combination of genes, it is a simple probability calculation to show how likely he is to be the biological father.

In practice, the HLA system is used in conjunction with a number of red cell blood group systems, including ABO, Rh,

MNSs, Kell, Duffy, and Kidd. The cumulative probability is computed and used to assess the likelihood of paternity. Using all of these systems results in a greater than 95% exclusion of falsely accused males.

Currently, the use of red-cell systems and HLA has been supplanted in most jurisdictions by the use of DNA markers. The markers used are found in numerous loci in the genome and have been found to be extremely polymorphic in various populations. Their frequencies have been studied extensively and allow for exclusion rates of up to 99.9%. Under such conditions paternity can be almost proved.

The HLA system has many applications in transplantation, in platelet transfusion, in the study of the immunogenetics of disease and in forensic matters such as parentage testing. Nevertheless, its most important biological function is probably in the orchestration of the immune response through self-recognition. As such, it is the only polymorphic alloantigenic system in humans whose function is known.

REFERENCES

1. Starzl T, Marchioro TL, Holmes JH, et al. Renal homografts in patients with major donor-recipient blood group incompatibilities. *Surgery* 1964;55:195–200.

2. United Network for Organ Sharing UNOS. 1998 Annual Report, The U.S. Scientific Registry of Transplant Recipients and The Organ Procurement and Transplantation Network. Published by: Department of Health and Human Services, Health Resources and Services Administration. 1998:119–124.

3. Cho YW, Terasaki PI. Long-term survival. In: Terasaki PI, ed. *Clinical Transplants 1988*. Los Angeles: UCLA Tissue Typing Laboratory, 1989:277–282.

4. Busson M, Raffoux C, Bouteiller AM, et al. Influence of HLA-A, B and DR matching on the outcome of kidney transplant survival in preimmunized patients. *Transplantation* 1984;38:227–230.

5. Mickey R, Cho Y, Carnahan E. Long term graft survival. In: Terasaki PI, ed. *Clinical Transplants 1990*. Los Angeles: UCLA Tissue Typing Laboratory, 1991:385–395.

6. Cicciarelli J, Corcoran S. An update on HLA matching, including HLA "epitope" matching: a new approach. In: Terasaki PI, ed. *Clinical Transplants 1988*. Los Angeles: UCLA Tissue Typing Laboratory, 1989:329–337.

7. Konoeda Y, Terasaki PI, Wakisaka A, et al. Public determinants of HLA indicated by pregnancy antibodies. *Transplantation* 1986;41:253–259.

8. Park MS, Barbett AA, Geer LI, et al. HLA epitopes detected by serology. In: Terasaki PI, ed. *Clinical Transplants 1990*. Los Angeles: UCLA Tissue Typing Laboratory, 1991:515–531.

9. Takemoto S, Terasaki P. HLA peptide matching. In: Terasaki PI, ed. *Clinical Transplants 1990*. Los Angeles: UCLA Tissue Typing Laboratory, 1991:497–513.

10. Keogh A, Kaan A, Doran T, et al. HLA mismatching and outcome in heart, heart-lung, and single lung transplantation. *Journal of Heart & Lung Transplantation* 1995;14:444–451.

11. Hosenpud JD, Edwards EB, Lin HM, et al. Joint Registry of the United Network of Organ Sharing, Richmond, VA, USA. Influence of HLA matching on thoracic transplant outcomes. An analysis from the UNOS/ISHLT Thoracic Registry. *Circulation* 1996;94:170–174.

12. Marcus BM, Duquesnoy RJ, Gordon RD, et al. Histocompatibility and liver transplant outcome: does HLA exert a dualistic effect. *Transplantation* 1988;46:372–377.

13. Halloran PF, Wadgymar A, Autenreid P. The regulation of expression of major histocompatibility complex products. *Transplantation* 1986;41:413–420.

14. Lobo PL, Spencer C, Douglas MT, et al. The lack of long-term detrimental effects on liver allografts caused by donor-specific anti-HLA antibodies. *Transplantation* 1998;55:1063–1066.

15. Sanfilippo F, Vaughn WK, Spees EK, et al. Cadaver renal transplantation ignoring peak-reactive sera in patients with markedly decreased pretransplant sensitization. *Transplantation* 1984;38:119–124.

16. Scornik JC, Brunson ME, Howard RJ, et al. Alloimmunization, memory, and the interpretation of crossmatch results for renal transplantation. *Transplantation* 1992;35:389s.

17. Tan EM. Drug induced autoimmune disease. *Fed Proc* 1974;33:1984–1987.

18. Kao KJ, Scornik JC, Small JS. Enzyme-linked immunoassay for anti-HLA antibodies: an alternative to panel studies by lymphocytotoxicity. *Transplantation* 1993;55:192.

19. Harmer AW, Heads AJ, Vaughn RW. Detection of HLA class I and class II specific antibodies by flow cytometry and PRA-Stat screening in renal transplant recipients. *Transplantation* 1997;63:1828.

20. Elandt-Johnson RC. Test for association and the sample correlation coefficient for a 2×2 contingency table. In: *Probability Model and Statistical Methods in Genetics*. New York: John Wiley & Sons, 1971:362–365.

21. Goeken NE. Outcome of renal transplantation following a positive crossmatch with historic sera: the second analysis of the ASHI survey. *Transplant Proc* 1985;17:2443–2450.

22. Noreen HJ. Crossmatch tests. In: Zachary AA, Teresi GA, eds. *ASHI Laboratory Manual*. 2nd ed. Lenexa, KS: American Society for Histocompatibility and Immunogenetics, 1990:307–320.

23. Scornik JC, Bray RA, Pollack MS, et al. Multicenter evaluation of the flow cytometry T cell crossmatch. Results from the ASHI-CAP proficiency testing program. *Transplantation* 1997;63:1440.

24. Bray RA. Flow cytometry crossmatching for solid organ transplantation. *Methods in Cell Biology* 1994;41:103–119.

25. Claas FHJ, van Rood JJ. Soluble HLA antigens and their peptides as immuneregulators and monitoring devices. *Human Immunology* 1999;60:398–400.

26. Zavazava N, Kronke M. Soluble HLA class I molecules induce apoptosis in alloreactive cytotoxic T lymphocytes. *Nature Medicine* 1996;2:1005–1010.

27. Borelli SD, Ferreira E, Oliveira AM, et al. Specific sHLA in healthy donors and donor-specific sHLA in renal transplant patients. *Human Immunology* 1999;60:430–434.

28. Yankee RA, Grumet FC, Rogentine GN. Platelet transfusion therapy. The selection of compatible platelet donors for refractory patients by lymphoyte HLA typing. *N Engl J Med* 1969;281:1208–1212.

29. Dahlke MB, Weiss KL. Platelet transfusion from donors mismatched for crossreactive HLA antigens. *Transfusion* 1984;24:299–302.

30. Bryant P, Vayntrub T, Schrandt H, et al. HLA Antibody enhancement by double addition of serum: use in platelet donor selection. *Transfusion* 1992;32:839–844.

31. Zmijewski CM. HLA and disease. *CRC Crit Rev Clin Lab Sci* 1984;20:285–370.

32. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet* 1955;19:251–253.

33. Haldane JBS. The estimation and significance of the logarithm of a ratio of frequencies. *Ann Hum Genet* 1956:20:309–311.

34. Thorsby E. Invited Anniversary Review: HLA associated diseases. *Human Immunology* 1997;53:1–11.

35. Monos DS, Mickelson E, Hansen JA, et al. Analysis of DR and DQ gene products of the DR4 haplotype in patients with IDDM: possible involvement of more than one locus. *Hum Immunol* 1988;23:289–299.

36. Todd JA, Bell JI, McDevitt HO. HLA DQ[fcbeta] gene contributes to susceptibility and resistance to insulin dependent diabetes mellitus. *Nature* 1987;329:599–604.

37. Morel PA, Dorman JS, Todd JA, et al. Aspartic acid at position 57 of the HLA DQB chain protects against type 1 diabetes: a final study. *Proc Natl Acad Sci USA* 1988;85:8111–8115.

38. Bias WB, Zachary AA. Genetic and statistical principles of paternity determination. In: Rose NR, DeMarcario EC, Rosner G, et al., eds. *Manual of Clinical Laboratory Immunology*. 4th ed. Washington, DC: American Society for Microbiology, 1986:901–912.

38
Bone Marrow Transplantation

Patrick G. Beatty

BACKGROUND

Part of "38 - Bone Marrow Transplantation"

Goals of Marrow Transplantation

For over 50 years, radiation and/or chemotherapy has been used in the treatment of malignant diseases (1, 2). Although these modalities can induce complete remissions, these remissions often were of short duration. Efforts to escalate drug and/or radiation doses in hopes of effecting a cure were frustrated by organ toxicity, particularly bone marrow damage. Despite these limitations, careful clinical studies led to the development of regimens capable of curing a substantial proportion of patients with Hodgkin's disease (3) and childhood acute lymphoblastic leukemia (4). However, for most patients, cure was impossible, as treatment doses capable of destroying the malignancy also would destroy marrow function.

Experiments carried out in secret during World War II, and openly afterwards, pointed towards a possible solution for the dose limiting toxicity of chemoradiotherapy. It was shown that lethally irradiated mice could be rescued with either spleen cells or marrow cells from a syngeneic animal (5, 6, 7). This observation led to the concept that patients with cancer might be cured of their malignancy by delivering very high-dose therapy, then rescuing them from the resulting hematologic toxicity by giving them normal bone marrow hematopoietic stem cells collected from another individual. Clinical experiments were carried out in desperately ill patients in the mid-1950s, delivering high dose radiation followed by intravenous infusion of marrow from another person (8). Although there were some cases of transient engraftment, there were no long-term survivors: all patients died either of transplant-related complications or relapsed malignancy (9). However, these endeavors did demonstrate that human marrow could be obtained and infused without ill effects to either donor or recipient. More successful were two transplants carried out in two patients with leukemia, each of whom had an identical twin (10). Each patient received total body irradiation, and then a marrow infusion from his/her identical twin. Both twins apparently engrafted, but unfortunately, both quickly relapsed with their leukemia.

It was clear from these early human experiments that marrow transplantation was more difficult than initially imagined. Indeed, there were no further clinical trials until the beginning of the appreciation of the importance of the human histocompatibility complex, HLA, and the evolving ability to test in a rudimentary fashion for the antigens encoded by the complex. Thus, in 1968, three children with congenital immunodeficiency diseases were transplanted from HLA-matched siblings (11, 12, 13). All three patients are currently alive and well nearly 30 years later (14). Shortly thereafter, Dr. E. Donnall Thomas treated a series of patients with advanced leukemia, using high-dose chemoradiotherapy followed by marrow rescue from HLA-identical siblings (15). Although there was a high death rate from transplant-related complications and relapse, a small percentage of these otherwise incurable patients was rendered disease-free, and remain so more than 20 years later. These encouraging early successes led to the systematic development of the clinical art and science of marrow transplantation, with gradual improvements taking place in the control of infections, the understanding and control of graft-versus-host-disease (GVHD), the delineation of optimal treatment protocols, the recognition of the importance of performing the transplant early in a patient's disease course, and perhaps most important, the development of an increasing ability to define histocompatibility differences between patient and donor (16, 17, 18, 19, 20, 21). These developments led to the gradual acceptance of marrow transplantation as the cost-effective treatment of choice for many otherwise fatal hematologic diseases (22).

Indications for Marrow Transplantation

There are four general indications for marrow transplantation: to replace a diseased myeloid and/or lymphoid system; to provide rescue from high dose chemoradiotherapy given for treatment of a malignancy; to deliver anticancer immunotherapy; and to effect therapy for defective genes.

The replacement of a diseased or damaged organ is analogous to solid organ transplantation. Indeed, replacement of a genetically defective immune system was the first successful indication for transplantation; and transplantation still remains the most effective therapy for a wide range of these disorders, including severe combined immune deficiency, Wiskott-Aldrich syndrome, and any other genetic or acquired disease that causes potentially fatal damage to either lymphoid or myeloid cells (23). In addition to the genetic syndromes, there is a spectrum of acquired diseases that may or may not have a defined etiology, such as chloramphenicol-induced severe aplastic anemia, accidental radiation damage after incidents such as the Chernobyl disaster, or

idiosyncratic reactions to a wide range of medications (24). A substantial proportion of acquired myeloid or lymphoid damage appears to be immunologically based, which occasionally can be cured or ameliorated with immunosuppressive therapy, without need for complete replacement by marrow transplantation (25). The majority of patients with nongenetic marrow/lymphoid failure appear to have no inciting cause. Whatever the etiology, all these defects can be approached by administering cytotoxic agents, then delivering healthy bone marrow from another individual. The results of such transplants generally are excellent. For instance, patients who suffer from newly diagnosed severe aplastic anemia, and who receive immediate transplant from an HLA-identical sibling, can expect long-term, disease-free survival rates above 90% (26).

The second and most common indication for transplantation is for treatment of chemotherapy-sensitive malignant disease. The approach to a patient who is known to have a chemoradiosensitive malignancy is to deliver potentially curative cytotoxic therapy at a level also known to be marrow ablative. These patients then are rescued with marrow from an appropriate donor, be it an HLA-matched sibling, an HLA-mismatched relative, or an HLA-matched or mismatched unrelated donor. The first disease successfully approached in this manner was acute leukemia. Patients with either acute lymphocytic or acute myelocytic leukemia were given high dose chemoradiotherapy, then rescued with marrow from an HLA-genotypically matched sibling. Although the early death rates were extremely high, approximately 10% of these patients became long-term, disease-free survivors (15). As understanding evolved concerning transplant complications, the death rate dropped to the point where it was possible to consider transplantation of patients who had fatal diagnoses, but who were still in remission. This led to a dramatic increase in the survival rates for patients with diagnoses such as acute myelocytic leukemia in first remission and chronic myelocytic leukemia in chronic phase (27, 28). Other malignant diseases that now are routinely approached with allogeneic marrow transplant include non-Hodgkin's lymphoma, Hodgkin's disease, multiple myeloma, myelodysplastic syndrome, and indeed any malignant disease of the hematopoietic system which is known to be therapy responsive, but is not usually curable with conventional dosing (29, 30, 31, 32).

The third indication for marrow transplantation, anti-cancer immunotherapy, only became evident after years of experience. As the marrow inoculum contains a complete immune system, it can potentially effect a potent immunotherapeutic effect on residual malignant cells. Although the mechanism has not been worked out, it has long been noted that there is a correlation between an increased incidence of GVHD and a decreased risk of leukemia relapse (33, 34). When efforts were aimed at decreasing the incidence of GVHD, such as depleting T-cells from the marrow inoculum, there was often an increased risk of malignancy relapse (35). Conversely, there have been numerous attempts to harness this graft-versus-leukemia effect by attenuating or even eliminating the standard post transplant GVHD prophylaxis (36). Generally these efforts have been fruitless: although there may be a decrease in the relapse rate, in most studies this was counteracted by an increased death rate from acute or chronic GVHD. More recently, clinical researchers are studying the effects of T-lymphocytes given several weeks after transplant in hopes of separating in time the graft-versus-leukemia effect from the GVHD effect (37). One hypothesis is that T-cells given at the time of transplant are more likely to cause GVHD because of the inflammatory cytokine milieu secondary to the chemoradiotherapy conditioning regimen.

The fourth indication for marrow transplantation is to administer gene therapy. As described for the first indication, bone-marrow transplantation can be curative for genetic defects of the hematopoietic system: the mechanism is simply the obliteration of the defective gene as it exists within the hematopoietic cell, and substitution of a normal gene, carried by a normal hematopoietic cell. It has been shown that diseases that do not at first glance appear to be related to the hematopoietic system can be cured by this approach (38, 39). For instance, the mucopolysaccharide deficiency diseases often involve macrophage function defects, which allow buildup of toxic products in the liver, brain, or other organs (40). These can be cured by introducing macrophages with a normal gene, as part of replacing the entire hematopoietic system with a marrow transplant. In the future, it may be possible to directly manipulate the diseased gene within a hematopoietic stem cell, and/or substitute a normal gene (41).

The Allogeneic Response in Marrow Transplantation

The major allogeneic marrow transplantation complication, which requires such detailed and precise histocompatibility testing, is that of GVHD (42). Whereas with other types of organ transplantation, graft rejection is the major issue in marrow transplantation, where an intact immune system is transplanted into a patient's body, it is necessary to consider immune reactivity in two vectors: graft-versus-host and host-versus-graft. In most patients who receive a marrow transplant, particularly those being treated for an underlying malignancy, treatment of that malignancy is so immunosuppressive that the chance of the patient's immune system surviving with sufficient strength to reject a marrow graft is minimal. However, and perhaps partly because of this extraordinarily high dose of chemoradiotherapy, there is a very high probability that the incoming immune system will recognize the antigens in the patient's body and cause severe or even fatal reactions. Although skin, liver, and gut are the most commonly afflicted organs, any organ of the body can be damaged. The immunogenetic nature of this reaction is most clearly demonstrated by the observation that GVHD is far less commonly seen when genotypically identical twins are used as donors (43). Further, the incidence of GVHD increases with increasing HLA disparity (44). However, although generalizations can be made about the relative risks given certain levels of disparity, GVHD risk for individual patients cannot be reliably predicted. Indeed, patients with no HLA disparities can have fatal GVHD, and patient/donor pairs with one or more major HLA differences can have none. Additional non-HLA factors that may be implicated in GVHD risk include concurrent infections (45); the recent delivery of high-dose chemoradiotherapy (with its “cytokine storm”); and the influence of non-HLA antigens, namely the minor histocompatibility antigens (46, 47, 48).

With the recognition of GVHD as the major barrier to wider use of bone marrow transplantation, major efforts have been made to ameliorate the problem. However, attempts at this have led to other, sometimes unexpected, complications. As mentioned above, it has been demonstrated that patients who survive GVHD are less likely to have a relapse of their underlying malignancy, the so-called graft-versus-leukemia effect. Another unexpected problem that became apparent when efforts were made to decrease GVHD is an increased risk of graft rejection (49, 50). Although the mechanism is not entirely clear, it appears that some GVHD may be required to obliterate the remnants of the host immune system, and hence render it incapable of rejecting the incoming marrow inoculum.

Although GVHD affecting the skin, liver, and gut is the most visible manifestation of the syndrome, the most common cause of death in patients with GVHD is infection (45). It appears that the phenomenon of GVHD disregulates the immune system, to the extent that it is incompetent to recognize pathogens. Furthermore, as the treatment of GVHD involves nonspecific immunosuppression, the patient's vulnerability increases.

Even if the patient and donor are genotypically matched for HLA, there remains a substantial risk of GVHD. The antigens that elicit GVHD in HLA-matched siblings are by definition “minor histocompatibility antigens.” Characterization of these antigens is just beginning (51, 52). It is unclear, particularly in humans, how many antigens there are, where they are encoded in the genome, and whether the same loci are important for each donor/patient pair (53, 54, 55, 56).

HLA-MATCHED SIBLINGS

Part of "38 - Bone Marrow Transplantation"

Histocompatibility Assessment

In determining the most appropriate histocompatibility assessment for a patient who is a candidate for marrow transplantation, it is necessary to consider the availability and relevance of three different categories of donors, usually in the following order of preference: HLA genotypically identical siblings, HLA mismatched relatives, and matched or mismatched unrelated donors.

In the late 1960's, the art of histocompatibility testing had advanced to the point where it usually was possible to define the segregation of HLA alleles within a family (57). As it was then just becoming appreciated, all of the elements of the major histocompatibility complex in humans (HLA) are encoded on a short stretch on the short arm of chromosome 6, and hence are inherited *en bloc*. Thus, with a relatively limited and primitive set of serologic reagents, plus cellular methodology, it was possible to determine which if any siblings had inherited the same two HLA haplotypes from their parents.

The goal of histocompatibility assessment when screening for an HLA-matched sibling donor is straightforward: identify which, if any, siblings have inherited the same HLA haplotypes from their parents. This requires typing all siblings, and if possible, both parents for HLA-A, -B, and -DR. Currently, it is most cost effective to type for HLA-A and -B by serologic methodology, and for HLA-DR by molecular methodology. As discussed below, the use of a mixed lymphocyte culture or other cellular-based technology has been made unnecessary by the precision of the new molecular methods.

Early in the development of histocompatibility testing, it became obvious that if cells from two siblings that are known to be HLA-identical are mixed, the cells would remain quiescent and would not divide (58). However, if there was a disparity between siblings for HLA-DR, there was a rapid development of blasts and cell division, which could be monitored easily by measuring tritiated thymidine uptake. This proved to be a useful test in the early days of HLA typing, when it often was difficult to clearly establish haplotype segregation. Thus, if a patient and a prospective donor were clearly negative in a mixed lymphocyte culture, but there was strong reactivity elsewhere in the family, the patient and donor could be presumed to be HLA genotypically matched. However, with the increasing precision of DNA typing, it is now possible to demonstrate with virtual certainty how haplotypes segregate within a family, thus making the MLC test obsolete for that purpose. Other cellular crossmatching tests include measurements of precursor frequency by a limiting dilution analysis (59). There is some indication that these tests might be capable of picking up minor histocompatibility antigen differences, but this has not been definitively demonstrated. Other more complicated tests include mixing donor lymphocytes with irradiated patient T-lymphocytes, then taking the subsequent cells and layering them upon patient fibroblast lines, and demonstrating fibroblast destruction (60). This does appear to be somewhat predictive of minor antigen disparity in HLA-matched siblings, but the cumbersome nature of the test, and the time involved for developing results, have led to limited acceptance of this method.

Minor Histocompatibility Antigens

Once HLA haplotype segregation has been established in a family, and it has been demonstrated that the patient and donor are genotypically HLA matched, it can be concluded that to any given level of discrimination, they are matched for all known HLA alleles and indeed for any antigens that have been encoded in that region but that have not yet been described. Thus, the risk of GVHD must be ascribed entirely, by definition, to minor histocompatibility antigens (61, 62). Although the existence of segregating non-HLA antigen systems can be well demonstrated in animal models, precise definition of the antigens has proven elusive (46, 47 and 48, 51, 52, 53, 54, 55 and 56). Because of the peculiar biology of these antigens, it has been difficult to develop reliable testing. It appears that the minor antigen is expressed on the cell surface as a degraded peptide fragment bound to a specific HLA molecule. These peptides can be readily recognized only by T-lymphocytes, which coordinately recognize both peptide and adjacent HLA amino-acid residues. Thus, it has proven virtually impossible to develop antibodies capable of recognizing these antigens. Recently, sophisticated biochemical methodology involving removing the peptides from the HLA binding groove has allowed sequencing of some of these peptides (51, 56). It appears they are fragments of various proteins that do not necessarily have any connection to the immune system. Another approach to the identification of minor histocompatibility antigens is by using genetic linkage analysis. Gubarev et al. (63, 64) describe localization

of two distinct loci, which encode minor histocompatibility antigens. T-lymphocyte clones were isolated from a patient with GVHD, and each tested against lymphoblastoid cell lines derived from numbers of large, genetically mapped pedigrees, each of which had been transfected with the appropriate HLA-restricting element (in this case HLA-B7). Each family member was phenotyped for presence or absence of the minor antigen by a cytolytic T-cell lysis assay, and a computer linkage analysis performed. Each clone localized a gene to within 30 to 60 centimorgans on chromosomes 11 and 22.

It will take considerably more effort to fully catalog these antigens, identify the loci that encode them (65, 66, 67 and 68), and determine the number of alleles per locus. Only when this has been accomplished will it be possible to first develop clinical testing methodology, and then ultimately determine how many loci exist, their degree of polymorphism, and which loci might be most important.

HLA MISMATCHED RELATIVES

Part of "38 - Bone Marrow Transplantation"

Histocompatibility Assessment

Given the genetics of the HLA system, specifically a 50% chance that two siblings have inherited the same haplotype from each parent, the total probability that any two siblings are HLA identical is 25%. Given the current average family size in the United States of 2.7 siblings, one of whom is the patient, there are 1.7 siblings to choose from. Therefore, the approximate probability that the average U.S. patient will have an HLA match within the family is 30% to 35%. The remainder of the patients who otherwise might have been considered for transplant must look beyond their siblings. For those patients who do not have an HLA-matched sibling within the family, it is reasonable to consider transplantation from a family member who shares one haplotype and is variably mismatched on the nonshared haplotype. The clinical data clearly indicate that with increasing disparity for HLA, whether it is the A, B, or DRB1 locus, there is increased risk of GVHD (44). The first goal of histocompatibility testing in such cases is to determine that patient and donor share one HLA haplotype by inheritance. With respect to the two nonshared haplotypes, the problem of positively identifying each antigen at each locus is obviously significantly more complex than that of simply identifying haplotype segregation. Thus, it becomes critically important to have reagents that are of sufficient quality to unambiguously define each antigen, so as to be certain of whether the patient is either matched or mismatched for that locus with the prospective donor. Ability to carry out this kind of testing evolved several years after the initial ability to define haplotypes; hence, there was a delay in testing the hypothesis that transplants from HLA mismatched relatives can be successful.

Thus, the histocompatibility assessment when considering HLA-mismatched relatives is considerably more complex than that involved when only HLA-matched siblings are considered. First, one must establish HLA haplotype segregation within the family: this requires testing for HLA-A, -B, and -DR. Once it is clear which family members share by inheritance one HLA haplotype, attention can be turned to the nonshared haplotypes. As these haplotypes, by definition, are not shared by inheritance, it is important to define with the greatest possible precision all antigens at each relevant locus (see below). Currently, this would include testing for HLA-DR and -DQ using molecular methodology; testing for HLA-C preferably by molecular methods; and testing for HLA-A, and -B currently by serology, but hopefully in the near future by molecular methods. As some HLA disparity can be presumed, it is critical to test by the lymphocytotoxicity assay for presence of patient-antidonor anti-HLA-specific antibodies. Although it is not clear whether such antibodies directly contribute to graft rejection, it is clear that a correlation exists (69). Perhaps these antibodies are surrogates for a more generalized increased alloreactive state.

Relative Impact of HLA Disparities

Determination of the relative importance of particular HLA loci to GVHD can be most readily accomplished in large cohorts of patients, treated in a similar fashion, in whom there is only one locus different between patient and donor. An early analysis carried out in Seattle showed no substantial difference in the risk of GVHD, whether the mismatched locus was HLA-A, -B, or the serologically defined -DR locus (44). More recently, using more precise molecular typing technology, Servida et al. from Seattle (70) have reported that disparity for HLA-DRB1 incurs a significantly greater risk for GVHD than disparity for HLA-A or HLA-B. There was a trend for lower GVHD risk for HLA-A disparity as opposed to HLA-B disparity.

It is possible that different modes of patient immunosuppression posttransplant might have an impact upon the relative hierarchy of risk of various loci. The current standard of care will accept a donor who differs for one of three loci: HLA-A, -B, or -DR. Although in such instances there is an increased risk of GVHD, there is not necessarily a substantial impact upon survival, mostly because of an increased graft-versus-leukemia effect. The relevance of HLA-C, -DQ, and -DP remains unclear in this setting. The use of two or three loci-disparate donors is controversial: because there is clearly an increased risk of GVHD, use of such donors therefore should be confined to clinical research trials attempting to ameliorate the inherent allogeneic risks (71, 72, 73).

HLA MATCHED/MISMATCHED UNRELATED DONORS

Part of "38 - Bone Marrow Transplantation"

Histocompatibility Assessment

When considering unrelated donors, the difficulty of positively identifying each antigen increases. Here we assume that there is no sharing of HLA haplotypes by inheritance. Thus, it is necessary to positively identify both alleles at each relevant HLA locus. Furthermore, it is not enough to simply determine match versus mismatch; one must also be certain of the exact allele designation as the patient has literally millions of possible donors to choose from (74), and it is clinically important to optimize this choice. For instance, there may be some mismatches that are less dangerous than others.

It quickly became apparent in the mid-1980s that the serologic

and cellular testing technologies were insufficient for optimal unrelated donor/patient matching. This has led to the rapid development of DNA-based allele-level typing, initially for class II (HLA-DR, -DQ, -DP) (75), and more recently for class I (HLA-A, -B, -C) (76, 77). It has become clear that DNA-based typing is far more accurate and reproducible than were previous methodologies (78, 18). It is likely only a matter of time before the cost of DNA typing lowers to the point at which this testing methodology completely supplants other technologies.

In considering transplantation using unrelated donors, it must be recognized that as opposed to the situation with related donors, the degree of match can only be presumed to extend to the level of resolution of the typing technology employed, and the number of loci tested. For instance, in considering HLA-matched siblings, if the patient and donor are both HLA-B12, and if these antigens can be demonstrated to be on a shared haplotype, they can be presumed to match; thus, it is immaterial whether the antigen can be split to demonstrate whether it is HLA-B44 or HLA-B45. In the case of an unrelated patient/donor pair where, of course, there are no haplotypes shared by inheritance, all such antigens need to be split to the limits of current technology, as they could be mismatched for these alleles. Thus it is necessary to carry out complete testing of all clinically relevant loci. Particularly as there now have been several additional loci demonstrated to be polymorphic in humans beyond HLA-A, -B, -C, and -DR, and particularly as the technology of HLA typing has advanced to the point that there are now dozens or even over a hundred alleles identified at each locus, it has become important to attempt to determine a hierarchy of importance of loci in the human allogeneic response. Addressing these questions requires a very large body of clinical data with precision typing. Only now are sufficient numbers of patients available to perform this expensive analysis (79, 80). Further, for most complete assurance for validity of this type of analysis, it is best to study patient/donor pairs in which there is only one locus difference. It appears safe from the available clinical data to date to conclude that the HLA-A, -B, and -DRB1 loci are of importance. There now are data that demonstrate that the HLA-DP locus has little, if any, significance in clinical GVHD (81). Disparity for HLA-DQ, however, appears to correlate with risk for AGVHD (82). The HLA-C locus is in the process of virtually complete redefinition using DNA typing (77, 83). Because of methodologic problems with serology related to the relatively low expression of HLA-C locus antigens, and their extensive cross reactivity with HLA-B locus alleles, it has always proven difficult to obtain reliable typing. With the availability of DNA-based methodology, a number of new alleles have been identified, and many of the old antigens have been split or redefined. The hypothesis has been raised that the HLA-C locus may have some importance in natural killer cell function, and perhaps some importance in graft rejection (84, 85).

In unrelated donor transplantation, the use of DNA-based technology for class II typing has virtually eliminated the use of serology because of its extreme accuracy and precision. Indeed, comparisons between DNA and serology in reliable laboratories have shown a very substantial error rate for HLA-DR serology (78, 86, 87, 88, 89 and 90). As most class I alleles now have been sequenced, it can be expected that DNA-based typing for the HLA-A, -B, and -C loci soon will move from the research and development phase into widespread application. It then will be instructive to reexamine the issue of the impact of minor HLA polymorphisms upon clinical outcome in light of using these far more precise tools (92, 93, 94 and 95).

The precise DNA-based technology to be utilized is still not resolved (75, 96). Initially, analysis of restriction-fragment-length-polymorphisms (RFLPs) was used; but this technology proved to be often insufficiently informative, expensive, and cumbersome. The invention of the polymerase chain reaction (PCR) led to the development of a number of related technologies exploiting differences or similarities at the DNA base-pair level. Which of these technologies will ultimately be selected for use in clinical practice will likely depend upon a compromise between cost, precision, and accuracy.

Cellular crossmatching technology initially was developed to help identify HLA-matched siblings; however, it has proven to be particularly unreliable in unrelated transplantation (97, 98 and 99). The explanation likely revolves around the fundamental issue in unrelated transplant, namely that the patient and donor can be expected to have only phenotypic, but not genotypic, identity. Thus, when cells from two HLA genotypically identical siblings are mixed in culture (MLC), one can expect virtually zero reactivity. However, even if a patient and an unrelated donor are fully matched for HLA-A, -B, and -DR, it is likely that there will be some disparities, which would be detectable in an MLC test. This low level of activity is complicated further by the fact that patient cells cannot be expected to be normally reactive, as a result of patient's underlying disease or the treatment thereof. When results of MLC testing have been analyzed carefully in unrelated transplantation, the assay has been found to be both poorly reproducible, and most significantly, virtually completely nonpredictive for development of GVHD. On the other hand, it is clear that high resolution HLA-DRB1 testing is very predictive of GVHD (100). As a result, the use of the MLC test has been virtually abandoned in this setting.

Impact of HLA Polymorphism on Finding Donors

Given the extreme polymorphism of HLA, and the large number of different loci (Table 38.1), it might seem foolhardy to hope that any given patient might be able to find a fully matched unrelated donor (101, 102 and 103). This problem has been operatively overcome by simplifying assumptions, and by considering the biology of the HLA system. It was presumed a decade ago that the most relevant loci would be HLA -A, -B, and -DRB1, although it is yet to be proven that this assumption is entirely valid. Indeed, given the increased risk of GVHD in transplantation using unrelated donors, it is quite possible that this assumption may be proven partially wrong. However, even assuming that these three loci encode the only antigens of importance, there are still an enormous number of possible HLA phenotypes. The biology of HLA, however, makes the possibility of finding an HLA match more manageable. As previously noted, HLA antigens are encoded on one small portion of the short arm of chromosome 6. As crossovers are infrequent, one can presume that HLA antigens are evolutionarily selected as a complete haplotype,

are inherited as a haplotype, and are preserved as a haplotype, a phenomenon termed “linkage disequilibrium.” Within different racial groups, certain haplotypes are far more common than would be expected by chance alone. The reason for this is unclear, but it might be postulated that this phenomenon is related to the evolutionary selection of individuals with haplotypes that protect against specific endemic diseases in their region of the world. Thus, for instance, among whites, the HLA-A1, -B8,-DR3 haplotype is extremely common. Perhaps an individual who possessed that haplotype would have had a high chance of surviving the Plague or smallpox in the Middle Ages. The relevance for our purpose is that a person with such a haplotype has a high chance of finding a donor even from amongst a relatively small file of volunteers. However, it appears that the majority of patients will have haplotypes that are neither particularly common nor in linkage disequilibrium. Thus, in order to have a reasonable probability of finding a well-matched donor, it is necessary to have available a very large file of HLA-typed potential donors. In the United States there are over 2.5 million donors available, with another 1 million available worldwide (104). Currently, approximately 65% of patients can find an acceptable donor (105). However, it is clear that patients belonging to racial groups that are not well represented in the registries have a dramatically decreased probability of finding a donor. This is not only because of the different distribution of HLA antigens and haplotypes within different racial groups: for African-Americans there appears to be an additional problem, namely that those peoples descended from those currently living in Africa appear to be far more polymorphic for HLA than are other racial groups (105). A hypothesis for this is the origin of humanity in Africa, with other racial groups who have migrated out of Africa passing through a genetic “bottleneck,” and thus having limited diversity. In any case, the implication is that in order to find HLA-matched donors for African-Americans, it will be necessary to recruit far more donors than are needed for other racial groups. Obviously, for these racial groups in particular, the development of technology to allow successful HLA-mismatched transplants from unrelated donors will be particularly important.

TABLE 38.1. HLA POLYMORPHISM (106)

Based on Number of Alleles Shown at Each Locus						
A	*B*	*C*	*E*	*G*		
142	285	99	5	14		
DRA	*DRB1*	*DRB3*	*DRB4*	*DRB5*	*DRB7*	
2	225	20	9	13	2	
DQA1	*DQB1*	*DPA1*	*DPB1*	*DMA*	*DMB*	*DOA*
20	42	17	81	4	65	8
Considering All Loci						
No. of possible haplotypes: 1.02×10^{23}				No. of possible genotypes: 1.04×10^{46}		
Only Considering HLA-A,B,DRA,DRB1						
No. of possible haplotypes: 1.82×10^{7}				No. of possible genotypes: 3.32×10^{14}		

REFERENCES

1. Goodman LS, Wintrobe MM, Dameshek W, et al. Nitrogen mustard therapy for Hodgkin's disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders. *JAMA* 1946;132:26–132.

2. Pusey WA. Cases of sarcoma and of Hodgkin's disease treated by exposures to X-rays: a preliminary report. *JAMA* 1902;38:166–170.

3. DeVita VT Jr, Simon RM, Hubbard SM, et al. Curability of advanced Hodgkin's disease with chemotherapy: long-term follow up of MOPP treated patients at the National Cancer Institute. *Ann Intern Med* 1980;92:587–595.

4. Mauer AM. Therapy of acute lymphoblastic leukemia in childhood. *Blood* 1980;56:1–10.

5. Jacobson LO, Simmons EL, Marks EK, et al. The role of the spleen in radiation injury and recovery. *J Lab Clin Med* 1950;35:746–770.

6. Lorenz E, Uphoff D, Reid TR, et al. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst* 1951;12:197–201.

7. Rekers PE, Coulter MP, Warren SL. Effect of transplantation of bone marrow into irradiated animals. *Arch of Surg* 1950;60:635–677.

8. Thomas ED, Lochte HL, Lu WC, et al. Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. *N Engl J Med* 1957;257:491–496.

9. Bortin MM. A compendium of reported human bone marrow transplants. *Transplantation* 1970;9:571–587.

10. Thomas ED, Lochte HL, Cannon, JH, et al. Supralethal whole body irradiation and isologous marrow transplantation in man. *J Clin Inves* 1959;38:1709–1716.

11. Bach FH, Albertini RJ, Anderson JL, et al. Bone-marrow transplantation in a patient with the Wiskott-Aldrich syndrome. *Lancet* 1968;ii:1364–1366.

12. de Koning J, van Bekkum DW, Dicke, KA, et al. Transplantation of bone-marrow cells and fetal thymus in an infant with lymphopenic immunological deficiency. *Lancet* 1969;i:1223–1227.

13. Gatti RA, Allen HD, Meuwissen HJ, et al. Immunological reconstitution of sex-linked lymphopenic immunological deficiency. *Lancet* 1968;ii:366–1369.

14. Bortin MM, Bach FH, van Bekkum DW, et al. 25th Anniversary of the first successful allogeneic bone marrow transplants. *Bone Marrow Transplantation* 1994;14:211–212.

15. Thomas ED, Buckner CD, Banaji M, et al. One hundred patients with acute leukemia treated by chemotherapy, total body irradiation, and allogeneic marrow transplantation. *Blood* 1977;49:511–533.

16. Armitage J.O. Bone Marrow Transplantation. *N Engl J Med* 1994;330:827–838.

17. Beatty PG. Bone marrow transplantation for the treatment of hematologic diseases: status in 1994. *Exp Hematol* 1995;23:277–288.

18. Beatty PG. Clinical and societal issues in blood and marrow transplantation for hematologic diseases. *Biology of Blood and Marrow Transplantation* 1995;1:94–114.

19. Beatty PG, Hansen JA. Bone marrow transplantation for the treatment of hematologic diseases: Status in 1992. *Leukemia* 1993;7:1074–1129.

20. Thomas ED, Storb R, Clift RA, et al. Bone marrow transplantation. *N Engl J Med* 1975;292:832–843, 895–902.

21. Goodrich JM, Bowden RA, Fisher L, et al. Gancyclovir prophylaxis to prevent cytomegalovirus disease after allogeneic marrow transplant. *Ann Intern Med* 1993;118:173–178.

22. Welch HG, Larson EB. Cost effectiveness of bone marrow transplantation in acute nonlymphocytic leukemia. *N Engl J Med* 1989;321:807–812.

23. Parkman R. The application of bone marrow transplantation to the treatment of genetic diseases. *Science* 1986;232:1373–1378.

24. Shadduck RK. Aplastic anemia. In: Beutler E, Lichtman MA, Coller BS, et al., eds. *Williams Hematology*, 5th ed. New York: McGraw-Hill, 1995:238–251.

25. Rosenfeld S, Kimball J, Vining D, et al. Intensive immunosuppression with antithymocyte globulin and cyclosporin as treatment for severe aplastic anemia. *Blood* 1995;85:3058–3065.

26. Anasetti C, Doney KC, Storb R, et al. Marrow transplantation for severe aplastic anemia. Long term outcome in fifty “untransfused” patients. *Ann Intern Med* 1986;104:461–466.

27. Thomas ED, Buckner CD, Clift RA, et al. Marrow transplantation for acute nonlymphoblastic leukemia in first remission. *N Engl J Med* 1979;301:597–599.

28. Goldman JM, Apperly JF, Jones L, et al. Bone marrow transplantation for patients with chronic myelogenous leukemia. *N Engl J Med* 1986;314:202–207.

29. Appelbaum FR, Sullivan KM, Buckner CD, et al. Treatment of malignant lymphoma in 100 patients with chemotherapy, total body irradiation, and marrow transplantation. *J Clin Oncol* 1987;5:1340–1347.

30. Phillips GL, Reece DE, Barnett MJ, et al. Allogeneic marrow transplantation for refractory Hodgkin's disease. *J Clin Oncol* 1989;7:1039–1045.

31. Gahrton G, Tura S, Ljungman P, et al. Allogeneic bone marrow transplantation in multiple myeloma. *N Engl J Med* 1991;325:1267–1273.

32. Anderson JE, Appelbaum FR, Fisher LD, et al. Allogeneic bone marrow transplantation for 93 patients with myelodysplastic syndrome. *Blood* 1993;82:677–681.

33. Antin J.H. Graft-versus-leukemia: no longer an epiphenomenon. *Blood* 1993;82:2273–2277.

34. Weiden PL, Sullivan KM, Flournoy N, et al. Antileukemia effect of chronic graft-versus-host disease. *N Engl J Med* 1981;304:1529–1533.

35. Marmont AM, Horowitz MM, Gale RP, et al. T-cell depletion of HLA-identical transplants in leukemia. *Blood* 1991;78:2120–2130.

36. Sullivan KM, Deeg JH, Sanders J, et al. Hyperacute graft vs. host disease in patients not given immunosuppression after allogeneic marrow transplantation. *Blood* 1986;67:1172–1175.

37. Naparstek E, Magler A, Cividalli G, et al. T-cell depleted allogeneic BMT for acute leukaemia using Campath-1 antibodies and post-transplant administration of donor's peripheral blood lymphocytes for prevention of relapse. *Br J Haematol* 1995;89:506–515.

38. Brenner MK, Rill DR, Holladay MS, et al. Gene marking to determine whether autologous marrow infusion restores long-term haemopoiesis in cancer patients. *Lancet* 1993;342:1134–1137.

39. Weinberg KI, Kohn DB. Somatic gene therapy for severe combined immunodeficiency. In: Chang P, ed. *Somatic Gene Therapy*. Boca Raton, FL: CRC Press, 1995.

40. Krivit W, Shapiro EG. BMT for storage diseases. In: Desnick RJ, ed. *Treatment of Genetic Diseases*. New York: Churchill-Livingstone, 1991:203–221.

41. Kohn DB, Weinberg KI, Nolta JA, et al. Engraftment of gene-modified umbilical cord blood cells in neonates with adenosine deaminase deficiency. *Nature Medicine* 1995;1:1017–1023.

42. Ferrara JLM, Deeg HJ. Graft-versus-host-disease. *N Engl J Med* 1991;324:667–674.

43. Rappeport J, Reinherz E, Mihm M, et al. Acute graft versus host disease in recipients of bone marrrow from identical twin donors. *Lancet* 1979;2:717–720.

44. Beatty PG, Clift RA, Mickelsen EM, et al. Marrow transplantation from related donors other than HLA-identical siblings. *N Engl J Med* 1985;313:765–771.

45. Bowden RA. Infections in patients with graft-vs.-host disease. In: Burakoff SJ, Deeg JH, Ferrara J, et al., eds. *Graft-vs.-Host Disease*. New York: Marcel Dekker, 1990:525–538.

46. Beatty PG. Minor histocompatibility antigens. *Exp Hematol* 1993;21:1514–1516.

47. Beatty PG. National Heart, Lung, Blood Institute workshop on the importance of minor histocompatibility antigens in marrow transplantation. *Exper Hematol* 1997;25:548–558.

48. Goulmy E. Minor histocompatibility antigens: from T Cell recognition to peptide identification. *Human Immunol* 1997;54:8–14.

49. Apperley FJ, Jones L, Hale G, et al. Bone marrow transplantation for patients with Chronic Myelogenous Leukemia: T-cell depletion with Campath-1 reduces the incidence of graft-versus-host-disease, but may increase the risk of leukemia relapse. *Bone Marrow Transplantation* 1986;1:53–66.

50. Martin PJ, Hansen JA, Storb R, et al. Human marrow transplantation: an immunological perspective. *Adv Immunology* 1987;40:379–438.

51. den Haan JM, Sherman NE, Blokland E, et al. Identification of a graft versus host disease-associated human minor histocompatibility antigen. *Science* 1995;268:1476–1480.

52. Marjit WAF, Kernan NA, Diaz-Barrientos TH, et al. Multiple minor histocompatibility antigen-specific cytotoxic T-lymphocyte clones can be generated during graft rejection after HLA-identical bone marrow transplantation. *Bone Marrow Transplantation* 1995;16:125–132.

53. de Bueger M, Goulmy E. Human minor histocompatibility antigens. *Transplant Immunology* 1993;1:28–38.

54. Martin PJ. Increased disparity for minor histocompatibility antigens as a potential cause of increased GVHD risk in marrow transplantation from unrelated donors compared with related donors. *Bone Marrow Transplantation* 1991;8:217–223.

55. Perreault C, Decary F, Brichu S. Minor histocompatibility antigens. *Blood* 1990;76:1269–1280.

56. Simpson E, Roopenian D. Minor histocompatibility antigens. *Curr Opin Immunol* 1997;9:655–661.

57. Beatty PG, Mickelsen EM, Petersdorf EW, et al. Histocompatibility 1991. *Transfusion* 1991;31:847–856.

58. Dupont B, Hansen JA. Human mixed-lymphocyte culture reaction: genetics, specificity, and biological implications. *Adv Immunology* 1976;23107–23202.

59. Kaminski G, Hows J, Goldman J, et al. Optimising a limiting dilution culture system for quantitating frequencies of alloreactive cytotoxic T-lymphocyte precursors. *Cell Immunology* 1991;137:88–95.

60. Vogelsang GB, Hess AD, Berkman AW, et al. An in vitro predictive test for graft-versus-host-disease in patients with genotypic HLA-identical bone marrow transplants. *N Engl J Med* 1985;313:645–650.

61. Mutis T, Gillespie G, Schrama E, et al. Tetrameric HLA class I-minor histocompatibility antigen peptide complexes demonstrate minor histocompatibility antigen-specific cytotoxic T-lymphocytes in patients with graft-versus-host disease. *Nature Medicine* 1999;5:839–842.

62. Goulmy E, et al. Mismatches of minor histocompatibility antigens between HLA-identical donors and recipients and the development of graft-versus-host disease after bone marrow transplantation. *N Engl J Med* 1996;334:281–285.

63. Gubarev MI, Jenkin JC, Leppert MF, et al. Localization to chromosome 22 of a gene encoding a human minor histocompatibility antigen. *J Immunol* 1996;157:5448–5454.

64. Gubarev MI, Jenkin JC, Otterrud BE, et al. Enumeration of loci encoding human minor histocompatibility antigens. *Exp Haemat* 1998;26:976–981.

65. den Haan JM, Sherman NE, Blokland E, et al. Identification of a graft-versus-host disease-associated human minor histocompatibility antigen. *Science* 1995;268:1476–1480.

66. Wang W, Meadows LR, den Haan JM, et al. Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein. *Science* 1995;269:1588–1590.

67. den Haan JM, Meadows LM, Wang W, et al. The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. *Science* 1998;279:1054–1057.

68. Wilke M, Pool J, den Haan JM, et al. Genomic identification of the minor histocompatibility antigen HA-1 locus by allele-specific PCR. *Tissue Antigens* 1998;52:312–317.

69. Anasetti C, Amos D, Beatty PG, et al. Effect of HLA compatibility of engraftment of bone marrow transplants in patients with leukemia or lymphoma. N Engl J Med 1989;320:197–204.

70. Servida P, Gooley T, Hansen JA, et al. Improved survival of haploidentical related donor marrow transplants mismatched for HLA-A or B versus HLA-DR. *Blood* 1996;88:10.

71. Aversa F, Tabilio A, Terenzi A, et al. Successful engraftment of T-cell-depleted haploidentical “three-loci” incompatible transplants in leukemia patients by addition of recombinant human granulocyte colony-stimulating factor-moblilized peripheral blood progenitor cells to bone marrow inoculum. *Blood* 1994;84:3948–3955.

72. Henslee-Downey PJ, Anhyankar SH, Parrish RS, et al. Use of partially mismatched related donors extends access to allogeneic marrow transplant. *Blood* 1997;89:3864–3872.

73. Aversa F, Tabilio A, Velardi A, et al. Treatment of high-risk acute leukemia with T-cell-depleted stem cells from related donors with one fully mismatched HLA haplotype. *New England Journal of Medicine* 1998;339:1186–1193.

74. Beatty PG, Kollman C, Howe CWS. Unrelated-donor marrow transplants: The experience of the National Marrow Donor Program. In: Cecka JM and Terasaki PI, eds. *Clinical Transplants 1995*. Los Angeles: UCLA Tissue Typing Laboratory, 1995:271–277.

75. Begovich AB and Erlich HA. HLA typing for bone marrow transplantation. *JAMA* 1995;273:586–591.

76. Petersdorf EW, Setoda T, Smith AG, et al. Analysis of HLA-B*44 alleles encoded on extended HLA haplotypes by direct automated sequencing. *Tissue Antigens* 1994;44:211–216.

77. Petersdorf EW, Stanley JF, Martin PJ, et al. Molecular diversity of the HLA-C locus in unrelated marrow transplantation. *Tissue Antigens* 1994;44:93–99.

78. Hegland J, Ng J, Hurley CK, et al. Pilot project to evaluate quality of large scale HLA-A, -B, and -DR typing of marrow donors. *Hum Immumol* 1995;44(suppl 1):126a.

79. Hurley CK, Ng J, Hegland J, et al. High resolution class II DNA typing of NMDP donor-recipient pairs: A progress report. *Human Immunol* 1996;47:80.

80. Hurley CK, Hegland J, Nj J, et al. High resolution class I DNA typing of NMDP donor-recipient pairs: a progress report. *Human Immunol* 1996;49:38.

81. Petersdorf EW, Smith AG, Mickelsen EM, et al. The role of HLA-DPB1 disparity in the development of acute graft-versus-host disease following unrelated donor marrow transplantation. *Blood* 1993;81:1923–1932.

82. Petersdorf EW, Longton EM, Anasetti C, et al. Definition of HLA-DQ as a transplantation antigen. *Proc Nat Acad Science* 1996;93:15358–15363.

83. Bunce M, Barnardo MCNM, Procter J, et al. High resolution HLA-C typing by PCR-SSP: identification of allelic frequencies and linkage disequilibria in 604 unrelated random UK Caucasoids and a comparison with serology. *Tissue Antigens* 1996;48:680–691.

84. Moretta A, Sivori S, Vitale M, et al. Existence of both inhibitory (p58) and activatory (p50) receptors for HLA-C molecules in human natural killer cells. *J Exp Med* 1995;182:875–884.

85. Petersdorf EW, Longton GM, Anasetti C, et al. Association of HLA-C disparity with graft failure after marrow transplantation from unrelated donors. *Blood* 1997;89:1818–1823.

86. Ng J, Hurley CK, Baxter-Lowe LA, et al. Large-scale oligonucleotide typing for HLA-DRB1/3/4 and HLA-DQB1 is highly accurate, specific, and reliable. *Tissue Antigens* 1993;42:473–479.

87. Ng J, Hurley CK, Carter C, et al. Large-scale DRB and DQB1 oligonucleotide typing for the NMDP registry: progress report from year 2. *Tissue Antigens* 1996;47:21–26.

88. Hurley CK. Acquisition and use of DNA-based HLA typing data in bone marrow registries. *Tissue Antigens* 1997;49:323–328.

89. Sintasath DM, Bei M, Steiner N, et al. Analysis of HLA-A and -B serologic typing of bone marrow registry donors using polymerase chain reaction with sequence-specific oligonucleotide probes and DNA sequencing. *Tissue Antigens* 1997;50:366–371.

90. Sintasath DM, Hegland JD, Ng J, et al. Analysis of serologic discrepancies of NMDP=s pilot project using PCR-SSOP and direct sequencing. *Human Immunol* 1996;49:44A.

91. Sasazuki T, Juji T, Morishima Y, et al. Effects of matching of class I HLA alleles on clinical outcome after transplantation of hematopoietic stem cells from an unrelated donor. Japan Marrow Donor Program. *New England Journal of Medicine* 1998;339:1177–1185.

92. Petersdorf EW, Gooley TA, Anasetti C, et al. Optimizing outcome after unrelated marrow transplantation by comprehensive matching of HLA class I and II alleles in the donor and recipient. *Blood* 1998;92:3515–3520.

93. Middleton PG, Taylor PRA, Jackson G, et al. Cytokine gene polymorphisms associating with severe acute graft-versus-host disease in HLA-identical sibling transplants. *Blood* 1998;92:3943–3948.

94. Zanone-Ramseier R, Gratwohl A, Gmür J, et al. Sequencing of two HLA-A blank alleles: implications in unrelated bone marrow donor matching. *Transplantation* 1999;67:1336–1341.

95. Davies SM, Kollman C, Anasetti C, et al. Engraftment and survival after unrelated donor bone marrow transplantation. *Blood* 2000;96:4096–4102.

96. Williams TM, Baxter-Lowe LA, Salazar M, et al. Comparison of direct DNA sequencing and SSOP methods for high resolution identification of class II alleles. *Human Immunol* 1996;47:113.

97. Mickelsen EM, Bartsch GE, Hansen JA, et al. The MLC assay as a test for HLA-D compatibility between patients and unrelated donors: results of a National Marrow Donor Program study involving multiple centers. *Tissue Antigens* 1993;42:465–472.

98. Mickelsen EM, Guthrie LA, Etzioni R, et al. Role of the mixed lymphocyte culture reaction in marrow donor selection: matching for transplants from related haploidentical donors. *Tissue Antigens* 1994;44:83–92.

99. Mickelsen EM, Longton G, Anasetti C, et al. Evaluation of the mixed lymphocyte culture (MLC) assay as a method for selecting unrelated donors for marrow transplantation. *Tissue Antigens* 1996;47:27–36.

100. Petersdorf EW, Longton G, Anasetti C, et al. Effect of HLA-DRB1 allele mismatching on clinical outcome following unrelated donor marrow. *Blood* 1994;84:393A.

101. Beatty PG, Dahlberg S, Mickelsen EM, et al. Probability of finding HLA-matched unrelated marrow donors. *Transplantation* 1988;45:714–718.

102. Bodmer JG, Marsh SGE, Albert ED, et al. Nomenclature for factors of the HLA system, 1995. Tissue Antigens 1995;46:1–18.

103. Schipper RF, D'Amaro D, Bakker JT, et al. HLA gene and haplotype frequencies in bone marrow donors worldwide registries. *Human Immunol* 1997;52:54–71.

104. Perkins HA, Kollman C, Howe CWS. Unrelated-donor marrow transplant: the experience of the National Marrow Donor Program. In: Terasaki PI and Cecka JM, eds. *Clinical Transplants*. Los Angeles: UCLA Press, 1994:295–301.

105. Beatty PG, Mori M, Milford E. Impact of racial genetic polymorphism upon the probability of finding an HLA-matched donor. *Transplantation* 1995;60:1–6.

106. Mason PM, Parham P. HLA class I region sequences, 1998. *Tissue Antigens* 1998;51:417.

Section 7 Hematology

Hematology - Introduction

Serhan Alkan

Section Chief

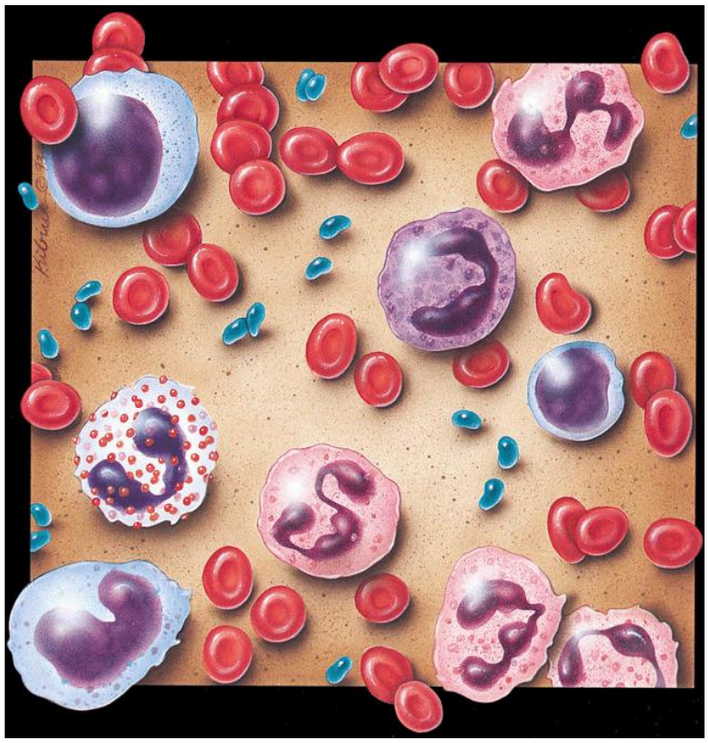


Figure.

39 Hematopoiesis and the Hematopoietic Growth Factors

40 Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders

41 Red Blood Cell Disorders

42 The Thalassemia and Hemoglobinopathy Syndromes

43 Acute Leukemias and Myelodysplastic Syndromes

44 Chronic Lymphoproliferative Disorders, Immunoproliferative Disorders, and Malignant Lymphoma

45 Chronic Myeloproliferative Disorders

Continued advancements in the field of hematopathology, particulary studies based on characterization of genetic abnormalities as well as progress made in automation, have resulted in major changes in clinical diagnosis and management of patients. The following chapters will provide an understanding of the basic approach to diagnosis of various disorders related to hematopoietic disorders. As a discipline of clinical pathology, hematopathology requires expertise in a variety of areas including instrumentation, clinical laboratory analysis, and traditional microscopic skills. These skills actually place hematopathology in the unique position of bridging the clinical laboratory sciences and anatomical pathology. Although the basic foundation of hematopathology remains unchanged, the technological advances that have characterized the field of medicine over the past decade have also changed diagnostic approaches in hematopathology. Indeed, the hematopathologist is now challenged with mastering both the traditional skills of laboratory medicine and morphology and newer diagnostic modalities and instrumentation. The medical community has embraced these new technological advancements and already incorporated some into the routine patient management.

The key to becoming a successful hematopathologist depends on the incorporation of these newer diagnostic modalities as well as expertise gained in traditional morphologic skills. During the past decade, many clinical and laboratory studies have validated the efficacy of utilization of new assays in hematopathology. Based on these studies and their documented successes in patient management, healthcare providers are increasingly incorporating this knowledge in their daily practice. Therefore it becomes crucial for physicians to be aware of both traditional as well as newer modalities available in the laboratory.

39

Hematopoiesis and the Hematopoietic Growth Factors

Stephen G. Emerson

INTRODUCTION

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

All of the cells observed in the adult peripheral blood derive from the bone marrow, arising through an extraordinary process termed hematopoiesis. This process results in the maintenance of the wide variety of distinct blood cells, despite the fact that each blood cell variety has its own distinct lifespan in the circulation. Our current understanding of this process is embodied in a theory called the stem-cell model of hematopoiesis. This model forms the foundation for truly all of our present thinking about normal hematopoiesis, pathologic hematologic disease states, and hematologic therapies.

Stem-Cell Theory

Unlike other tissues of mesodermal origin, which as a rule have little turnover in their constituent cell populations, the cellular components of the blood are constantly undergoing cell death and replacement by new cells. While red blood cells last approximately 4 months in the circulation, platelets last only about 1 week and granulocytes less than 10 hours. It is estimated that every day 1×10^{11} blood cells are lost to wear and tear and are replaced with an equal number of new blood cells. To fulfill the continual need for replacement blood cells, hematopoiesis occurs actively throughout our lifetimes. As a result, the blood-forming tissues are among the most mitotically active, along with the gastrointestinal epithelium and epidermis. The recognition of this high turnover rate for blood cells has led to the development of the stem-cell theory of blood cell development and maintenance, or hematopoiesis (1). This chapter discusses this theory in detail, as well as its implications for clinical practice.

Hematopoietic Growth Factors

One of the most exciting advances in the study of hematopoiesis in the last decade has been the study of the role of hematopoietic hormones in the control of the process of blood-cell differentiation. These hematopoietic hormones, also known as hematopoietic growth factors, appear to control every step in the process of the development of new blood cells. At the present time, we know principally a group of such hematopoietic growth factors that are stimulatory to hematopoietic stem cells and their progeny. However, some inhibitory factors have been discovered and appear to play an equally important role in the negative control of this process. Other hormones that are felt to have primarily nonhematopoietic functions also are known to affect the process of the production of new blood cells. This chapter concisely describes the role of both stimulatory and inhibitory hematopoietic growth factors.

Hematopoietic Microenvironment

The term hematopoietic microenvironment refers to the “stromal” elements of the organs in which hematopoiesis occurs, that is, the cellular and noncellular elements that do not directly give rise to the blood cells but rather provide a solid three-dimensional structural matrix in which the hematopoietic stem cells and their progeny proliferate and differentiate until they migrate into the blood stream. The role of the hematopoietic microenvironment in the control of the development of blood cells is believed to be of paramount importance in the process of hematopoiesis. Both the stromal cells and their secreted matrix proteins appear to influence the process of hematopoiesis as profoundly as the soluble, secreted hematopoietic growth factors, and a summary of current understanding of the contribution of the hematopoietic microenvironment is included in this chapter.

HISTORICAL CONSIDERATIONS

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

Early Studies

The scientific study of hematopoiesis is a recent phenomenon. Although the cellular nature of blood was discerned by van Leeuwenhoek and reported in 1674, it was not until the middle of the 19th century that interest was aroused in the origin of the cellular components of blood. The earliest observations on hematopoiesis were the studies of Weber and Kolliker, who determined in 1846 that, in the fetus, the liver is the principal site of hematopoiesis. In 1868, two scientists working independently, Ernst Neumann and Giulio Bizzozero, discovered that in the adult, erythropoiesis, the process of the differentiation of red blood cells, occurs within the bone marrow, with subsequent release of newly formed erythrocytes into the bloodstream. Bizzozero also recognized that white cells were formed in the bone

marrow, and he coined the term “platelet” for thrombocytes. The origin of the blood platelets from megakaryocytes in the bone marrow was not uncovered until 1906, however, when James Homer Wright, the inventor of the stain used most frequently for blood and hematopoietic tissues, noted the shedding of platelets from the cytoplasm of megakaryocytes.

The identification of the bone marrow as the principal hematopoietic organ in the adult was followed, at the close of the 19th century and the first half of the 20th century, by a proliferation of various theories of the cellular origin of blood cells. While it was universally understood during that time that the various blood cells must derive from some sort of ancestral cell or cells, the nature of the precursor cell(s) remained highly controversial.

Two basic theories were advanced, each with highly opinionated proponents, and each with a myriad of variations. The first theory, the monophyletic theory, postulated that all different forms of blood cells were derived from a common ancestral cell, the totipotent hematopoietic stem cell. The term totipotent indicated its capability of differentiating into any and all of the various forms of blood cells. The monophyletic theory was first proposed by Artur Pappenheim at the turn of the century (2) and was supported vigorously by some of the most prominent medical scientists of this century, including the Russian anatomists/embryologists Vera Dantschakoff and Alexander Maximow, Maximow's pupil William Bloom, and Hal Downey, a hematologist and the editor of an important work called the *Handbook of Hematology*.

In opposition to the monophyleticists, a second group arose, the polyphyleticists, who posited the existence of independent lines of hematopoietic cell development for various types of blood cells. For example, Paul Ehrlich, the first to apply histologic staining techniques to the study of blood tissues, proposed the “dualist” theory, viz., that granulocytes derived from a primitive “myelocyte,” whereas the lymphocytes had a completely separate origin (3). Others who were considered polyphyleticists included the Swiss hematologist Otto Naegeli, Robert Schilling of the University of Wisconsin, and Florence Sabin. Each described various modifications of the polyphyletic theory, all of which, however, basically revolved around the notion that the different morphologically recognizable lineages of blood cells derived from different populations of precursor cells that did not have the capability of giving rise to several different types of blood cells.

Experimental evidence to settle this vexing issue did not come until the middle of the 20th century, when evidence of the multilineage capability of hematopoietic stem cells was convincingly demonstrated in transplantation experiments in lethally irradiated mice (4, 5). The monophyletic viewpoint has now largely been vindicated, and at the present, most experts in hematopoiesis agree on the existence of a totipotent hematopoietic stem cell that has the capability, as a single cell, of giving rise to cells of all the different lineages of blood cells, including lymphoid, myeloid (granulocytes and monocytes/macrophages), erythroid (red blood cells), and megakaryocytic lineages (6).

HEMATOPOIESIS IN THE EMBRYO AND FETUS

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

Role of the Yolk Sac

The fertilized egg first develops the beginnings of blood tissue while still in the embryonic stage (7). The first step toward the development of blood tissue is thought to occur in the yolk sac, where undifferentiated cells called mesoblasts are found and are believed to migrate there from the primitive streak of the embryo. The mesoblasts are highly mitotically active, and will subsequently differentiate into cells that are clearly related to the mature blood cells of the adult, called “primitive erythroblasts,” as well as into cells called primitive endothelial cells, which give rise to vascular channels in the yolk sac. Within hours after migration, the yolk-sac mesoblasts have generated, by a process of cell division and differentiation, primitive erythrocytes, primarily nucleated but including a minority that are nonnucleated, all of which acquire hemoglobin and thus lend a reddish color to the clumps of yolk-sac cells where blood-cell formation is occurring. The clumps of hemoglobinized cells are visible to the naked eye, hence the name “blood islands” to describe these localized areas of embryonic hematopoiesis in the yolk sac. Megakaryocytes also are found in the blood islands and are presumably derived from the mesoblasts. Other mesoblasts appear to differentiate into a type of cell called the “hemocytoblast.”

A second stage of hematopoiesis in the yolk sac occurs in the embryos of some mammals. In human embryos, this second stage is present but is not as vigorous as in rabbits, the mammal in which the embryogenesis of blood cells has been most extensively studied. In the second stage of yolk-sac hematopoiesis, hemocytoblasts differentiate into “definitive” erythroblasts, which subsequently acquire hemoglobin and are called “definitive” or “secondary” normoblasts. These may lose their nuclei by a process of extrusion and become “definitive” erythrocytes. Vascular channels form in the blood islands, and eventually connect to form a network of blood vessels. This network of primitive blood vessels early on contains the primitive erythroblasts and hemocytoblasts, and later definitive erythroblasts and erythrocytes. By the end of the third week of embryonic development in the rabbit, all the hematopoietic activity of the blood islands has subsided and the process of hematopoiesis has gradually shifted to the liver.

The Embryonic Body Mesenchyme

A minor role in early embryonic hematopoiesis is played by primitive mesenchymal cells in the body cavity itself, particularly in the head mesenchyme. Small numbers of mesenchymal cells of the body cavity develop into erythroblasts, megakaryocytes, granulocytes, and phagocytic cells analogous to their counterparts in the adult. Quantitatively, the number of cells produced is small, and large accretions of blood cells similar to the blood islands of the yolk sac do not form in the body cavity mesenchyme. The small clusters of blood cells that do form quickly degenerate and no further blood formation occurs here to any significant degree after the third week. However, the formation of blood cells in the mesenchyme is illustrative of the fact that mesenchymal cells have potential for differentiation into hematopoietic cells, and this potential may in theory reside with mesenchymally derived cells throughout life.

Emergence of the Liver As the Principal Site of Hematopoiesis in the Embryo

Beginning at around the 12-mm stage of the human embryo, blood formation is seen to occur in the embryonic liver. The liver

soon becomes the dominant site of hematopoiesis and remains active in hematopoiesis until birth. As the endodermal cords of the liver primordium grow into the septum transversum, they encounter wandering mesenchymal cells with the appearance of lymphocytes. These lymphocytoid-wandering cells are subsequently trapped between the primordial liver endodermal cords and the endothelial cells of ingrowing capillaries. They give rise to hemocytoblasts similar to those of the yolk sac. These hemocytoblasts soon form foci of hematopoiesis similar to the blood islands of the yolk sac, wherein secondary erythroblasts are formed in large numbers, which subsequently divide and differentiate into definitive erythrocytes through the progressive acquisition of hemoglobin and loss of the cell nucleus. Although definitive erythrocytes may be seen in the liver at the 12-mm stage, they do not emerge into the circulation in any great numbers until much later. Thus, by the 70-mm stage, the majority of circulating erythrocytes in the embryo are secondary (definitive) erythrocytes.

Megakaryocytes also appear to form from the hemocytoblasts in the embryonic and fetal liver. Granulocytic cells are found in the embryonic liver, but they appear to develop not from the hemocytoblasts but perhaps directly from the lymphocytoid-wandering cells themselves. The embryonic liver is not the dominant site of hematopoiesis for long, being replaced in importance by the embryonic bone marrow beginning around the 25- to 30-mm stage in the human.

The Embryonic Bone Marrow and Myelopoiesis

Bone formation in the embryo occurs at varying times for different bones. The earliest bones to form are the long bones of the appendicular skeleton. Initially, a cartilaginous model of each bone is formed. The central core of the diaphysis of each long bone subsequently becomes ossified, and soon an area of bone resorption develops followed by the ingrowth of mesenchymal cells from the periosteum. These mesenchymal cells are accompanied by the ingrowth of capillaries. The mesenchymal cells continue to increase in number by continued influx of other mesenchymal cells as well as by division of those already within the newly forming marrow cavity. They also elaborate a noncellular ground substance, or “stroma,” which fills the developing marrow cavity. Cells identical to the hemocytoblasts of the liver and yolk sac develop from these early marrow mesenchymal cells. As in the yolk sac and liver, these give rise to megakaryocytes and erythroid cells. However, there is a second population of cells, called the “lymphoid-wandering cells,” which appear nearly identical to lymphocytes. These cells predominate in the population of cells in the very early marrow cavity, and are thought to give rise to myeloid cells, including neutrophils, basophils, and eosinophils. The embryonic marrow differs markedly from the earlier centers of hematopoiesis in that the generation of these myeloid cells is especially vigorous and dominates the embryonic marrow hematopoietic activity. The process of formation of the early myeloid cells, or “myelopoiesis,” occurs first in the central portion of the marrow cavity and spreads outward from there to eventually include the entire marrow cavity. Erythropoiesis occurs slightly later in the embryonic marrow and generally is admixed with the process of myelopoiesis. Small foci of erythropoiesis thus can be seen among the many maturing cells of myeloid lineage. After birth, hematopoiesis ceases in the liver, and the bone marrow continues to be the principal site of hematopoiesis for the remainder of life.

Hematopoiesis in the Spleen of the Embryo and Fetus

The last major site of hematopoiesis to form in embryonic life is the spleen. Although the spleen itself forms much earlier, wandering mesenchymal cells do not begin to invade the spleen until around the 70-mm stage in humans. These then differentiate into typical hemocytoblasts, which give rise to cells of primarily erythroid lineage. The spleen is thus a center of erythropoiesis until the time of birth, when erythropoiesis gradually ceases. Although some myelopoiesis occurs in the embryonic and fetal spleen, it is relatively insignificant in comparison. Much later, during the fifth month of gestation, the white pulp of the spleen forms by the differentiation of mesenchymal cells that have grouped around the splenic arterioles. The formation of the splenic lymphocytes appears to occur as a process completely separate from the origin of erythropoiesis in this organ.

Other Sites of Hematopoiesis in the Embryo and Fetus

The embryonic thymus develops as an outgrowth of the third branchial pouch. The thymic epithelium is invaded by wandering mesenchymal cells, which begin multiplying rapidly and differentiating into lymphocytes. During this process, small numbers of erythroid and myeloid cells are formed in the thymus, but the primary process is that of lymphopoiesis. The lymphocytes formed in this organ will constitute a distinct class of lymphocytes with a special function: that of cell-mediated immunity.

The lymph nodes develop as outpouchings from the primitive lymphatic vessels, which become surrounded by accretions of mesenchymal cells. Subsequently, these seem to round up and become similar in appearance to the lymphocytes of the adult. A few of the mesenchymal cells give rise to cells of other lineages, such as erythrocytes, granulocytes, and megakaryocytes, but this is a transitory phenomenon, and, as in the thymus, the principal process is that of lymphopoiesis.

Summary of Embryonic and Fetal Hematopoiesis

In all the hematopoietic organs of the embryo and fetus, a similar process takes place. Wandering mesenchymal cells are attracted to a particular site, by processes still not clearly understood, and become transformed into cells recognizable as hematopoietic precursors. These embryonic hematopoietic precursors appear to be capable of multilineage differentiation, but at any one site, the process of hematopoiesis may be dominated by the formation of a particular lineage, presumably under the influence of the local environment. The various sites of embryonic hematopoiesis seem to be active only at specific times during development and follow a pattern of programmed involution, except for the bone marrow, which continues as the

principal location of hematopoiesis in the adult, and the lymph nodes, spleen, thymus, and other lymphatic tissues, which continue to be active in lymphopoiesis.

HEMATOPOIESIS IN THE ADULT

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

After birth, the major site of hematopoietic activity shifts gradually from the liver and spleen to the bone marrow cavities of nearly all bones of the axial and appendicular skeleton. The marrow acquires a reddish color like that of blood once hematopoietic activity begins, reflecting the vigorous production of erythrocytes that contain hemoglobin. The bone marrow cavity serves primarily as a site for the production of nonlymphoid blood cells, whereas lymphopoiesis in the adult occurs primarily in the spleen, lymph nodes, thymus, and the gut-associated lymphoid tissue, including the tonsils, adenoids, and Peyer's patches. Thus, when examined with the light microscope, the adult marrow will be seen to be composed primarily of erythroid and myeloid precursor cells, together with scattered megakaryocytes and a population of cells known as “stromal cells,” which are crucial for the maturation of the precursor cells and release of the fully differentiated cell types into the circulation.

As each individual ages, the marrow of bones of the appendicular skeleton gradually loses its red appearance and is transformed into yellow marrow, a reflection of the progressive replacement of hematopoietic tissue by adipose tissue. Thus, by early adulthood, the long bones no longer bear red marrow but are completely replaced by nonhematopoietic yellow marrow, and the primary sites of red marrow are confined to the sternum, ribs, vertebrae, and pelvis. Although the stimulus for this progressive transformation of red to yellow marrow is unknown, in pathologic conditions associated with vigorous hematopoietic activity, the transformation may fail to take place and the red marrow may actually expand into bones not normally associated with hematopoietic activity, such as the diploic cavities of the cranial bones. The liver, spleen, and lymph nodes also may be locations of “extramedullary hematopoiesis” in such situations. An extreme example occurs in individuals with thalassemia major, a disease in which erythropoiesis is unusually brisk throughout life, resulting in a characteristic expansion of the marrow spaces of all the cranial bones and long bones and enlargement of the liver and spleen. This is so pronounced that the diploe of the calvarium has a characteristic “hair-on-end” appearance in radiographs of the skull as a result of expansion of the marrow space, and maxillary hyperplasia results in a characteristic facies with prominent cheekbones and malocclusion of the teeth because the maxilla is disproportionately larger than the mandible.

THE STEM CELL MODEL OF HEMATOPOIESIS

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

The cell that gives rise to all other types of blood cells, the stem cell, is so rare that it has never been clearly morphologically identified. Rather, the existence of stem cells is inferred by functional assays that demonstrate the ability of single cells to generate multiple hematopoietic lineages. Thus, stem cells currently are defined not by their appearance but by their function (8). Stem cells are known to be extremely rare cells, although quantifying them is somewhat imprecise as a result of different degrees of rigor applied to the definition of what constitutes a stem cell. The most generous estimate is that stem cells occur in human bone marrow with a frequency of one per 1 million nucleated bone marrow cells, while more conservative estimates place this figure at one per 10 million.

The tremendous production rate of hematopoietic cells requires that the bone marrow produce as many cells as it contains roughly every other day. To maintain this rate throughout life, the bone marrow must contain cells that have the ability to generate vast numbers of mature cells continuously, that is, without losing the ability to do so. This *self-renewal* ability is critical to the concept of the stem cell. Presently, there are two theories as to how this might occur. According to the first theory, every stem-cell division is asymmetric, producing one undifferentiated stem cell and one more differentiated cell that is committed to producing mature blood cells (9). In the second theory, each stem cell division produces either two additional stem cells or two more mature cells. The stem cell pool is thus maintained not by precise asymmetric divisions within each stem cell, but rather by a balance between the number of stem-cell divisions yielding more stem cells and divisions yielding more mature cells.

At the point the stem cell leaves the self-renewing pool to populate the differentiating pool, it is still an unrecognizable blast cell with the capacity to produce cells of all lineages. With each subsequent division, the daughter progenitor cells become more and more restricted in their commitment to the production of specific blood-cell lineages. That is, if one isolates progenitor cells and permits them to propagate and differentiate, they will generate collections of cells that are of only one or a few lineages. The more differentiated the progenitor cell, the fewer lineages are produced and the smaller the number of cells produced (10). These concepts, which have been supported by several decades of *in vivo* and *in vitro* experiments, have now defined the hierarchical stem-cell model of hematopoiesis (Fig. 39.1).

The Hematopoietic Microenvironment

If maintained in a simply nutritive environment, stem cells will die without differentiating or dividing. To support the process of hematopoietic self-renewal and differentiation, stem cells and their progeny must be maintained in the close proximity of nonhematopoietic mesenchymal cells, called *stromal cells*. These cells, which are composed of a heterogeneous group of fibro-blasts, endothelial cells, and adipocytes, line the endosteal surfaces in the bone marrow cavity. These cells appear to supply two closely related requirements for the hematopoietic cells, soluble hematopoietic growth factors and membrane-bound attachment molecules (11).

The *hematopoietic growth factors (HGF), or colony-stimulating factors (CSFs)* are a class of glycoprotein hormones that obligately regulate the division and differentiation of hematopoietic cells. These hormones are required for survival, proliferation, differentiation, and function of all the hematopoietic cells. Although initially discovered as spontaneously secreted products of T-cell tumors, it is clear that these hormones are normally the products of bone marrow stromal cells as well as T lymphocytes and monocytes.

CSFs are produced in a two-tiered process. First, small amounts of certain CSFs — interleukin 6 (IL-6), granulocyte-macrophage colony-stimulating factor (GM-CSF), and stem-cell factor (SCF) — are produced constitutively by bone-marrow stomal cells, probably in response to stimulation by plasma proteins. The production of these CSFs is responsible for basal hematopoiesis, maintaining blood counts in the normal ranges (12).

CSF secretion is greatly increased above the basal levels in response to infection. Bacterial and viral products activate monocytes, which then secrete interleukin 1 (IL-1), tumor necrosis factor a (TNFα), and granulocyte colony-stimulating factor (G-CSF), as well as their own macrophage colony-stimulating factor (M-CSF). These products, in turn, stimulate additional CSF secretion. IL-1, together with antigenic stimulation of specific receptors, activates T cells to secrete GM-CSF and interleukin 3 (IL-3) (13). IL-1 and TNFα each stimulates fibroblasts and endothelial cells in the bone marrow stromal microenvironment to increase their secretion of IL-6 and GM-CSF, and also to secrete large quantities of G-CSF (Fig. 39.2) (14). These cytokines thereby directly increase the numbers of circulating neutrophils, monocytes, and plasma cells as well as activating these same mature cells. The generation of each specific lineage of mature blood cells is in this manner regulated by a specific set of hematopoietic growth factors. Although the sets of hematopoietic growth factors that induce specific mature blood-cell subsets overlap, each is characteristically distinct.

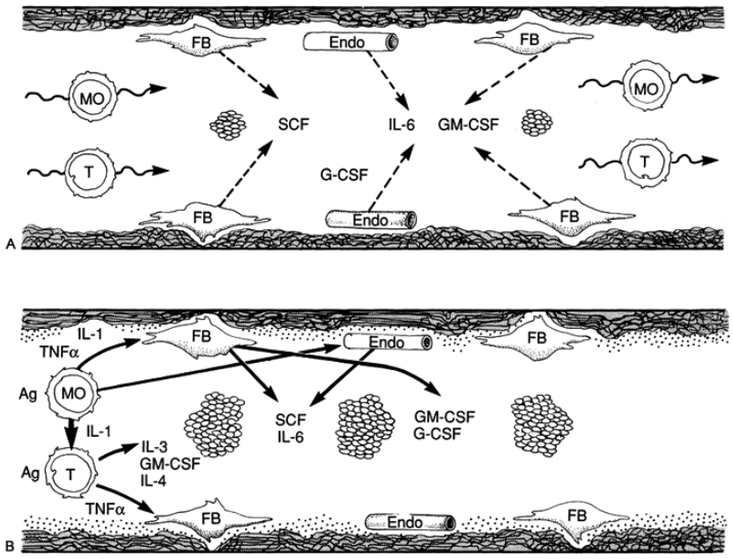


FIGURE 39.1. A: Basal hematopoiesis. **B:** Antigen-amplified hematopoiesis.

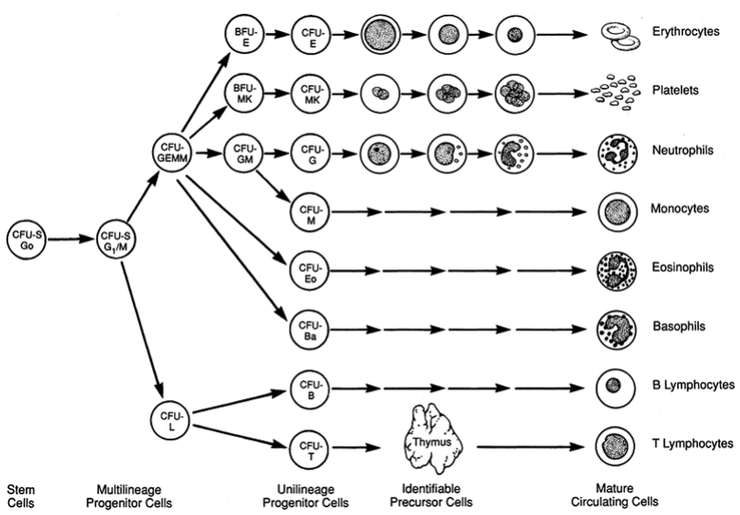


FIGURE 39.2.

Erythropoiesis

The final stages of erythroid differentiation are regulated largely by *erythropoietin*, a glycoprotein produced in response to tissue hypoxia in the fetal liver and adult kidney. Of the 18 or so cell divisions that take place during the time that a stem cell generates a mature red blood cell, the final eight to 10 divisions are strongly induced by erythropoietin. The transcription of the erythropoietin gene in renal peritubular endothelial cells and hepatoblasts is regulated by oxygen-sensitive transcription factors that upregulate gene expression with declining O_2 delivery. Overproduction of erythropoietin, observed in some cases of renal-cell carcinoma and hepatoma, leads directly to erythroid polycythemia.

The preceding cell divisions, which give rise to erythropoietin-sensitive erythroid progenitor cells, are largely erythropoietin-independent. These proliferation and maturation events are instead induced by granulocyte-macrophage colony-stimulating factor (GM-CSF) and *stem-cell factor (SCF)*, both of which are produced locally within the marrow microenvironment by bone-marrow stromal cells. In addition, these steps can be amplified specifically by the secretion of interleukin 3 (IL-3) by activated T lymphocytes (15).

Granulopoiesis

Much like erythroid differentiation, the final stages in neutropoiesis and monopoiesis are induced by *granulocyte colony-stimulating*

factor (G-CSF) and *macrophage colony-stimulating factor (M-CSF)*, respectively. The early divisions, which direct multipotential progenitors to become committed to individual lineages, are regulated by the synergistic interactions of GM-CSF, SCF, and IL-3. As described above, while there is a constant level of basal secretion of CSF by the bone-marrow stromal fibroblasts that line the marrow endosteal surfaces, the secretion of GMCSF and G-CSF is dramatically upregulated in the presence of inflammation in response to the secretion of IL-1 and TNFα by monocytes.

For the production of eosinophils, *interleukin 5 (IL5)* and, to a lesser extent, IL-3 and GM-CSF, play major inductive roles. Basophils and mast cells are directly stimulated by SCF and IL-3. In both of these instances, the initial afferent signals that trigger the release of these cytokines is not yet well understood.

Megakaryopoiesis

The earliest stages in the development of megakaryocytic progenitor cells also appear to be induced by IL-3 and GM-CSF, in conjunction with SCF. What induces the later stages of megakaryopoiesis is not yet certain, but recent evidence suggests that thrombopoietin (TPO) and interleukin 11 (IL-11) play important roles (16). IL-11 and TPO also play roles in the terminal budding events that lead to increased platelet counts.

B Lymphopoiesis

As with the myeloid lineages, the development of B cells begins by the differentiation of pluripotent stem cells into undifferentiated but committed B-cell progenitors. The initial stages in the proliferation and differentiation of these B-cell progenitors are induced by *interleukin 7 (IL-7)* and SCF. Once recognizable pre-B cells and B cells are generated, further differentiation and divisions are induced by stimulation through the immunoglobulin antigen receptor, through the F_cg receptor, and through stimulation by soluble *interleukin 4 (IL-4)* and interleukin 6. Once antibody-producing plasma cells are generated, additional proliferation as well as antibody secretion is stimulated by IL-6 and GM-CSF.

T Lymphopoiesis

Once pre-T cells undergo the complex processes of negative and positive selection in the thymus that generate self/nonself discrimination, the resulting mature T cells are subject to antigen- and cytokine-induced activation and expansion. Stimulation with interleukin 2 (IL-2) as well as antigen leads to preferential

expansion of CD8 T cells, while direct stimulation of the antigen receptor TCR/CD3 along with CD28 leads to preferential activation of CD4 cells.

CLINICAL USE OF HEMATOPOIETIC GROWTH FACTORS

Part of "39 - Hematopoiesis and the Hematopoietic Growth Factors"

Since the first human hematopoietic growth factors were cloned and isolated 8 years ago, preclinical and clinical trials have rapidly led to their introduction into the clinic for routine and experimental use. At the present time, erythropoietin, G-CSF, GMCSF, and IL-11 have all been approved for human use by the Food and Drug Administration (FDA), while IL-3, SCF, and IL-6 have begun phase I and II trials.

Erythropoietin

The *anemia of renal failure* is directly responsive to treatment with erythropoietin. Doses of 50 U/kg three times weekly lead to prompt reticulocytosis, as long as the patients have sufficient iron, folate, and B_{12} stores and have no other source of ongoing inflammation. The only side effect of such treatments is hypertension, if the hemoglobin level rises too high. The introduction of erythropoietin therapy in this manner has made thousands of renal dialysis patients nontransfusion dependent and has greatly improved the quality of their lives (17).

Other anemias, such as in the setting of chronic diseases such as cancer, autoimmune immunodeficiency syndrome (AIDS), and rheumatologic diseases, can also respond to erythropoietin. However, the doses required are higher (150 U/kg or more three times weekly), and even at high doses the responses are variable. In general, the higher the baseline erythropoietin level circulating in the patient's plasma, the less the chance of response to erythropoietin therapy.

A third, growing area of application of erythropoietin therapy is to *autologous donation* of red blood cells prior to elective surgery. By administering erythropoietin under a controlled setting, hematologists can stimulate a mild erythrocytosis, which then allows safe phlebotomy and storage. In this way, it is possible to stimulate and store several units of red blood cells prior to any elective surgery, thus eliminating the need for and risk of allogeneic red-cell transfusion. The only limitations on this therapy are the organizational abilities of hematologists and blood banks. Given the large fraction of blood products that go for urgent and emergent surgeries, allogeneic donation and blood transfusion will still be essential. However, for an increasing fraction of cases, autologous, erythopoietin-stimulated donation and transfusion will be the preferred route for transfusion support.

Granulocyte Colony-Stimulating Factor

When the peripheral neutrophil count is depressed, increased levels of circulating G-CSF can be detected by sensitive assays such as enzyme-linked immunosorbent assay (ELISA). However, the amount of G-CSF produced is suboptimal to stimulate rapid granulopoiesis. Therefore, if additional G-CSF is supplied as a pharmaceutical, the neutrophil count will recover more rapidly and to a higher level. The first major setting in which G-CSF therapy has been applied has been in *chemotherapy-induced neutropenia*. Delivery of G-CSF subcutaneously, in a dosage of 4 to 10 (μg/kg/day), beginning approximately 1 week following chemotherapy, helps lessen the severity of chemotherapy-induced nadirs. Controlled studies have now shown that such treatment helps prevent infections and hospitalizations. Similarly, if patients suffer an infection during a chemotherapy-induced neutropenic nadir, therapy with G-CSF will accelerate their recovery from the nadir. The only side effect observed with G-CSF is bone pain, which occurs in 10% to 15% of patients and can be controlled easily with analgesics.

Pharmacologic doses of G-CSF are also effective in raising the neutrophil count in several cases of chronic neutropenia, such as congenital neutropenia (Kostmann's disease), idiopathic neutropenia, and immune-mediated neutropenias, such as thyroglobulin-lymphoproliferative disease (18). In these cases, however, the doses of G-CSF required can be substantially higher, and the responses are not as uniform.

Granulocyte-Macrophage Colony-Stimulating Factor

Like G-CSF, GM-CSF increases the neutrophil count *in vivo*. However, it has a substantially broader range of activity and also increases the monocyte count and eosinophil count as well. Increases in reticulocytes and platelets also have been reported, but these are less reliable. Given its broad spectrum of activity, GM-CSF has been approved by the FDA for *acceleration of recovery of hematopoietic function following bone-marrow transplantation*. In this setting, GM-CSF clearly shortens the time to recovery, resulting in notably decreased morbidity of the transplantation procedure (19).

In addition, GM-CSF almost certainly has the same salutory effect as G-CSF in preventing and treating chemotherapy-induced nadirs. Although it is possible that its spectrum of side effects (chills, fevers, third-spacing of plasma) might make its use in this setting more problematic than G-CSF, at this time there has been no side-by-side comparison of GM-CSF and G-CSF in any setting.

One recent application of GM-CSF and G-CSF that deserves particular notice is that of progenitor-cell mobilization. In the first trials of GM-CSF, it was found that while the density of bone marrow progenitor cells rose slightly, the density of circulating progenitor cells rose dramatically, often 50- to 100-fold over baseline. Later, very similar data were obtained for G-CSF (and the same is probably true for IL-3 and SCF as well) (20). Based on these data, several bone-marrow transplant centers have used decreasing numbers of peripheral blood leukapheresis collections (from a baseline of eight to 10 down to two to four) to successfully support reengraftment following autologous peripheral blood transplantation. Although there are many questions remaining regarding the cause and the meaning of this effect, as well as the role of progenitor-cell mobilization in transplantation, the application of hematopoietic growth factors to “hemotherapy” in transplantation medicine will be an exciting venue in the decade ahead.

Interleukin-11

Although GM-CSF, IL-3, and IL-6 each have megakaryocytic colony stimulating activity *in vitro*, they have little if any thrombopoietic effect in humans when administered *in vivo*. IL-11, in contrast, does not raise platelet count when administered to patients who became thrombocytopenic following high dose cytotaxin and adriamycin chemotherapy (21). However, the effect of IL-11 is not as uniform as that observed with G-CSF and GM-CSF for neutropenia, so its use to date in the clinic has been more limited.

Stem cell factor (SCF) has to date had little *in vivo* success as a thrombopoietic agent. In contrast, SCF may be more useful as an adjunct to G-CSF for the mobilization of hemotopoietic progenitors and stem cells. However, it should be realized that both SCF and IL-11 are very new drugs, and the full ranges of their clinical utilities are still being determined.

REFERENCES

1. Emerson SG. The stem cell model of hematopoiesis. In: Hoffman R, Benz E, Shattil S, et al., eds. *Hematology, Basic Principles and Practice*. Baltimore: Williams & Wilkins, 1991:72–81.

2. Pappenheim A. Abstammung and Entstehung der roten Blutzelle [Abstract]. *Virchows Arch* 1898;151:89.

3. Ehrlich P. Uber die specifischen Granulationed des Blutes. *Arch Anat Physiol Abt* 1879;2:571.

4. Ford CE, Hamerton JL, Barnes DWH, et al. Cytological identification of radiation chimeras. *Nature* 1961;177:452.

5. Till JE, McCulloch EA. Direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res* 1961;14:213.

6. Lemischka IF, Raulet CH, Mulligan RC. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 1986;45:917–927.

7. Lipton JM, Nathan DG. The anatomy and physiology of hematopoiesis. In: Nathan DG, Oski FA, eds. *Hematology of Infancy and Childhood*. Philadelphia: WB Saunders, 1987:128–158.

8. Metcalf D, Moore MAS. Haematopoietic cells. In: Neugberger A, Tatum EL, eds. *Frontiers of Biology*. Amsterdam: NorthHolland, 1971:25.

9. Holtzer H, Biehl J, Antin P, et al. Quantal and proliferative cell cycles: how lineages generate cell diversity and maintain fidelity. *Prog Clin Biol Res* 1983;134:213–227.

10. Emerson SG, Sieff CA, Wang EA, et al. Purification of fetal hematopoietic progenitors and demonstration of recombinant multipotential colony stimulating activity. *J Clin Invest* 1985;76:1286–1290.

11. Tsai S, Sieff CA, Nathan DG. Stromal cell associated erythropoiesis. *Blood* 1986;67:1418-1426.

12. Guba SC, Sartor CI, Gottschalk LIZ, et al. Bone marrow stromal fibroblasts secrete IL-6 and GM-CSF in the absence of inflammatory stimulation. *Blood* 1992;80:1190–1198.

13. Guba SC, Stella G, Turka LA, et al. Regulation of IL-3 gene induction in normal human T cells. *J Clin Invest* 1989;84:1701–1706.

14. Fibbe WE, van Damme J, Billiau A, et al. IL-1 induces human marrow stromal cells in long-term culture to produce G-CSF and GM-CSF. *Blood* 1988;71:430–435.

15. Emerson SG, Thomas S, Ferra JL, et al. Developmental regulation of erythropoiesis by hematopoietic growth factors: analysis on populations of BFU-E from bone marrow, peripheral blood, and fetal liver. *Blood* 1989;74:49–55.

16. Chang M, Suen Y, Meng G, et al. Differential mechanisms in the regulation of endogenous levels of thrombopoietin and interleukin-11 during thrombocytopenia: insight into the regulation of platelet production. *Blood* 1996;88:3354–3362.

17. Eschbach JW, Egrie JC, Downing MR, et al. Correction of the anemia of end stage renal disease with recombinant human erythropoietin. *N Engl J Med* 1987;316:73–78.

18. Bonilla MA, Gillio AP, Ruggiero M, et al. Effects of recombinant humans G-CSF on neutropenia in patients with congenital agranulocytosis. *N Engl J Med* 1989;320:1574–1580.

19. Nemunaitis J, Singer JW, Buckner CD, et al. Long-term follow-up of patients who received recombinant human granulocyte-macrophage colony stimulating factor after autologous bone marrow transplantation for lymphoid malignancy. *Bone Marrow Transplant* 1991;7:49–52.

20. Kessinger A, Armitage JO. The evolving role of autologous peripheral stem cell transplantation following high-dose therapy for malignancies [editorial]. *Blood* 1991;77:211–213.

21. Issacs C, Robert NJ, Bailey FA, et al. Randomized placebo-controlled study of recombinant human interleukin-11 to prevent chemotherapy-induced thrombocytopenia in patients with breast cancer receiving dose-intensive cyclophosphamide and doxorubicin. *J Clin Oncol* 1997;15:3368–3677.

40

Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders

Curtis A. Hanson

This chapter covers a variety of topics, all included under the broad category of laboratory hematology. This area is really the substantive, day-to-day component of any routine clinical hematology laboratory. Such a laboratory must have a solid foundation in (a) basic morphologic identification of peripheral blood and bone marrow cells; (b) the complete blood count (CBC) and its components; (c) the differential leukocyte count; (d) basic hematology instrumentation and automation, as related to the CBC and differential count; and (e) a variety of miscellaneous, hematology-associated assays, such as reticulocyte count and erythrocyte sedimentation rate (ESR).

In addition to these morphologic and technological considerations, the laboratorian must have an understanding of the clinical disorders, both benign and malignant, associated with abnormalities of the CBC and differential leukocyte count. Conversely, the interpretation of hematologic data by the clinician must be based on an awareness of how the data were generated and how to avoid the production of spurious results. Thus, to gain the greatest clinical benefit from hematologic studies, a mutual understanding of both the laboratory and clinical components is required.

This chapter is divided into three broad areas: (a) peripheral blood and bone marrow morphology, as this is the basis for understanding normal hematopoiesis as well as clinical hematologic disease; (b) blood cell counting, including the CBC and its components, as well as the differential leukocyte count, reticulocyte count, and the ESR; and (c) nonmalignant disorders of white blood cells, including quantitative increases or decreases in any white blood cell count (WBC) component, benign morphologic disorders, and functional disorders of leukocytes.

PERIPHERAL BLOOD AND BONE MARROW MORPHOLOGY

Part of "40 - Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders"

The morphologic evaluation of the peripheral blood and bone marrow involves the evaluation of all components of each sample. Examination of the peripheral blood must therefore consist of evaluating the erythrocytes (red blood cells), leukocytes (white blood cells), and platelets (thrombocytes). The leukocyte component consists of five basic cell types: neutrophils, lymphocytes, monocytes, eosinophils, and basophils. Neutrophils, eosinophils, and basophils can also be referred to as granulocytes or polymorphonuclear leukocytes (PMNs). Among neutrophils there are both segmented and band neutrophil types. Each of these different cell components has a particular morphologic appearance, as well as having unique and essential cellular functions.

Peripheral Blood

Neutrophil (Segmented). The segmented neutrophil is an intermediate-size cell (10 to 15 μm) with three to five nuclear segments. Each of the segments is connected by a thin filament. The nuclear chromatin of the neutrophil is coarsely clumped, with no visible nucleoli. The cytoplasm of the neutrophil contains abundant small, azurophilic (secondary) granules that vary in size (Fig. 40.1).

Neutrophil (Band). The band neutrophil is slightly larger than the segmented neutrophil and has a U-shaped nucleus with coarsely clumped chromatin. Instead of a thin filament separating the lobes, as seen in the segmented neutrophils, a band neutrophil has a much thicker band connecting the nuclear segments. This defining criterion is somewhat vague and has led to significant variability and lack of reproducibility between technologists in determining band counts (see the subsection on neutrophilia in this chapter). Cytoplasmic granulation is identical to that found in the segmented neutrophil.

Lymphocyte. The lymphocyte is a round- or oval-shaped cell that varies from small (7 to 10 μm) to intermediate (10 to 20 μm) in size. The nucleus is usually round to slightly indented or folded, having smooth, mature chromatin; parachromatin is not typically visible. Nucleoli are faint. Cytoplasm will vary from scant to abundant and pale blue to basophilic in color. Occasional azurophilic granules may be identified. These granulated lymphocytes are thought to represent either natural killer cells or lymphocytes with cellular cytotoxic action.

Monocyte. The monocyte is an intermediate to large cell (15

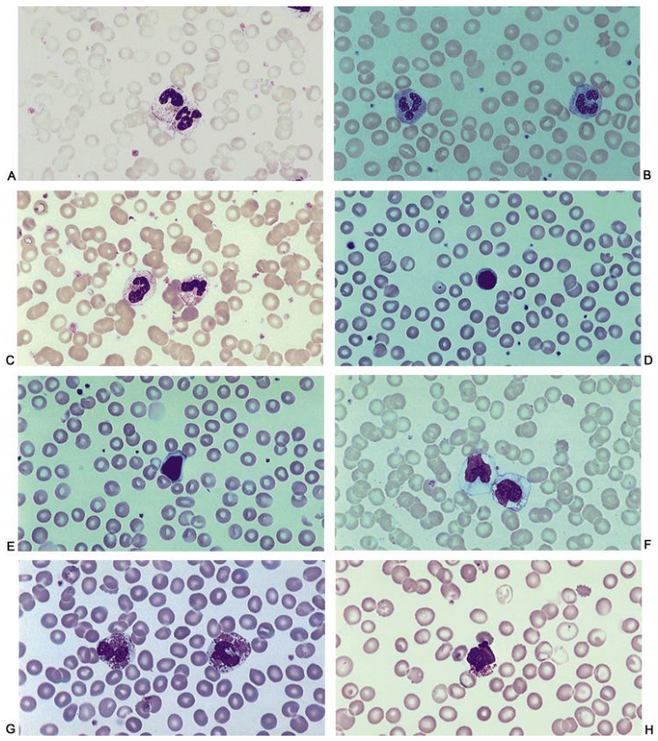


FIGURE 40.1. Peripheral blood smear from a normal individual. **A, B**, Segmented neutrophil; **C**, band neutrophil; **D, E**, lymphocyte; **F**, monocyte; **G**, eosinophil; **H**, basophil.

to 20 μm) with an indented and irregularly folded nucleus. The nucleus consists of what has been described as “raked” chromatin. Nucleoli are typically not seen. The cytoplasm of a monocyte is “dirty,” being grayish blue and occasionally having fine to very small azurophilic granules. Occasional vacuoles may also be identified.

Eosinophil. The eosinophil is an intermediate-size (10 to 15 μm) cell characterized by large, refractile, reddish orange granules. These eosinophilic granules are much larger than those found in neutrophils and are quite distinct in appearance. The nucleus is typically segmented into two or occasionally three lobes. The nuclear chromatin is coarsely clumped with no visible nucleoli.

Basophil. The basophil is similar in size to the neutrophil and eosinophil and contains characteristically large, round, blue-purple granules spread over the nucleus and cytoplasm. This granulation frequently obscures the nuclear features of the cell. The nucleus may contain from one to three lobes and typically has no visible nucleolus. The chromatin of the basophil is slightly different from the chromatin of an eosinophil or neutrophil and is not as coarsely clumped as those two cells. Rather, the basophil nuclear chromatin is smooth, having little visible parachromatin.

Bone Marrow

Granulopoiesis

The normal cellular maturation of the erythrocytic, granulocytic, and monocytic sequence is illustrated in Fig. 40.2. The maturation of the granulocytic series is characterized by sequential steps of reductions in nuclear and cytoplasmic volume, progressive clumping of nuclear chromatin, nuclear lobulation and segmentation, and the sequential loss and gain of primary and secondary granules, respectively (Fig. 40.3). The cells in the first half of the morphologic differentiation scheme (myeloblasts, promyelocytes, and myelocytes) are capable of replication and of undergoing mitosis. The cells in the latter half of myeloid differentiation (metamyelocyte, band neutrophil, segmented neutrophil, eosinophil, and basophil) cannot divide and undergo mitosis.

Myeloblast. The myeloblast is the youngest identifiable cell in the granulocytic series. Myeloblasts are medium-size cells (12 to 16 μm), having high nuclear:cytoplasmic ratios, finely reticular chromatin, and prominent nucleoli. The scant cytoplasm is deep blue and will show rare to few primary, azurophilic granules.

Promyelocyte. The promyelocyte is the largest cell of the granulocytic series, measuring 15 to 22 μm. The prominent azurophilic primary granules are the characteristic feature of this cell. Its nucleus remains round to oval, having fine to perhaps slightly condensed chromatin and typically one to two nucleoli. The amount of cytoplasm in the promyelocyte is slightly more than is found in the myeloblast and is basophilic to pale blue in color. No secondary granules are identified. The primary granules typically overlie both the cytoplasm and the nucleus. At the junction between promyelocyte and myelocyte development, the secondary granules appear in the cytoplasm as a pale, yellowish blush, typically in the Golgi area. It may be difficult to appreciate the specific secondary granules; rather, the early secondary granules give a “patchy,” light blue appearance to the cytoplasm.

Myelocyte. The myelocyte is slightly smaller than the promyelocyte with a central to eccentrically located, round to oval nucleus. The chromatin of the nucleus has begun to clump, but no definite indentation in the nucleus has occurred. Nucleoli are absent or inconspicuous. A moderate amount of cytoplasm is present. The cytoplasm of the myelocyte contains both primary and secondary granules. The secondary granules will appear pink to light blue or purple on a Wright-Giemsa stain. The Golgi hof may be quite prominent in these cells. The amount of cytoplasm relative to the nucleus is increased compared with the promyelocyte stage.

Metamyelocyte. The metamyelocyte is the first stage beyond the mitotically active stages of granulopoiesis. The metamyelocyte is only slightly larger than the mature neutrophil (10 to 18 μm) and shows a slightly indented or kidney bean–shaped nucleus. Chromatin is quite dense with no nucleolus evident. Obviously, secondary granules at this point are quite prominent and far outnumber the larger, darker, primary granules.

Neutrophil/Eosinophil/Basophil. The band and segmented neutrophils, eosinophils, and basophils were previously discussed.

Monocytopoiesis

Monocytes and macrophages constitute the so-called mononuclear phagocyte system (Fig. 40.2). These cells arise from committed stem cells capable of generating either monocytic or granulocytic precursors. After differentiation into mature monocytes, circulating monocytes may eventually migrate to target tissues, where they become macrophages. These macrophages may be present in multiple tissue sites, including the spleen, liver, and lungs.

Monoblast. The monoblast is a large cell (14 to 22 μm) with a round, eccentrically placed nucleus. The nucleus has fine chromatin and prominent nucleoli. There is relatively abundant cytoplasm for such a large, blastic-appearing cell, with certainly more cytoplasm than a typical myeloblast. The cytoplasm is pale blue and may rarely contain some faint azurophilic granules.

Promonocyte. The promonocyte is a poorly defined cell, intermediate between a monoblast and a mature monocyte. Like the monoblast, the promonocyte is a large cell with an indented to slightly irregular nucleus. Nuclear folding does not reach the degree of indentation seen in the mature monocyte and thus is intermediate between a monoblast and a monocyte. The nucleus typically shows one to two faint nucleoli. The nuclear chromatin has begun to show slight condensation and a coarseness beyond that of the typical monoblast. The cytoplasm retains the gray-blue color of a monocyte and typically shows some fine azurophilic granules and occasional vacuoles.

Monocyte. The monocyte was discussed previously.

Erythropoiesis

The process of erythropoiesis begins at the pluripotent stem cell through the intermediate stages to the committed erythroid stem cell unit (Fig. 40.1). The earliest identifiable component of the erythroid series is the pronormoblast. This cell undergoes sequential divisions through the basophilic normoblast to eventually

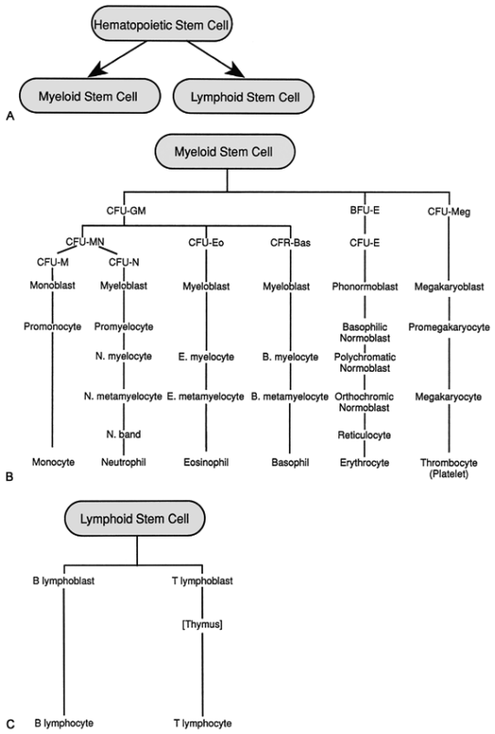


FIGURE 40.2. Schematic line drawing of the normal. **A**, Hematopoietic; **B**, myeloid; **C**, lymphoid maturation sequence. CFU, colony-forming unit; BFU, burst-forming unit; GM, granulocyte/monocyte; M, monocyte; N, neutrophil; Eo, eosinophil; Bas, basophil; E, erythroid; Meg, megakaryocyte.

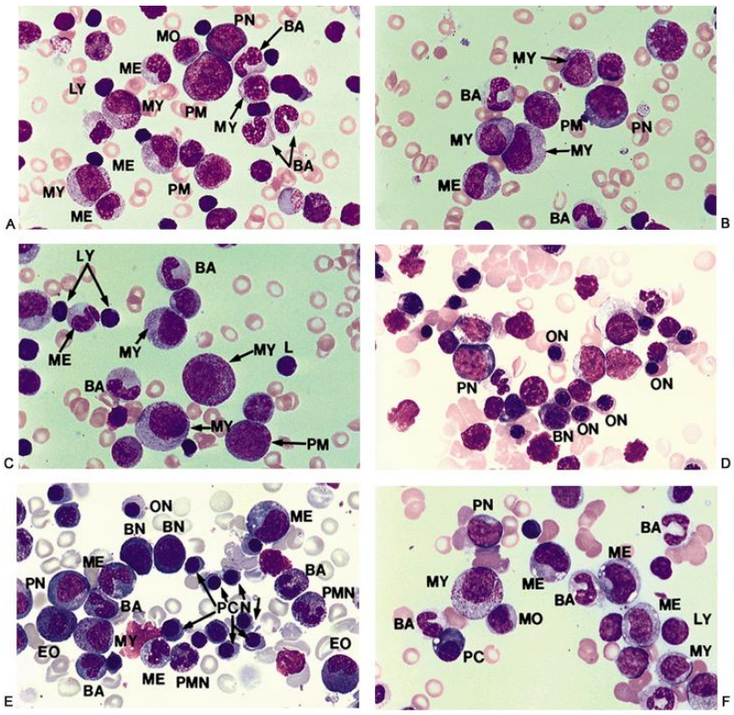


FIGURE 40.3. A–F: Photomicrographs of bone marrow aspirate smears show erythroid and granulocytic maturation. MB, myeloblast; PM, promyelocyte; MY, myelocyte; ME, metamyelocyte; MO, monocyte; PN, pronormoblast; BN, basophilic normoblast; PCN, polychromatophilic normoblast; ON, orthochromatic normoblast; PC, plasma cell.

the reticulocyte and a mature red blood cell. Each division results in a decrease in cell size, nuclear condensation, nuclear pyknosis, and eventual extrusion of the nucleus and hemoglobinization of the cytoplasm. The morphologic characteristics of the normal erythroid series are noted below.

Pronormoblast. The pronormoblast is the largest erythroid precursor (15 to 22 μm), having a uniformly round to slightly oval nucleus. The chromatin is finely reticular with one to two prominent nucleoli. The cytoplasm is deeply basophilic and scant in quantity. There is typically a well-defined chromatin rim around the nucleus.

Basophilic Normoblast. The basophilic normoblast is decreased in size (10 to 16 μm) relative to the pronormoblast, having slightly more cytoplasm and a slightly more condensed chromatin pattern. Nucleoli are inconspicuous. The cytoplasm retains the deeply basophilic color seen in the pronormoblast.

Polychromatophilic and Orthochromatic Normoblast. These two cell types constitute the final two sequences of nucleated erythroid

precursors in the bone marrow. They continue to show the reduction in size seen in erythroid differentiation (10 to 15 μm and 8 to 12 μm, respectively). Hemoglobinization becomes evident at these stages as the cytoplasm turns from basophilic to blue-gray to grayish pink and eventually to a definite reddish-orange color. The nucleus shows further chromatin condensation and becomes quite pyknotic in appearance in the orthochromatic stage.

Reticulocyte. The reticulocyte is the first stage of erythroid development that lacks a nucleus. The cytoplasm has a slight gray-blue tinge to the prominent orange-staining color. Vital staining (e.g., with new methylene blue) will show fine basophilic reticulum in the red blood cell, which is the defining feature of a reticulocyte. This represents residual RNA, remaining from previous precursor stages.

Lymphopoiesis

Lymphocytes play an important role in antigen recognition and the subsequent immune response. The earliest lymphocyte precursors undergo the initial maturation steps in the bone marrow before migration to extramedullary sites; lymphocytes can be antigenically and functionally split into B- and T-cell types (Fig. 40.1). The T-lymphocyte precursors migrate from the bone marrow to the thymus, where further differentiation and maturation occur. The thymus provides an appropriate environment such that immunocompetent T lymphocytes are produced. It is in the thymus that T-cell selection occurs, a process whereby cells that would react with “self” are removed; cells that express the preferred T-cell receptors and antigens and have appropriate histocompatibility recognition antigens are selected for further differentiation. From the thymus, mature T lymphocytes are released to the peripheral lymphoid tissues, including the paracortical regions of lymph nodes, the periarteriolar regions of spleen, and other nonlymphoid sites such as in the mucosa of the gastrointestinal tract and pulmonary tree.

The other major component of the lymphoid system is the B lymphocyte. B cells differentiate to maturity within the bone marrow. The normal development and differentiation of B cells begin at the hematopoietic stem cell, go through the early stages of B-cell differentiation while still in the bone marrow, acquire surface immunoglobulin as mature B lymphocytes, and end with plasma cell production of cytoplasmic immunoglobulin. As the mature B cells are released from the bone marrow, they migrate to B-cell areas of lymph nodes and spleen, such as the follicles, mantle zones, and medullary regions. Other B-cell areas include the follicles and follicular areas of mucosa-associated lymphoid tissue.

Lymphopoiesis can be accurately and precisely defined by immunophenotyping studies, which detect surface antigens present at various stages of lymphoid development. Morphology offers a limited perspective of lymphopoiesis and can identify only the lymphoblast, lymphocyte, and plasma cell components.

Lymphoblast. The lymphoblast is a small- to intermediate-size cell (8 to 15 μm) with a round nucleus and scant cytoplasm. The chromatin is fine but slightly coarser than that found in myeloblasts. It has inconspicuous to prominent nucleoli and no observable cytoplasmic granules.

Lymphocyte. The morphology of the lymphocyte was discussed previously.

Plasma Cell. The plasma cell is a small cell (7 to 10 μm) with an eccentrically placed nucleus. The nucleus has coarsely clumped chromatin that has been described as resembling a clock face. The cytoplasm is a basophilic, cornflower-blue color and has a prominent, perinuclear hof.

Megakaryopoiesis

Megakaryopoiesis involves a much different morphologic sequence of differentiation than the other myeloid cell lines (Fig. 40.1). The megakaryoblast is closely related to the early erythroid precursor and morphologically has some resemblance to a primitive pronormoblast. Erythropoietin, which stimulates bone marrow erythroid production, also stimulates megakaryopoiesis and the production of platelets. The megakaryoblast undergoes endomitosis, which involves chromosomal division without cytoplasmic and cellular division. As endomitosis and nuclear segmentation occur, recognizable megakaryocytic precursors can be identified. Endomitosis continues until the megakaryocyte contains 8, 16, or 32 nuclear lobes. This process is accompanied by an increase in cell size and cytoplasm. Eventually the cytoplasm becomes more granular, and small platelets, which are membrane-bound portions of granular cytoplasm, break off and eventually enter the circulation.

Megakaryoblast. The megakaryoblast is a poorly defined cell that may not be easily recognizable as a distinct entity within a normal bone marrow. It may either be a small blast (8 to 15 μm) or much larger in size (18 to 30 μm). The megakaryoblast has a round to oval nucleus, prominent nucleoli, and variable amounts of blue to deeply basophilic cytoplasm. The chromatin may be finely reticular in larger cells or dark and dense in smaller cells.

Megakaryocyte. The megakaryocyte is the largest nucleated cell within the bone marrow and ranges in size from 35 to 100 μm. The megakaryocyte typically has a multilobated nucleus with 8, 16, or 32 nuclear lobes. No nucleoli are evident. The cytoplasm appears light blue and contains finely dispersed granules.

BLOOD CELL COUNTING

Part of "40 - Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders"

The complete hematologic evaluation of a peripheral blood sample can provide important clinical and diagnostic information about the three major components of the peripheral blood: (a) the erythrocytes or red blood cells, (b) the leukocytes or white blood cells, and (c) the platelets. These studies are useful in screening patients for potential hematologic disorders as well as providing diagnostic information for specific diseases as described in this and other chapters of this section.

The process of performing a basic hematologic analysis of peripheral blood involves four primary steps: (a) collection and processing of the peripheral blood sample, (b) determination of the CBC, including cell concentrations and red blood cell indices, (c) determination of the differential WBC (either automated or manual); and (d) blood film examination for potential morphologic changes/abnormalities. Although these steps are

discussed in some detail here, it may not be appropriate or necessary to perform all four steps on all samples. Rather, there must be appropriate use of laboratory resources according to the presenting clinical indications.

Specimen Collection and Processing

The handling and processing of blood samples can be a forgotten component in providing quality results for the hematology laboratory. Table 40.1 lists some of the preanalytic and analytic errors that can affect hematologic results. Understanding these potential sources of errors is essential knowledge for both laboratorians and clinicians. These include specimen collection, choice of anticoagulant, specimen clotting, storage temperature, time between acquisition and analysis, hemolysis, adequate specimen mixing, hyperlipemia, and other metabolic/biochemical abnormalities. Recognition of these factors by laboratory personnel is essential in providing accurate clinical hematologic data for their clinical colleagues.

TABLE 40.1. THE COMPLETE BLOOD COUNT (CBC): SOURCES OF SPURIOUS RESULTS

CBC Component	Causes of Spurious Increase	Causes of Spurious Decrease
White blood cells (WBC)	Cryoglobulin Heparin Monoclonal proteins Nucleated red cells Platelet clumping Unlysed red cells	Clotting Smudge cells
Red blood cells (RBC)	Cryoglobulin Giant platelets High WBC (>50,000/μL)	Autoagglutination Clotting Hemolysis (*in vitro*) Microcytic red cells
Hemoglobin (Hb)	Carboxyhemoglobin (>10%) Cryoglobulin Hemolysis (*in vivo*) Heparin Hyperbilirubinemia Lipemia Monoclonal proteins	Clotting
Hematocrit (Hct) (automated)	Cryoglobulin Giant platelets High WBC (>50,000/μL) Hyperglycemia (>600 mg/dL)	Autoagglutination Clotting Hemolysis (*in vitro*) Microcytic red cells
Hematocrit (microhematocrit)	Hyponatremia Plasma trapping	Excess ethylenediaminetetraacetic acid Hemolysis (*in vitro*) Hypernatremia
Mean corpuscular volume	Autoagglutination High WBC (>50,000/μL) Hyperglycemia	Cryoglobulin Giant platelets Hemolysis (*in vitro*) Microcytic red cells Swollen RBC
Mean corpuscular hemoglobin	High WBC (>50,000/μL) Spuriously high Hb Spuriously high RBC	Spuriously low Hb Spuriously high RBC
Mean corpuscular hemoglobin concentration	Autoagglutination Clotting Hemolysis (*in vitro*) Hemolysis (*in vivo*) Spuriously high Hb Spuriously low Hct	High WBC (>50,000/μL) Spuriously low Hb Spuriously high Hct
Platelets	Cryoglobulin Hemolysis (*in vitro* and *in vivo*) Microcytic red cells Red cell inclusions White cell fragments	Clotting Giant platelets Heparin Platelet clumping Platelet satellitosis

Modified from Combleet J, Spurious results from automated hematology cell analyzers. *Lab Med* 1983;14:509.

Venipuncture/Anticoagulants

Peripheral blood for hematologic tests is typically obtained from a venipuncture specimen. Blood may also be drawn from pricking the fingertip, heel, great toe, or earlobe. The blood should be collected into a tube containing an anticoagulant and thoroughly mixed. The choice of anticoagulants for hematologic studies include ethylenediaminetetraacetic acid (EDTA) (sodium or potassium salts), trisodium citrate, and heparin, with EDTA being the standard anticoagulant of choice for hematologic studies. Citrate and EDTA are, in effect, calcium chelators and thus remove calcium, which is an essential ingredient for the coagulation

process. Heparin inhibits thrombin formation by complexing with antithrombin III.

EDTA specimens may produce significant morphologic artifacts on blood films if allowed to stand for more than 2 to 3 hours, although instrument analysis is stable for as long as 8 hours at room temperature. If analysis cannot be performed within 8 hours, clinical specimens can be refrigerated for as long as 24 hours at 4°C with little degradation of instrument results. Specimens with EDTA kept at room temperature for more than 8 hours may show an artifactual increase in the mean corpuscular volume (MCV) and a decrease in both the mean corpuscular hemoglobin concentration (MCHC) and ESR.

Citrate is most commonly used for coagulation and special platelet studies but may be useful for routine hematologic studies if EDTA-dependent platelet agglutination is present; blood counts should be increased by a factor of 1.1 if citrate is used to account for the sample dilution that occurs with the use of sodium citrate. Heparin is a sufficient anticoagulant for red blood cell analysis but may cause clumping of platelets or leukocytes. These imitations of available anticoagulants have basically mandated that most hospital-based hematology work be performed on-site and that referral hematologic specimens be acquired within a relatively short time of analysis to minimize laboratory and morphologic artifact.

Complete Blood Count

The CBC is the backbone of any hematologic evaluation. The CBC includes a determination of (a) red blood cell data, such as total red blood cell count (RBC); hemoglobin/hematocrit; red blood cell indices, including MCV, MCHC, and mean corpuscular hemoglobin (MCH); and the red cell distribution width (RDW); (b) white blood cell data; and, usually, (c) platelet count and, sometimes, mean platelet volume (MPV). A differential WBC may also be part of a routine CBC in some laboratories, although a more judicial use of the differential count is probably warranted.

The CBC in the broadest sense can provide important baseline information about the functional state of the bone marrow. Recognition of a "-penia," or deficiency of a blood component, is indication of either a marrow-production problem or peripheral destruction activity. Likewise, a "-cytosis," or elevation of a blood component, could indicate either a normal marrow response to a peripheral stimulation or the peripheral manifestation of an uncontrolled malignant proliferation. Thus, in general, the CBC is useful in broadly screening for hematologic disease (Table 40.2). The hemoglobin and/or hematocrit quantify the degree of anemia or polycythemia and the MCV/MCHC/RDW can be useful in further subclassifying the type of anemia (e.g., normocytic, microcytic, macrocytic). The WBC can provide diagnostic or follow-up information regarding either a benign response or malignant process (see later in this chapter). Obtaining a platelet count is the first step in the evaluation of the hemostatic process. Any of the CBC components can be useful in monitoring patient responses to therapy (e.g., iron, vitamin B_{12}/folate, marrow toxicity owing to drugs, chemotherapy).

TABLE 40.2. INDICATIONS FOR ORDERING THE COMPLETE BLOOD COUNT

Ambulatory population	
General population (screening)	Not useful
Specific subgroups (e.g., pregnant women, elderly, immigrants.)	Possibly useful
Hospital population	
No abnormality suspected	Rarely useful
Hematologic abnormality suspected	Useful
Repetitive testing	Useful in some patients at appropriate intervals

The appropriate ordering of the CBC is not something that has been seriously debated in the medical literature. Indeed, the role of the CBC has been considered a virtual icon in the evaluation of hospitalized as well as ambulatory patients. The challenge to the medical community will be to develop practice parameters for CBC utility in various patient settings. The CBC as a screening device in the ambulatory setting has limited utility in the general population because of the low prevalence of asymptomatic disease that would be detected by a routine CBC. Indeed, even the detection of mild asymptomatic abnormalities with the CBC, such as mild anemia, may not offer any real clinical benefit to the patient. Screening of particular subgroups, such as pregnant women, institutionalized geriatric persons, and possibly immigrants from Third-World countries, may be useful because the prevalence of hematologic disease is higher in those groups than in the general population. However, no studies have been done in these groups to justify the benefit of CBC screening for disease.

It is in hospitalized patients that the routine CBC has been most abused both as an admission screen and as a repetitive follow-up test. If no hematologic abnormality is suspected clinically, the CBC is rarely useful on patients admitted to a hospital. Those clinical situations in which a routine CBC would be of questionable value include elective surgeries in which only minor blood loss is anticipated, for patients undergoing minor diagnostic procedures, and for patients who have no clinical indication of a hematologic abnormality. If, however, a hematologic abnormality is suspected clinically, the CBC is obviously indicated and may be useful in confirming either a primary or reactive hematologic disorder that may have an impact on disease diagnosis and management.

It is also difficult to document the effectiveness of repetitive CBC testing. Unfortunately, it has become common practice to order repetitive CBCs (e.g., daily) for patients for whom no benefit may be gleaned. If a patient clinically responds to treatment for a hematologic disease, then ordering a repetitive CBC would not be beneficial unless there is indication that such treatment has not been effective. This aspect of defining appropriate testing intervals for the CBC has not been studied extensively and remains a major challenge.

Table 40.2 provides a summary of the utility and indications for obtaining a CBC. Although laboratories have generally not sought to control the flow of CBC samples, it seems inevitable that some limitation of its use will evolve as health care cost containment grows in importance. The greatest challenge to laboratorians will be devising practice standards that outline more effectively the indications for routine and follow-up CBCs for both hospitalized and ambulatory patients.

Hematology Analyzers/Instrumentation

Clinical hematology laboratories have undergone significant changes in the arena of automation as technological advances have continually evolved. The era of manual CBC determinations has thankfully ended, with automated blood analysis now a routine part of virtually all hematology laboratories. Automated red blood cell and white blood cell counters emerged in the clinical laboratories in the 1960s. The addition of platelet counters, seven-parameter CBC analyzers, and three-parameter differential leukocyte counters appeared in the 1970s. The 1980s were characterized by single, stand-alone instruments capable of performing 10-parameter CBCs. The greatest change in the hematology laboratory came in the latter part of the 1980s and the early 1990s with the development of more sophisticated and accurate leukocyte differential counters. Changes that have emerged over the past few years include bar-code readers, closed-tube sampling, automated sampling systems, and “walk-away” analysis stations. Bar-code readers have allowed for efficient and high throughput of specimens and provide more accurate patient and specimen identification. Closed-tube sampling has been a tremendous benefit in this era of universal precautions and the recognition of high-risk blood samples. The “walk-away” features of many instruments has been made possible because of the built-in mixing systems as well as simple robotics for specimen aspiration and handling.

It is improbable that more clinically relevant parameters will be extrapolated from the CBC. Therefore, one would predict that the next decade or two will show advances and changes related to two areas: increasing automation and walk-away capabilities of the instrument with the eventual use of robotics to handle and process specimens and increasing use of computer analytic capabilities to control utilization of hematologic testing as well as providing more innovative ways to display and interpret laboratory data.

One can envision the evolution of robotic techniques that would allow specimens to be transported from a central receiving area to the hematology instrument based on appropriate specimen bar coding. Technologist interaction would be needed at the point of troubleshooting or manual slide making (another potential automated area). The power of today's computers to store and analyze large clinical databases will also permit the evolution of autoverification of patient results. In addition, by interrogating the patient database, computer capabilities may allow the laboratorian better control over the utilization of the CBCs and differential WBCs; in particular, controlling the frequency of repetitive CBCs and differential WBCs. Last, more sophisticated laboratory information systems will, it is hoped, allow the laboratories to present their data in a more meaningful clinical fashion. For example, the capability to present data graphically or in combination with results from other laboratories (for example, hemoglobin values with Coombs' test, transfusion record, and bilirubin) could be a major enhancement in how laboratory data are interpreted clinically.

Total RBC

For most medical technologists, the hemacytometer is merely a reminder of how technology has improved their lives. It is safe to write that virtually all modern laboratories use some form of automated counting to determine the total RBC. Nearly all modern instruments use either electrical impedance methods or laser light-scatter characteristics to determine the RBC. Although frequently ignored in the evaluation of a CBC, the RBC is the basis for calculating the hematocrit, MCH, and MCHC. In addition, the RBC may have value by itself when attempting to distinguish iron deficiency anemia from thalassemia in patients with an unknown microcytic anemia. In iron deficiency, the RBC diminishes in proportion with the hemoglobin concentration. This is in contrast with thalassemia, in which the RBC may be normal to increased relative to the degree of anemia as shown by the hemoglobin value. Several mathematical formulas have been derived in an attempt to distinguish these disorders more accurately; none has proved perfect, and most have been too complex to remember. A simple formula reported by Baker and Cornbleet (1) takes the MCV divided by the RBC; a value greater than 13 favors iron deficiency and a value less than 13 favors thalassemia.

Automated counting methods for red blood cells have been based primarily on electrical impedance or light-scattering techniques. These methods allow both the counting of total cells and determining the cell size (i.e., the MCV) of the red blood cells. Analyzer from the Beckman Coulter, Inc., Fullerton, California uses the electrical impedance method, whereas light-scattering techniques have been pioneered with the Bayer Advia, Tarrytown, New York analyzers. The electrical impedance method requires that cells pass an aperture through which a current is passing. As the cells cross the electrical current, the change in the electrical resistance that can be detected is proportional to the cell size. In light-scattering techniques, photomultiplier tubes detect changes in light scatter as the cells pass through a flow cell. Various cell types can be distinguished from these light-scatter characteristics based on cell size and granularity.

Very few clinical situations will result in false elevations or decreases in the total RBC. Red blood cell autoagglutination and extreme red blood cell microcytosis may lead to spurious decreases in the reported total RBC. In the latter case, the small red blood cells may be counted as large platelets and not be included in the red blood cell histograms, whereas autoagglutination can result in the red blood cell “clump” being counted as a single cell. Conversely, very high WBCs as well as cryoglobinemia may both lead to false elevations in total RBC. High WBCs, typically if greater than 100.0×10^9 cells/L, can significantly add to the total RBC, whereas normal WBCs have little impact on the total RBC.

Hemoglobin

The primary function of the red blood cell is transporting oxygen and carbon dioxide. Hemoglobin within the cell takes up oxygen within the lungs and releases it in tissues in exchange for carbon dioxide. This regulated delivery and control of oxygen and carbon dioxide exchange is obviously an extremely vital component of the mammalian system. The role of oxygen delivery depends on the binding of hemoglobin to oxygen. The equilibrium between oxygen and hemoglobin varies according to oxygen tension (PO_2). This dissociation curve (the so-called oxygen

TABLE 40.3. HEMOGLOBIN DERIVATIVES: ABSORPTION MAXIMAS

Hemoglobin	Abbreviation/Terms	Maxima (nm)
Hemoglobin	Hb	431,555
Oxyhemoglobin	HbO_2	415,542
Carboxyhemoglobin	HbCO	420,539
Sulfhemoglobin	SHb	—
Carboxysulfhemoglobin	SHbCO	—
Hemiglobin	Hi/methemoglobin	406,500
Hemiglobincyanide	HiCN/cyanmethemoglobin	421,540

Modified from van Assendelft OW. *Spectrophotometry of haemoglobin derivatives*. Assen, The Netherlands: Royal Van Gorcum Ltd., 1970.

dissociation curve) demonstrates the saturation process of hemoglobin relative to the conditions found in both lung (binding of oxygen) and peripheral tissue capillaries (release of oxygen).

Hemoglobin constitutes more than 90% of the red blood cell and is composed of two pairs of globin chains ($\alpha 2B_2$) and four heme groups containing ferrous iron. A decrease in hemoglobin concentration is referred to as an anemia, whereas an increase in hemoglobin concentration is called polycythemia. Some clinical conditions can lead to the formation of oxidized hemoglobin (for example, Heinz body anemias); this oxidized hemoglobin is called methemoglobin, which has a reduced capacity for carrying carbon dioxide. Sulfhemoglobin is an uncommon form of hemoglobin that results from the addition of sulfur to the hemoglobin molecule during the process of oxidation. Sulfhemoglobin cannot carry oxygen and typically accounts for less than 1% of all hemoglobins. Clinically, sulfhemoglobin has been reported in some patients receiving sulfa-based drugs, as well as in some patients with clostridium bacteremia. Carboxyhemoglobin normally accounts for less than 1% of the overall hemoglobin concentration and is produced during the process of heme metabolism to bilirubin.

The interaction of carbon monoxide and hemoglobin deserves specific mention. The affinity of carbon monoxide to hemoglobin is much greater than with oxygen (more than 200 times greater). Thus, even small concentrations of carbon monoxide in the air will preferentially bind to hemoglobin and thus prevent oxygen exchange and transportation. Acute carbon monoxide poisoning is the most dramatic cause of elevated carboxyhemoglobin levels. Long-term carbon monoxide exposure as a result of smoking, or smoke exposure, may also lead to elevation of carboxyhemoglobin and a "left shift" in the oxygen dissociation curve. The differentiation of these various hemoglobin derivatives is classically based on spectrophotometric determination. The various hemoglobins have particular absorption spectra, shown in Table 40.3.

Hemoglobin is most commonly determined by spectrophotometry in modern instrumentation using a cyanomethemoglobin procedure; this is the standard method by which virtually all hematology instruments measure hemoglobin. This procedure involves oxidizing all hemoglobin to methemoglobin by using excess ferric iron (potassium ferricyanide). Excess cyanide ions (potassium cyanide) subsequently leads to the formation of cyanated methemoglobin (which maximally absorbs light at 540 nm). Other forms of hemoglobin besides cyanomethemoglobin, including hemoglobin, oxyhemoglobin, carboxyhemoglobin, and methemoglobin, will also absorb light at 540 nm (Table 40.3). Falsely elevated (real or apparent) hemoglobin absorbances can occur owing to hyperlipemia, fat droplets (associated with hyperalimentation), hypergammaglobulinemia, cryoglobulemia, or a leukocytosis (more than 50.0×10^9 cells/L). Improperly collected blood specimens (venipuncture or fingerstick) can also cause falsely elevated (hemoconcentrate) or falsely decreased (hemodilute) levels. Other laboratory methods for the determination of hemoglobin have been based on detecting oxyhemoglobin or by detecting the iron content of whole blood. Neither method is widely performed because of the lack of standardization or the inherent complexity of the process.

The automated determination of hemoglobin has bypassed many of the inherent errors of manual hemoglobin detection. However, variation in results can still occur from a variety of situations. As discussed earlier, sample collection problems or specimen abnormalities (e.g., hyperlipemia) can give falsely elevated or decreased hemoglobin levels. Hemoglobin concentrations are reported in grams per deciliter or grams per liter and vary according to age, gender, and ethnic background (Table 40.4). Cord blood or newborn capillary blood can be more concentrated (as much as 2 g/dL) than venous blood; premature infants have lower hemoglobin values than term infants. The average hemoglobin concentration drops dramatically during the first month of life and does not significantly increase until after puberty. Adult females have hemoglobin values of 1.0 to 2.0 g/dL lower than adult males. It also has been recognized that blacks of both genders have slightly lower mean hemoglobin levels than whites. All this normal variation mandates that different normal ranges be utilized for reporting hematologic results and that the clinical interpretation of CBC findings be done in conjunction with all appropriate clinical findings.

TABLE 40.4A. COMPLETE BLOOD COUNT: NORMAL VALUES

Age	Hb[a] (g/dL)	Hct (%)	RBC ($\times 10^{12}$-L)	MCV (fl)	MCH (pg)	MCHC (g/dL)	RDW (%)
At birth	13.5–19.5	42–60	3.9–5.4	98–118	31–37	30–36	—
1 day	14.5–22.5	45–67	4.0–6.6	95–121	31–37	29–37	—
1 wk	13.5–21.5	42–66	3.9–6.3	88–126	28–40	28–38	—
1 mo	10.0–18.0	31–55	3.0–5.4	85–123	28–40	29–37	—
6 mo	9.5–13.5	29–41	3.1–4.5	74–108	25–35	30–36	—
1 yr	10.5–13.5	33–39	3.7–5.4	70–86	23–31	30–36	—
6 yr	11.5–13.5	34–40	3.9–5.3	75–87	24–30	31–37	—
12 yr	11.5–13.5	35–45	4.0–5.2	77–95	25–34	31–37	—
Adult female	11.7–15.7	34.9–46.9	3.8–5.2	80.8–100	26.5–34.0	31.4–35.8	<15
Adult male	13.5–17.5	39.8–52.2	4.4–5.9	80.5–99.7	26.6–33.8	31.5–36.3	<15

[a] Mean Hb level in blacks of both sexes and all ages may be 0.5–1.0 below the mean for whites.
Hb, hemoglobin; HCT, hematocrit; RBC, red blood cells; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red cell distribution width.
Data from Baker J, Cornbleet PJ. Erythrocyte disorders. In: Howanitz JH, Howanitz PJ, eds. *Laboratory medicine—test selection and interpretation*. New York: Churchill-Livingstone, 1991:447–498.

TABLE 40.4B. WHITE BLOOD CELL AND LEUKOCYTE DIFFERENTIAL COUNTS: NORMAL VALUES

		Neutrophils ($\times 10^9$/L)						
Age	WBC ($\times 10^9$/L)	Total	Band	Segmented	Lymphocytes	Monocytes	Eosinophils	Basophils
At birth	18.1 (9.0-30.0)	11.0 (6.0-26.0) 61%	1.61 9.1%	9.4 52%	5.5 (2.0–11.0) 31%	1.05 (0.40-3.1) 5.8%	0.40 (0.02–0.85) 2.2%	0.10 (0–0.64) 0.6%
24 hr	18.9 (9.4-34.0)	11.5 (5.0-21.0) 61%	1.75 9.2%	9.8 52%	5.8 (2.0–11.5) 31%	1.10 (0.20-3.1) 5.8%	0.45 (0.05–1.00) 2.4%	0.10 (0–0.30) 0.05%
1 wk	12.2 (5.0-21.0)	5.5 (1.5-10.0) 45%	0.83 6.8%	4.7 39%	5.0 (2.0–17.0) 41%	1.10 (0.30-2.7) 9.1%	0.50 (0.07–1.10) 4.1%	0.05 (0–0.25) 0.04%
1 mo	10.8 (5.0-19.5)	3.8 (1.0–9.0) 35%	0.49 4.5%	3.3 30%	6.0 (2.5–16.5) 56%	0.70 (0.15–2.0) 6.5%	0.30 (0.07–0.90) 2.8%	0.05 (0–0.20) 0.5%
6 mo	11.9 (6.0-17.5)	3.8 (1.0–8.5) 32%	0.45 3.8%	3.3 28%	7.3 (4.0–13.5) 61%	0.48 (0.10–1.3) 4.8%	0.30 (0.07–0.75) 2.5%	0.05 (0–0.20) .04%
1 yr	11.4 (6.0-17.5)	3.5 (1.5–8.5) 31%	0.35 3.1%	3.2 28%	7.0 (4.0–10.5) 61%	0.55 (0.05–1.1) 4.8%	0.30 (0.05–0.70) 2.6%	0.05 (0–0.20) 0.4%
6 yr	8.5 (5.0–14.5)	4.3 (1.5–8.0) 51%	0.25 (0–1.0) 3.0%	4.0 (1.5–7.0) 48%	3.5 (1.5–7.0) 42%	0.40 (0–0.8) 4.7%	0.23 (0–0.65) 2.7%	0.05 (0–0.20) 0.6%
12 yr	8.0 (4.5–13.5)	4.4 (1.8–8.0) 55%	0.25 (0–1.0) 3.0%	4.2 (1.8–7.0) 52%	3.0 (1.2–6.0) 38%	0.35 (0–0.8) 4.4%	0.20 (0–0.55) 2.5%	0.04 (0–0.20) 0.5%
Adult	7.4 (4.5–11.0)	4.4 (1.8–7.7) 59%	0.22 (0–0.7) 3.0%	4.2 (1.8–7.0) 56%	2.5 (1.0–4.8) 34%	0.30 (0–0.8) 4.0%	0.20 (0–0.45) 2.7%	0.04 (0–0.20) 0.5%

Data presented as mean (±2 S.D.)
WBC, white blood cell.
Modified from Combleet J, Astanita R, Wolf PL. White blood cell and platelet disorders. In: Howanitz JH, Howanitz PJ, eds. *Laboratory medicine—test selection and interpretation*. New York: Churchill-Livingstone, 1991:553–618.

Hematocrit

The hematocrit is simply the ratio of the volume of the red blood cells to the volume of the whole blood. Although hematocrits can be determined directly by centrifugation ("spun" hematocrits), most hematocrits are now calculated directly from the RBC and MCV: hematocrit = RBC (cells/L) × MCV (liter/cell) (Table 40.5). Thus, any factor leading to spurious errors in RBC or MCV determination will also lead to spurious hematocrit results. Because the hematocrit is thus derived from two separate factors and all their sources of error, it should be clear that the hematocrit is a less accurate measure of anemia than the direct determination of hemoglobin concentration.

Manual or spun methods for determining hematocrits are still used in some circumstances and are based on centrifuging capillary blood tubes at 10,000 to 15,000*g* to separate the fluid and cell components of the blood. Spun hematocrits may give spuriously high results if a significant concentration of plasma becomes trapped in the red cell layer. This phenomenon can be found in samples with polycythemia, macrocytosis, spherocytosis, hypochromic anemias, and red blood cell fragment disorders (such as sickle cell anemia or burn patients). Improper mixing and excessive dilution with anticoagulant can also lead to inaccurate spun hematocrit levels.

Red Blood Cell Indices

The red blood cell indices are a key component of the CBC. These include the MCV, MCH, and MCHC, which are used to determine the size, cellular hemoglobin, and hemoglobin concentration of the red blood cells, respectively. The MCV is the most useful of the red blood cell indices and is especially important in classifying anemias as either normocytic (normal MCV), microcytic (decreased MCV), or macrocytic (increased MCV) (see Chapter 41 for more details). The MCV is determined by the distribution of the red blood cell histogram (Fig. 40.4). The mean of the red blood cell distribution histogram, based on electrical impedance, is the MCV and the coefficient of variation (or sometimes the standard deviation) is the RDW (discussed later).

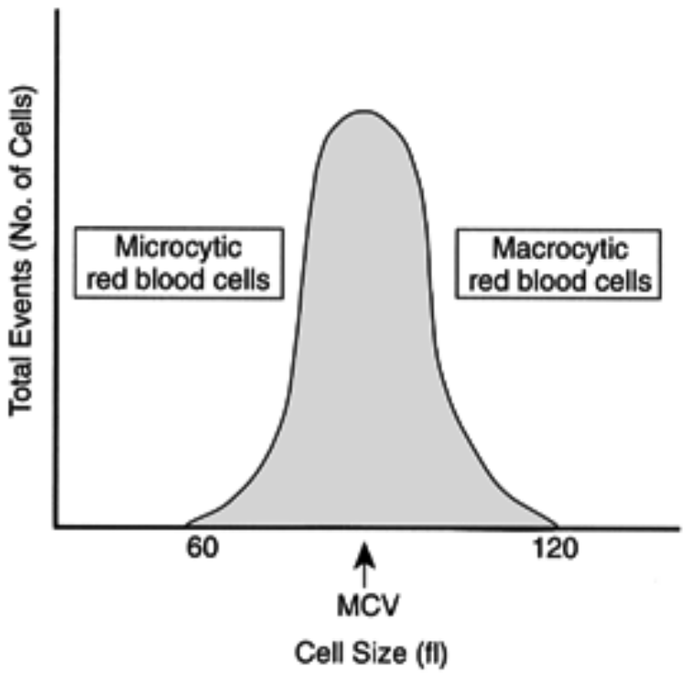


FIGURE 40.4. Red cell distribution histogram from an automated blood analyzer. The number of events counted (*y* axis) is plotted versus the cell size or volume (*x* axis). The median value of the histogram distribution is the mean corpuscular volume (MCV) of the sample. The coefficient of variation (or in some cases, the standard deviation) is equivalent to the red cell distribution width. Microcytic red blood cells, large platelets, and debris fall to the left portion of the curve, and macrocytic red blood cells, and small leukocytes fall to the right.

The MCH is the hemoglobin concentration per cell and is calculated from the hemoglobin and RBC (Table 40.5). It is of little use clinically and typically follows the movement of the MCV value. The MCHC is the average hemoglobin concentration per total red blood cell volume (Table 40.5) and is of minimal diagnostic use. A decrease in MCHC provides little clinical information and may not be noted until, for example, late in the course of an iron deficiency anemia. An increased MCHC is indicative of instrumentation or specimen problems and is a useful tool in the quality-control process. When MCHCs (hemoglobin/hematocrit) were determined manually (before automated instruments were widely used), a low MCHC was more commonly encountered in iron deficiency. This erroneous finding was the result of spuriously elevated spun hematocrits with the plasma trapping seen in microcytic anemias. Thus, the MCV is the only one of its red blood cell indices that provides consistently

TABLE 40.5. COMPONENTS OF THE COMPLETE BLOOD COUNT

CBC	Traditional Units	International Units
Hemoglobin	g/dL	g/L
Hematocrit	dL/dL (%)	L/L (%)
Red blood cells (RBC)	10^6 cells/μL	10^{12} cells/L
White blood cells	10^3 cells/μL	10^9 cells/L
Platelet count	10^3 plt/μL	10^9 platelets/L
Mean corpuscular volume (MCV)	fl/cell	fl (10^{-15}L/cell)
Mean corpuscular hemoglobin (MCH)	pg/cell	pg (10^{-12}g)/cell
Mean corpuscular hemoglobin concentration (MCHC)	g/dL	g/L
Derived CBC Values		
Hematocrit (Hct) (L/L;%) = [MCV (in liters per cell)] × [RBC (in cells/L)]		
MCH (pg/cell) = [HB (in g/L)] / [RBC (in cells/L)]		
MCHC (g/dL) = [Hb (in g/dL)] / [Hct (in L/L)]		

CBC, Complete blood count.

useful clinical information, with the MCH and MCHC useful primarily for quality-control purposes.

Red Cell Distribution Width

The RDW provides some insight and quantification into the variation in red cell size, or anisocytosis. It is derived from the red blood cell histogram and represents the coefficient of variation (or, in some cases, the standard deviation) of the red cell histogram distribution curve (Fig. 40.4). Bessman et al. (2) proposed a system of classifying anemias based on RDW and suggested that it may be a more sensitive indicator of a change in cell size than purely the MCV (Table 40.6). In general, an elevated RDW has been associated with anemias from various deficiencies such as iron, vitamin B_{12}, or folate. This is to be contrasted with the normal RDW that classically characterizes the microcytic anemias seen in thalassemia. However, as in most laboratory tests, this single red blood cell parameter has been shown to have more than its share of false positives and false negatives relative to this classification scheme. Indeed, some cases of thalassemia may have an elevated RDW, and cases with anemia of chronic disease that have slightly lower MCVs appear to have slightly elevated RDWs; myelodysplastic syndromes and some malignancies may also be associated with elevated RDWs. Thus, the RDW must be interpreted in conjunction with other CBC data in interpreting the abnormal hemogram.

TABLE 40.6. UTILITY OF RDW AND MCV IN CLASSIFICATION OF ANEMIA

MCV	RDW Normal	RDW High
Low	Thalassemia trait	Iron deficiency
	Transfusion	S-ß-thalassemia
	Chemotherapy	Hemoglobin H
	Malignancy	RBC fragmentations
	Hemorrhage	
	Hereditary spherocytosis	
	Posttraumatic splenectomy	
Normal	Normal	Transfusion
	Chronic disease	Early iron, B_{12}, or folate deficiency
		Homozygous hemoglobinopathy
		Myelofibrosis
		Sideroblastic anemia
High	Aplastic anemia	Folate deficiency
	Liver disease	B_{12} deficiency
		Cold agglutinin
		Hemolytic anemia
		Chemotherapy

Modified from Bessman JD, Gilmer PR Jr, Gardner FH. Improved classification of anemias by MCV and RDW. *Am J Clin Pathol* 1983;80:322–326.

RDW, red cell distribution width; MCV, mean corpuscular volume; RBC, red blood cell.

WBC—Total

The total WBC (or leukocyte count) consists of the PMNs (segmented and band forms), lymphocytes, monocytes, eosinophils, basophils, and potentially any other circulating hematopoietic cells. Historically, nucleated red blood cells are not included and are typically subtracted from the total WBC. WBCs are performed on EDTA-anticoagulated blood. With automated hematology counters, the total WBC, like the total RBC, is determined by either electrical impedance methods or light-scatter techniques. Hemacytometers are not used routinely in hematology laboratories for the WBC but may be used if the automated counters fail to provide accurate results, such as in leukopenic or leukemic patients. In either method, the red blood cells are lysed before the determination of the WBC.

The WBC can provide important diagnostic and monitoring information in patients with primary hematologic disease or acute/chronic infectious processes. The WBC also provides appropriate therapeutic data after administration of chemotherapy, radiation therapy, or antimicrobial agents. Because the WBC is composed predominantly of PMNs, any change in the total WBC typically reflects a change in the total PMN count. However, the WBC by itself is neither highly sensitive nor specific for an acute or chronic infectious process. In other words, an elevated WBC should not be considered as evidence of an infectious process unless there is clinical evidence of such and, conversely, a low WBC does not rule out an infectious process when clinical findings favor an infectious process. Other causes of elevated WBC include trauma, surgery (24 to 48 hours after), hemorrhage, delivery, tissue necrosis, and corticosteroids and other medications.

Heparinized blood should not be used for determining the WBC because of its interference with lysing reagents and its association with platelet agglutination in some specimens. As mentioned previously, citrate anticoagulant may be useful in cases with EDTA-dependent platelet agglutination; a correction factor of 1.1 is needed to account for the extra volume of anticoagulant used.

With the electrical impedance method, the red blood cells are first lysed and the remaining white blood cells pass through a narrow aperture, one cell at a time, which changes the electrical current across that opening. The change in current is proportional to the cell number and cell size. As with the total RBC, light-scatter characteristics may be used to determine the WBC as the light scatter relates to cell size and cellular content. As with any CBC component, various clinical and specimen characteristics may lead to spurious WBC results. Nucleated red blood cells, cryoglobulin, platelet clumps, large platelets, and unlysed red blood cells may all lead to false elevations of the total WBC. As mentioned previously, nucleated red blood cells are usually excluded from the corrected WBC; the following formula is used: corrected WBC = (measured WBC × 100)/[100 + (*n* red cells/100 white cells)].

White Blood Cell Count—Differential

The differential WBC (or differential leukocyte count) remains one of the most frequently performed tests in the clinical hematology laboratory. Until the early 1970s, the only accepted method for performing a differential count was microscopic examination of a Romanowsky-stained peripheral blood smear. Traditionally, the differential WBC includes the counting and categorization of typically 100 white blood cells based on morphologic criteria, with the results being expressed as a percentage of each cell type identified. Absolute concentrations of the various cell types are calculated by multiplying the percentage by the total WBC. In addition to the differential count, morphologic evaluation of the smear provides the opportunity to evaluate morphologically all components of the peripheral blood, including red blood cells, white blood cells, and platelets. This is a key advantage of the manual differential in that it allows the detection of a variety of disorders that might otherwise be lost in a totally automated system.

The manual performance of a differential count, i.e., the manual differential, is a time-consuming, labor-intensive, and relatively expensive procedure in this era of technological advances and automation. In addition to these technical and administrative disadvantages, the traditional manual differential count has other medical and scientific limitations, including poor sensitivity, specificity, and predictive value; it is imprecise because of sampling error and statistical probabilities and is prone to subjective judgmental errors. Nonetheless, the manual differential count has remained the gold standard of differential WBCs.

To interpret accurately the morphologic features of a peripheral blood film, several technical steps are required. First, the film must be well made with a good “feathered” edge and preferably with space around all slide edges. There should be a uniform distribution of cells and uniformly distributed Wright's or Wright-Giemsa stain. Although these basic requirements seem rather obvious, they can be easily forgotten in the laboratory, which can lead to inaccurate differential counts.

The area behind the feather edge of the blood film is the site of accurate counting. In this area, the red cells should show typical central pallor without being clumped. The smear is examined in a logical pattern, moving from one edge to another. Typically, 100 consecutive white blood cells are identified and classified by the morphologic reviewer. Other morphologic abnormalities are also noted during this process.

It is important to realize the statistical impact that a 100-cell count has on overall differential accuracy. The greater the number of cells counted, the greater the statistical precision of a differential count. Table 40.7 illustrates the 95% confidence interval limits for various blood cell percentages as determined by the total cell count obtained. Basically, the fewer cells that are counted, the less precise the blood cell percentage is. This is especially true with the cells that account for relatively lower percentages, as would be seen with neutrophil bands, monocytes, eosinophils, and basophils. For example, a band count of 5% obtained at one time would not represent a statistically significant change in the percentage of bands compared with a 10% count the following day if a 100-cell differential count is performed. This can be contrasted to obtaining either a 500- or 1,000-cell differential count in which a change from 5% to 10% would be considered statistically significant. In addition to this statistical variation dependent on total cell counts obtained, one must keep in mind the large number of errors that can result from inadequately collected or processed blood samples, poor-quality blood films, inconsistent staining, or judgmental errors by the technologist or physician performing the differential count. All these components emphasize the deficiencies of the manual differential count. Thus, the historical reliance on the manual differential count as the gold standard of the hematology laboratory must be viewed skeptically, considering the major sources of error that can potentially contribute to the determination of the differential WBC.

TABLE 40.7. CONFIDENCE LIMITS (95%) FOR PERCENTAGES OF CELLS REPORTED IN A MANUAL DIFFERENTIAL COUNT

% Reported	*n* = 100	*n* = 1,000	*n* = 10,000
0	0.0–3.6	0.0–0.4	0.0–0.1
1	0.0–5.4	0.5–1.8	0.8–1.3
2	0.2–7.0	1.2–3.1	1.7–2.3
3	0.6–8.5	2.0–4.3	2.6–3.4
4	1.1–9.9	2.9–5.4	3.6–4.5
5	1.6–11.3	3.7–6.5	4.5–5.5
6	2.2–12.6	4.6–7.7	5.5–6.5
7	2.9–13.9	5.5–8.8	6.5–7.6
8	3.5–15.2	6.4–9.9	7.4–8.6
9	4.2–16.4	7.3–10.9	8.4–9.6
10	4.9–17.6	8.2–12.0	9.4–10.7
15	8.6–23.5	12.8–17.4	14.3–15.8
20	12.7–29.2	17.6–22.6	19.2–20.8
25	16.9–34.7	22.3–27.8	24.1–25.9
30	21.2–40.0	27.2–32.9	29.1–31.0
35	25.7–45.2	32.0–38.0	34.0–36.0
40	30.3–50.3	36.9–43.1	39.0–41.0
45	35.0–55.3	41.9–48.1	44.0–46.0
50	39.8–60.2	46.9–53.1	49.0–51.0
55	44.7–65.0	51.9–58.1	54.0–56.0
60	49.7–69.7	56.9–63.1	59.0–61.0
65	54.8–74.3	62.0–68.0	64.0–66.0
70	60.0–78.8	67.1–72.8	69.0–70.9
75	65.3–83.1	72.2–77.7	74.1–75.9
80	70.8–87.3	77.4–82.4	79.2–80.8
85	76.5–91.4	82.6–87.2	84.2–85.7
90	82.4–95.1	88.0–91.8	89.3–90.6
95	88.7–98.4	93.5–96.3	94.5–95.5
100	98.2–100.0	99.6–100.0	99.9–100.0

n, total cells counted.
Modified from Rümke CL, Bezemer PD, Kuik OJ. Normal values and least significant differences for differential leukocyte counts. *J Chron Dis* 1975;28:661–668.

Because of the significant limitations of the differential WBC, the clinical hematology laboratory has incorporated the use of numerous instruments with the capability to perform differential WBCs. Two basic methodologies have been used as automated differential counters in the clinical laboratory. These are digital

image analysis systems based on cell recognition from a stained blood smear and the determination of different leukocyte cellular characteristics that permit separation into subtypes by using various flow cell–related techniques. This latter methodology has become the standard in automated differential counts and is based on the use of either electrical impedance or light-scattering techniques.

Automated differential leukocyte counts as performed by modern hematology analyzers are more accurate, more precise, more economical, faster, and safer than traditional manual differential counts. Although the goal of any automated differential leukocyte counting system is to replace the manual differential count, this has not been fully accomplished in the clinical laboratory. The automated differential WBC is superior to the manual differential count in many aspects, but in some cases it fails to provide important morphologic detail that only the manual differential/review can provide. For example, red cell morphologic findings, categorization of immature granulocytes, and recognition of intracellular or extracellular organisms (e.g., malaria) can be accurately identified only on a manual smear. Thus, all modern clinical hematology laboratories have to deal with how an automated differential leukocyte count can be incorporated into the routine practice of that particular laboratory.

Most laboratorians would agree that the initial differential count for every patient, whether hospitalized or ambulatory, should be performed by automated methods. The key to successful implementation is based on the ability of the instrument to recognize both quantitative and qualitative abnormalities and to flag particular cases for further review. These flagged results may then lead to either a manual differential count or at least a manual review of the stained blood smear. Figure 40.5 demonstrates an approach to incorporating automated differential counts with manual reviews to best meet the patient's needs while still maintaining appropriate utilization of laboratory resources. It is the ability to recognize these qualitative abnormalities that determines the success of an automated differential count in a particular laboratory. The laboratorian must expect a false-negative rate that approaches 0%, while enduring a false-positive rate that is not too high. Obviously, the kinds of patients evaluated by a clinical hematology laboratory and the overall percentage of abnormal differential counts will also influence the degree of success of the automated differential. Thus, the type of instrument, the flagging criteria chosen, and the patient population served by the laboratory will determine the role of automated differential counts.

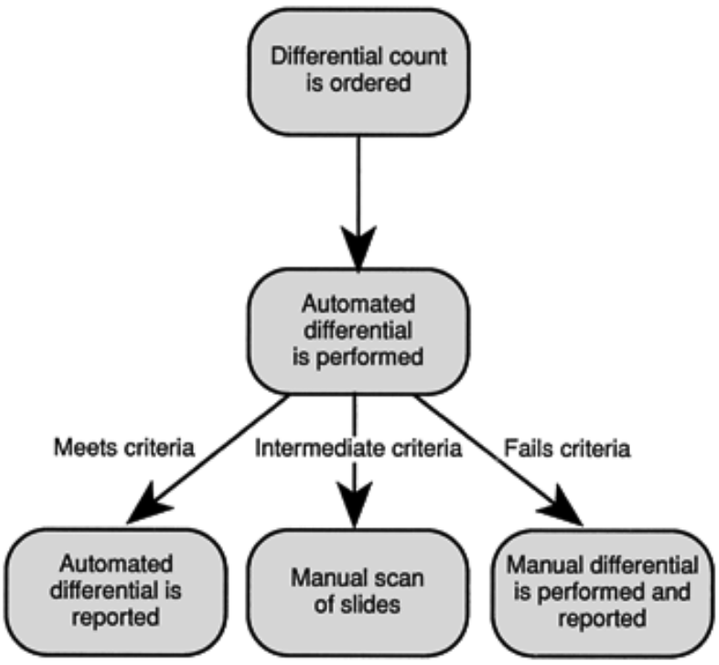


FIGURE 40.5. Schematic drawing shows the laboratory logistics of incorporating automated leukocyte differential counts, manual differential counts, and slide scanning.

Several hematology analyzers are available for the clinical laboratory. Each one approaches the issue of automated differentials in slightly different ways, but all use a single-cell flow apparatus that allows interrogation of each individual cell for a variety of potential characteristics. These characteristics permit separation of the various cell types. Technicon's H6000, H1, and H2 analyzers are flow cytometers that use myeloperoxidase cytochemistry coupled with light-scattering characteristics. Others, including Coulter's STKS, and TOA's Sysmex NE8000, use a combination of electrical impedance, low-frequency conductivity, and light-scattering characteristics to categorize white cells. Other newer analyzers, including the Cobas Argos 5 Diff by Roche, uses electrical impedance, cytochemistry, and optical absorbance,

and the CelDyn 3000 from Abbott uses a complex multiangle, light-scattering technology to categorize white cells. As mentioned previously, image analysis–based systems have not survived their introduction and are no longer being actively marketed to the clinical laboratories.

The blood cell histograms as produced by the modern automated hematology analyzers provide a wealth of information about the CBC and differential count. However, these scattergrams and histograms have not gained popularity among clinicians for various reasons. It is neither important nor necessary for clinicians to review routinely these histograms, but laboratorians may find the histograms invaluable for quality-control purposes as well as for evaluating difficult diagnostic cases and recognizing problem cases that need further review. Figure 40.6 provides examples of histograms from the Coulter STKS automated analyzer and depicts the typical distribution of the various peripheral blood components.

Additional parameters that can be generated by the Technicon analyzers include the mean peroxidase index, which is a measure of the myeloperoxidase staining intensity of neutrophils. Low levels of peroxidase index indicate a myeloperoxidase deficiency; the clinical importance of an elevated mean peroxidase index, however, has not been clarified. In addition, the Technicon also determines a lobularity index for neutrophils. A decreased lobularity index suggests less neutrophil lobulation and, therefore, a granulocytic left shift.

The appropriate use of the WBC differential is a serious concern for most hematology laboratories (Table 40.8). As with other laboratory tests, there should be different criteria for ordering differential WBCs for ambulatory patients as opposed to hospitalized patients. There is sufficient evidence to indicate that the differential WBC is not useful in the routine screening of ambulatory patients. Because the prevalence of hematologically manifested disease is low in this group of patients, significant clinical disease is typically not identified from the differential count unless it has also been suspected on other clinical grounds. Even when differential counts outside the normal range are identified in the ambulatory population, only rarely is a specific disease process diagnosed. Therefore, in the outpatient group, a differential WBC should be performed only in patients in whom the information may provide important diagnostic, prognostic, or therapeutic decisions.

In hospitalized patients, there are many clinical situations in which an abnormal differential count will correlate with a particular clinically important disease. Subsequent sections of this chapter outline the clinical situations associated with either cytopenias or cytoses of particular leukocyte components. As in the outpatient group, if no abnormality is suspected, a differential WBC will not be useful because of the low prevalence of disease in this particular group. If, however, an abnormality is suspected, then the differential WBC may provide some important information. For example, if there is a suspicion of a primary hematologic disorder, a differential count may be the primary means of making a specific diagnosis. Likewise, an unexpected leukocytosis or leukopenia found on a CBC may be more specifically elucidated if a leukocyte differential count is obtained. In patients with suspected infections or fevers, a differential count may be useful in documenting a neutrophilia if other data are inconclusive. It typically does not contribute diagnostic information if a leukocytosis has already been identified on a CBC, thus indicating a reactive process. Conversely, the differential WBC is seldom abnormal if the total WBC is normal, even if there is an ongoing infectious process. Thus, it should be recognized that a normal WBC and a differential WBC do not exclude the possibility of an infectious process.

A larger question that remains is the frequency at which to repeat a differential WBC in hospitalized patients. Repetitive testing may have diagnostic value if an infection has not shown improvement or if new symptoms evolve that are suspicious of either an infectious or hematologic disorder. Certainly, if management decisions are based on differential counts, for example, in a leukopenic or hematologic patient, then appropriate repetitive testing would be indicated. Overall, there has not been a consensus as to the appropriate testing interval for the differential count, but it would appear unlikely that daily differential counts in nonleukopenic patients would be indicated.

Platelet Count

A platelet count provides the starting point in the functional evaluation of the hemostatic system. An abnormality in the platelet count can lead to significant bleeding complications in a patient and may be indicative of a variety of underlying malignant or nonmalignant conditions. A diminished platelet count may be the result of either a marrow production problem or a peripheral destructive process. In cases in which megakaryopoiesis is decreased and platelet production diminished, evaluation of the bone marrow may reveal an infiltrative malignant process, whether metastatic or hematologic in nature. In addition, various drugs and some viral infections may lead to a reduction in platelet production. In patients receiving chemotherapeutic regimens, platelets are commonly diminished to very low levels. Peripheral destructive processes of platelets are common and are frequently entertained as part of a differential diagnosis. These are primarily immune-based thrombocytopenias but may occasionally involve splenic sequestration of platelets. Further discussion of the role of platelets in hemostasis is included in Section VIII of this book, Coagulation. This chapter does not deal with the platelet assays that are used to evaluate the bleeding patient. Rather, the laboratory's analysis of the total platelet count as obtained from hematology analyzers is described here.

EDTA is the preferred anticoagulant for platelet analysis and must be mixed thoroughly with the blood to avoid clotting and platelet clumping. If clotting does occur, a spuriously decreased platelet count may result. As discussed previously, in individuals with EDTA-dependent platelet agglutinin, citrate is the preferred alternative anticoagulant.

Today's automated hematology analyzers routinely obtain platelet counts over a wide range of values. However, manual hemacytometer counts are still essential in patients with low platelet counts, e.g., those patients with fewer than 50.0×10^9 platelets. Various red and white blood cells, platelet, and instrument artifacts may interfere with platelet counting. Newer instrumentation enhancements such as hydrodynamic focusing and pulse editing have helped to eliminate some of the crossover abnormalities seen in red cells and platelets. Some instruments,

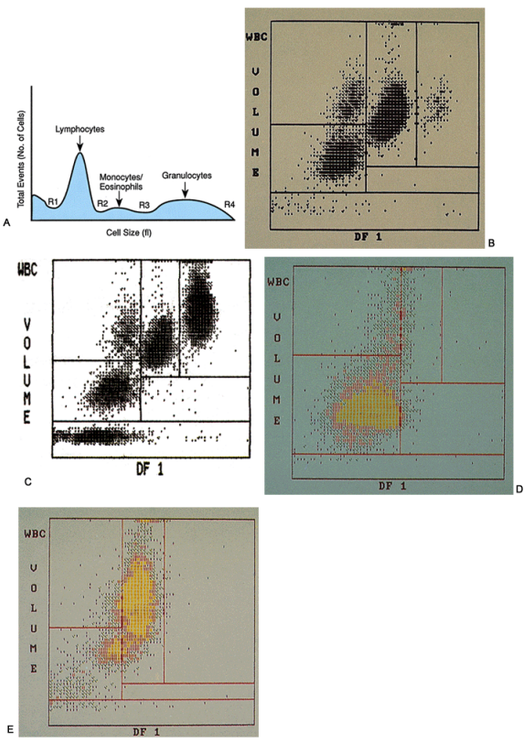


FIGURE 40.6. Blood cell histograms (Coulter STKS). **A:** Example of a white blood cell count (WBC) differential histogram based on volumetric studies. The various components of the WBC are shown on the graph. R1 to R4 correspond to flags generated by the instrument representing distribution abnormalities warranting manual observation. R1, nucleated red blood cells (RBCs), platelet clumps, large platelets, cryoglobulin, small lymphocytes, unlysed RBCs; R2, reactive lymphocytes, lymphoblasts, basophils, clotted sample; R3, eosinophilia, monocytosis, blasts, clotted sample; R4, granulocytosis. **B:** STKS—normal differential; **C:** STKS histogram—eosinophil (52%); **D:** STKS histogram—chronic lymphocytic leukemia; **E:** STKS histogram—acute myelogenous leukemia.

TABLE 40.8. INDICATIONS FOR ORDERING THE LEUKOCYTE DIFFERENTIAL COUNT

Indication	Usefulness
Ambulatory population	
General population (screening)	Not useful
Specific subgroups	Possibly useful (see below)
Hospital population	
No abnormality suspected	Not useful
Abnormality suspected	
Newly suspected infection or new fever	Useful in some patients
Suspicion of a primary hematologic disorder	Useful
Leukocytosis or leukopenia	Useful in some patients
Repeat tests	Useful in some patients at appropriate intervals

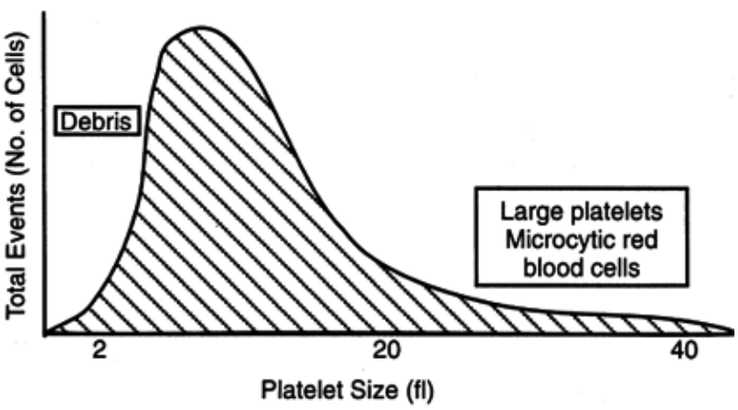


FIGURE 40.7. Platelet distribution histogram from 2 to 40 fl. The left portion of the curve may be contaminated with debris; the right portion may be contaminated with large platelets and microcytic red blood cells.

such as those manufactured by Coulter, count platelets within a particular size range, e.g., 2 to 20 fl, to eliminate any possible interference by red blood cells (Fig. 40.7). The instrument's computer then fits the results to a lognormal curve and extrapolates the curve to cover the normal range of platelet size of 0 to 70 fl.

Artifacts that can specifically interfere with platelet counting include red or white blood cell fragments/debris or electronic noise that can be found at the low end of platelet distribution. Interfering microcytic red blood cells or giant platelets may skew the distribution curve to the right. Platelet clumping is the most common cause of mistaken thrombocytopenia in the automated laboratory. Clumps of platelets will look to the instrument like large, single platelets or even white blood cells and will usually be excluded from the platelet distribution curve. In such cases, phase microscopy is necessary to obtain an accurate platelet count.

Some instruments will also calculate MPVs. The MPVs represent the mean volume (size) of platelets as determined by the platelet distribution histogram. Some studies indicated that the presence of large platelets, i.e., those with a high MPV, is suggestive of younger platelets that are found in peripheral destructive processes such as immune thrombocytopenias. The MPV may falsely increase or decrease with EDTA anticoagulation and therefore must be interpreted with caution. This inconsistency has limited its routine use. As discussed in the previous platelet section, the MPV will be unreliable if cytoplasmic fragments, electronic noise, giant platelets, or microcytic red blood cells artifactually interfere with platelet determination. Figure 40.8 demonstrates that the typical reference range of the MPV is inversely related to the platelet count. In other words, as the platelet count decreases, the MPV typically increases. The MPV is found to be high in patients with thrombocytopenia owing to idiopathic thrombocytopenic purpura (ITP) and other peripheral destructive processes, whereas patients with thrombocytopenia owing to marrow suppression typically have decreased MPV values (Table 40.9). Despite these possible uses, the clinical value of the MPV has not been clearly elucidated and must not be considered a gold standard in the evaluation of the thrombocytopenic patient.

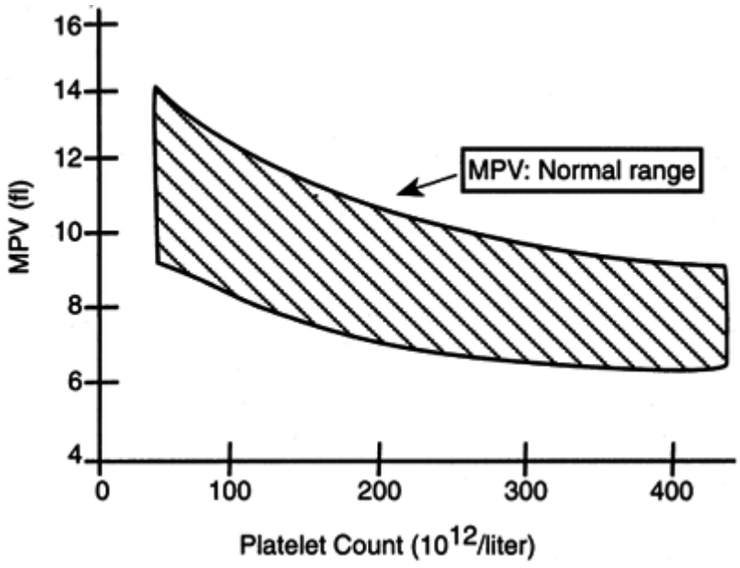


FIGURE 40.8. Mean platelet volume (MPV). The MPV varies inversely with total platelet count. Those samples having MPV values above the normal range for their total platelet count represent cases with adequate marrow production and, therefore, thrombocytopenia owing to peripheral destructive processes. Those samples with MPV values below the normal range for their total platelet count represent cases with probable marrow suppression and diminished platelet production.

TABLE 40.9. MEAN PLATELET VOLUME: CLINICAL CORRELATIONS

Low MPV	Normal to High MPV	High MPV
Marrow suppression	Hyperdestruction with marrow compensation	Hereditary disorders
Chemotherapy	Immune-related (ITP, drug-induced)	Bernard-Soulier syndrome
Megaloblastic anemia	Mechanical (consumptive coagulopathies, vasculitis)	May-Hegglin anomaly
Aplastic anemia		Miscellaneous
Marrow infiltration	Hemorrhage (major)	α-and ß-thalassemia trait (unknown cause)
Sepsis	Sepsis (without marrow suppression)	Myelodysplastic syndrome
Hypersplenism (variable)		Myeloproliferative disorders (in some cases)
Hereditary disorders Wiskott-Aldrich syndrome		

MPV, mean platelet volume.
Modified from Cornbleet PJ, Astanita R, Wolf PL. White blood cell and platelet disorders. In: Howanitz JH, Howanitz PJ, eds. Laboratory medicine—test selection and interpretation. New York: Churchill-Livingstone, 1991:553–618.

Reticulocyte Count

A reticulocyte can be broadly defined as an erythrocyte that still contains residual ribosomes and organelles but has ejected its nucleus from the previous orthochromatic erythroblast stage. Reticulocytes may take on various morphologic appearances depending on the amount of residual ribosomes and organelles, or reticulum (hence, the term reticulocyte). The definition of a reticulocyte from a technologist's perspective has classically been described as an erythrocyte containing a few (two or three) scattered dots. Although this may describe most of circulating reticulocytes, reticulocytes may rarely appear as diffusely granular cells. The life span of a reticulocyte has been estimated to be 3 to 4 days, but probably only 24 hours of the life span is spent in the peripheral circulation. Clinically, the reticulocyte percentage can be used as an indicator of erythropoiesis and is often utilized for evaluating patients with anemia. Although reticulocyte counts do not absolutely correlate with bone marrow erythroid activity,

the ease of obtaining and performing such a count has led to its common usage as a marker of red cell production. A normal or decreased reticulocyte count in moderate to marked anemia is strong evidence that the bone marrow is not responding in an appropriate fashion, as would be seen in iron, folate, or vitamin B_{12} deficiency, or as a result of a bone marrow infiltrative process. In contrast, an increased reticulocyte count generally reflects a rapid erythroid turnover, as would be seen in acute blood loss or acute or chronic hemolysis. In other words, the reticulocyte level can be used as a general indicator of bone marrow erythropoiesis and release.

Most current laboratory microscopic methods make the reticulocyte visible by precipitating the residual ribosomal RNA material with a dye such as new methylene blue or brilliant creosol blue; the precipitated RNA (or reticulum) will form small clumps that will stain with the dye and be visible as blue cytoplasmic dots or filaments (Fig. 40.9). Usually, a 1,000-erythrocyte count is performed manually through the microscope and the percentage of stained red blood cells derived. This manual determination of reticulocyte counts is a very imprecise method and open to subjective interpretation by the technologist. The imprecision and inaccuracy of reticulocyte counts have been well documented in several studies, with coefficients of variation ranging from 25% to more than 50%. These inaccuracies are probably related to (a) sampling error, (b) interobserver bias, (c) supervital staining of other cytoplasmic elements, such as nuclear debris, sidersomes, Heinz bodies, etc., and (d) variation in the quality of the smear and stain.

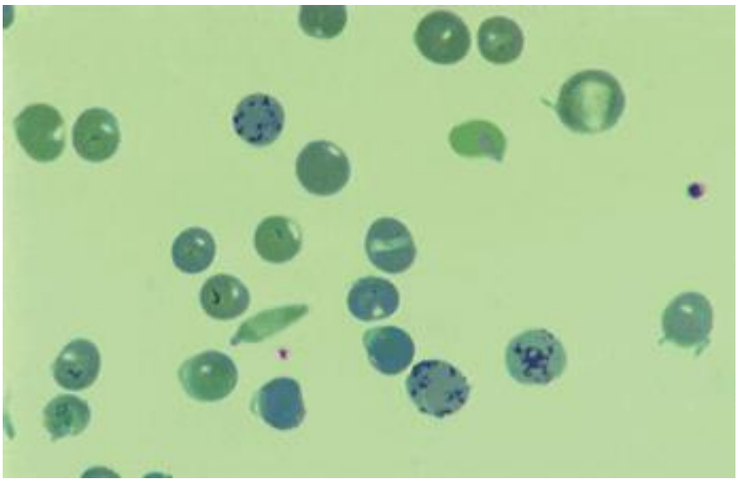


FIGURE 40.9. Peripheral blood smear with reticulocytes; staining is with new methylene blue dye. The blue granules represent precipitated, residual RNA.

Automated reticulocyte counting methods, such as image analysis and flow cytometry, have gained respectability and use in the clinical laboratories. Both these procedures remove much of the subjective interpretation involved in manual reticulocyte counting, allow evaluation of large numbers of red blood cells to avoid sampling error, and provide a standard and uniform analysis. Flow cytometric procedures depend on the binding of a suitable fluorescent dye to residual erythrocyte RNA. Although numerous dyes have been developed that accomplish RNA binding, not all have been successful in meeting the laboratory requirements for reticulocyte counting. To date, auramine O and thiazole orange are probably the two preferred fluorescent dyes for reticulocyte counting. Studies with both these fluorescent dyes have shown good to excellent correlation with results obtained from manual counts of cells and have consistently shown a tight reproducibility with duplicate samples. Flow cytometric detection methods offer many advantages over manual counting methods and will undoubtedly become the method of choice for reticulocyte analysis. Improved fluorescent dyes and increasing automation will continue to promote this transition.

Erythrocyte Sedimentation Rate

The ESR is one of the time-honored traditional tests in the hematology laboratory that has perhaps garnered too much importance in the clinical laboratory. In the laboratory, the ESR measures the distance a red blood cell falls in a vertical tube over a given period of time. The principle behind this process is the increased negative charge that the various inflammatory proteins (fibrinogen, α-, β-, and γ-globins) exact on the surrounding red blood cells. This promotes red blood cell separation and thus a more rapid fall of the red blood cell in the analysis tube. The resulting higher stack of measured red blood cells is interpreted as an elevated ESR.

Specimen collection and the anticoagulant used are crucial in the determination of the ESR. The Westergren method has become the standard procedure for determining the ESR (3). In this procedure, the blood specimen is anticoagulated with

sodium citrate and placed in a 30-cm glass tube with a 2.5-mm internal diameter. The modified Westergren procedure uses EDTA as the anticoagulant and may be the easiest method to use in the hematology laboratory because such samples can be used routinely for other hematologic parameters. The EDTA tube must be full and used within 2 hours if kept at room temperature; refrigerated samples may keep for as long as 12 hours if warmed before analysis.

The Wintrobe method has also historically been a popular method for ESR determination (4). The Wintrobe method also uses EDTA blood but requires only 1 mL of blood (as opposed to 2 mL with the Westergren procedure), uses a shorter but wider diameter tube than with the Westergren procedure, and may be preferred in the pediatric population because of the decreased amount of blood that is required.

Various factors, both clinical well as laboratory based, are known to disrupt the ESR. A slightly tilted ESR tube, high room temperature, or the use of heparin as an anticoagulant are all known to increase the ESR level. Conversely, a clotted blood sample, low room temperature, too short an ESR tube, or prolonged delay (longer than 2 hours if at room temperature) in analysis may all lead to a false decrease in ESR values. Clinical factors may also complicate the interpretation of the ESR. Anemia, hypercholesterolemia, chronic renal failure, as well as inflammatory disease may all produce an elevated ESR. Conversely, fragmented red blood cells (e.g., sickle cell anemia, burn patients), spherocytes, microcytic red blood cells, steroids, hypofibrinogenemia, and other miscellaneous entities may all lead to a decrease in the ESR value.

Clinically, an elevated ESR has been used as evidence for an inflammatory process. However, false-positive and false-negative results abound, and a wasteful hunt for a nonexistent underlying condition could result if only the ESR is relied on. The only consistent diagnostic use for the ESR, albeit begrudgingly, is in the diagnosis and monitoring of temporal arteritis and polymyalgia rheumatica. What then would constitute a valid indication for requesting an ESR? As with most laboratory tests, the ESR should not be used as a screening device in the healthy, asymptomatic population. No study has shown a significant contribution of an elevated ESR in detecting unsuspected disease in the asymptomatic patient. In the symptomatic patient, interpretation of ESR is complicated by the numerous factors that can lead to falsely elevated or falsely decreased values. Physicians often obtain an ESR in patients whose history and physical findings do not suggest any specific cause for their illness. Several studies showed that an ESR is not useful in those patients, such as in those for whom a diagnosis of a specific disease is not confirmed by other clinical data. In general, the ESR must be markedly elevated in such patients to be diagnostically useful, which is an exceedingly rare event in patients with no clear evidence of serious disease. A subsequent ESR several months later is the first step in evaluating an initial ESR elevation, rather than the physician instituting an extensive search for occult disease.

As previously indicated, the ESR is almost always increased in patients with temporal arteritis and polymyalgia rheumatica, and a normal ESR virtually excludes the diagnosis of temporal arteritis. The ESR may be useful in distinguishing inflammatory arthritic processes from other causes of joint symptoms, such as osteoarthritis. However, it is neither absolutely specific nor sensitive for rheumatoid arthritis and must be interpreted with great caution and obviously in conjunction with other clinical and laboratory findings. Some studies also suggested the role of ESR as an indicator of Hodgkin's disease, and other studies suggested a possible role as a marker for relapse of disease. Patients with a markedly elevated ESR greater than 100 mm/h usually have underlying malignancy, acute infection, or some type of connective tissue disease. However, unexplained elevations in the ESR less than this amount probably do not warrant evaluation and most do not have any associated disease process.

In summary, the ESR is a laboratory test that has remained essentially unchanged over its 75-year laboratory history. As such, its role in clinical diagnosis should be tempered with the reality that it is associated with numerous false-positive and false-negative results. A more judicial use of this screening test is definitely warranted.

BONE MARROW EXAMINATION

Part of "40 - Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders"

Components

Examination of the bone marrow ideally consists of evaluating three major components (Table 40.10): (a) the peripheral blood, (b) the bone marrow aspirate smear, and (c) a bone marrow tissue section (trephine biopsy and/or aspirate clot). Not all three components may contribute to a diagnosis in a given case, but each may potentially provide important diagnostic information and must be a part of the total examination process. Examination of the peripheral blood includes reviewing the CBC, differential count, reticulocyte percentage, and blood smear morphology. The bone marrow aspirate provides useful information for various studies, including (a) morphology and cytochemical stains, (b) flow cytometric immunophenotyping, (c) cytogenetic analysis, (d) molecular biological studies, and (e) microbiological

studies, if needed. Bone marrow tissue sections provide important information regarding the overall cellular distribution and marrow architecture and may consist of either a decalcified trephine biopsy or a clotted aspirate specimen (Fig. 40.10). Miscellaneous laboratory studies are also essential in interpreting a bone marrow study. These may include studies to evaluate red cell problems such as iron studies, B_{12}/folate levels, hemoglobin electrophoresis, and Coombs' test; lactate dehydrogenase, serum/urine immunoelectrophoresis, microbiological or virologic culture results, and other pertinent laboratory tests that may also provide important laboratory data needed to evaluate properly a bone marrow study.

TABLE 40.10. BONE MARROW EXAMINATION

Peripheral blood
CBC
Differential count
Reticulocyte count
Blood smear morphology
Bone marrow aspirate
Morphology
Cytochemical stains
Flow cytometric immunophenotyping
Cytogenetic analysis
Molecular biological studies
Bone marrow tissue section
Trephine biopsy
Clot section
Laboratory studies—miscellaneous
Iron studies
B_{12}/folate
Hemoglobin electrophoresis
Immunoelectrophoresis
LDH
Coomb's test
Microbiology culture/serology results
Miscellaneous

CBC, Complete blood count; LDH, lactate dehydrogenase.

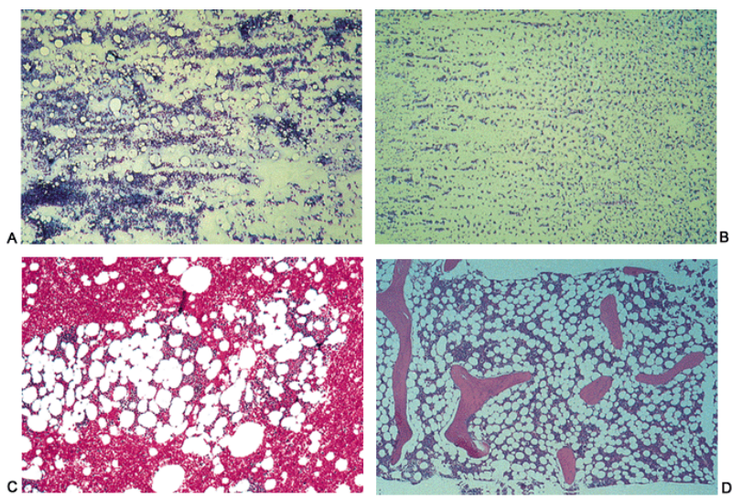


FIGURE 40.10. Examples of bone marrow preparation samples. **A**, Bone marrow aspirate spicule; **B**, bone marrow aspirate buffy coat smear; **C**, bone marrow aspirate clot section fixed in formalin; **D**, bone marrow biopsy section fixed in B5 fixative.

Indications for Bone Marrow Evaluation

Indications for examining a bone marrow are outlined in Table 40.11. The most common reason is for the evaluation of a known or suspected malignancy. A bone marrow study is definitely mandated if a hematologic malignancy is suspected. This suspicion may arise from either clinical or physical examination findings (e.g., splenomegaly or lymphadenopathy) or if abnormal or immature circulating cells are found in the peripheral blood smear. The bone marrow may also be involved in metastatic malignancies of nonhematopoietic origin. Evaluation of patients with either non-Hodgkin's lymphomas or Hodgkin's disease usually indicates that a bone marrow study be performed to adequately stage the patient for the presence of lymphoma. Bone marrow studies are also essential in monitoring the effect of therapy for patients with hematologic malignancies, as well as monitoring the effectiveness of bone marrow transplantation.

In nonspecialized institutions and practices, bone marrow studies are frequently required to evaluate patients with unexplained cytopenias. Patients who have unexplained thrombocytopenias or neutropenias can benefit by having a bone marrow

aspirate and biopsy performed to evaluate fully whether the cytopenia is a reflection of decreased marrow production or, by inference, increased peripheral destruction. Bone marrow studies for the evaluation of an anemia are not indicated as a primary procedure unless other clinical, peripheral blood morphology and biochemical studies failed to reveal the etiology of the anemia. Other indications for a bone marrow biopsy include obtaining tissue for culture in patients with fever of unknown origin.

Interpretation of Bone Marrow Studies

Bone marrow cellularity is best determined on the bone marrow tissue sections but can be estimated from the bone marrow aspirate smear. Marrow cellularity varies with age and is expressed as an estimate of the percentage of the bone marrow area occupied by cells to the total area of cells plus background elements and fat. The normal cellularity decreases with age and for a newborn is approximately 75% to 100%, 50% to 90% in the adolescent, 30% to 80% in the adult, and 20% to 50% in the elderly (older than age 65) (Fig. 40.11). Determination of cellularity is useful when evaluating patients with unexplained cytopenias; for example, increased cellularity in a patient with a cytopenia would suggest ineffective hematopoiesis (i.e., decreased cellular release) or increased peripheral destruction (Fig. 40.12). In contrast, decreased cellularity in an individual with a particular cytopenia would suggest a marrow production problem (Fig. 40.12).

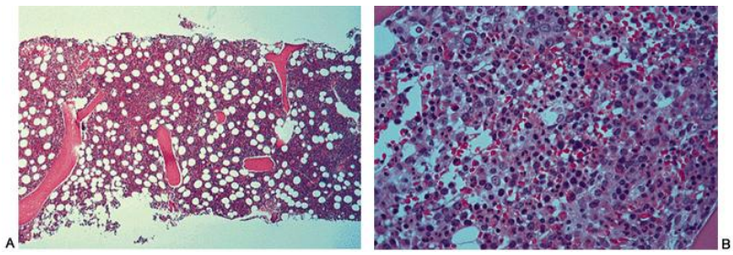


FIGURE 40.11. Bone marrow biopsy sections demonstrate normal cellularity. **A:** Approximately 40% to 50% cellularity in an otherwise healthy 60-year-old man; **B:** virtually 100% cellular marrow from a newborn boy.

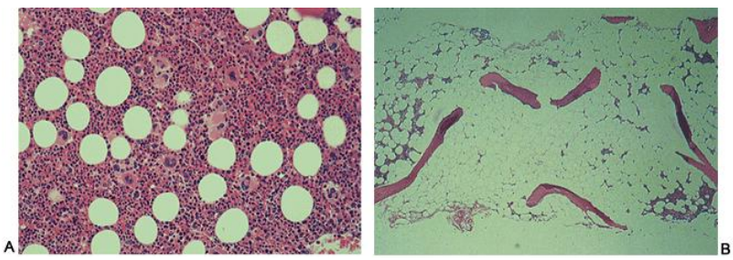


FIGURE 40.12. Bone marrow biopsy sections. **A:** Megakaryocytic hypercellularity from an individual with immune thrombocytopenia purpura; bone marrow from an individual with granulocytic hyperplasia owing to Kostmann syndrome. **B:** Markedly hypocellular bone marrow after chemotherapy in an individual with acute lymphoblastic leukemia.

The next step in evaluating bone marrow is determining the distribution of cells within the marrow. The adequacy of megakaryocyte production is easily estimated by evaluating either the marrow biopsy, aspirate clot sections, or an aspirate spicule. Typically, one to two megakaryocytes can be seen on a 40× or 50× objective field. The determination of increased or decreased numbers of megakaryocytes relative to the overall platelet count can suggest that a production or destruction problem may exist.

The next step in the evaluation of bone marrow is assessing myeloid and erythroid development both quantitatively and qualitatively. There should be a normal distribution of all cell types within each of the granulocytic and erythroid lines, indicating a normal maturational process. A 500-cell differential count of the bone marrow aspirate is useful in quantifying the

TABLE 40.11. INDICATION FOR BONE MARROW EVALUATION

Evaluation for malignancy
Primary hematologic disorder
Staging for Hodgkin's disease
Staging for non-Hodgkin's lymphoma
Metastatic tumor (nonhematopoietic)
Monitoring therapy
Postchemotherapy
Postbone marrow transplantation
Evaluation of cytopenias
Marrow production problem
Peripheral destruction
Inadequate/ineffective marrow release
Miscellaneous
Culture (fever of unknown origin)

amount of granulocytic and erythroid precursors within the marrow (Table 40.12). This myeloid:erythroid ratio ranges from 2:1 to 4:1 in the adult and is useful in determining whether there are relative hyper- or hypoplasias. Again, this may provide useful information in evaluating a patient with a neutropenia or anemia. In addition to quantitative abnormalities, qualitative abnormalities of myeloid development must also be assessed. Dysplastic changes in erythroid or granulocytic cells is not an uncommon finding and must be looked for when evaluating a bone marrow aspirate.

TABLE 40.12. BONE MARROW DIFFERENTIAL CELL COUNTS: NORMAL VALUES

Bone Marrow Cells	Childhood (%)	Adult (%)
Normoblasts (total)	23.1	21.5 (14.2–30.4)
Pronormoblasts	0.5 (0.0–1.5)	0.6 (0.2–1.4)
Basophilic	1.7 (0.2–4.8)	2.0 (0.7–3.7)
Polychromatophilic	18.2 (4.8–34.0)	12.4 (12.2–24.2)
Orthochromatic	2.7 (0.0–7.8)	6.5 (2.0–22.7)
Granulocytes (total)	57.1	56.0 (45.1–66.5)
Myeloblasts	1.2 (0.0–3.2)	1.0 (0.5–1.8)
Promyelocytes	1.4 (0.0–4.0)	3.4 (2.6–4.6)
Myelocytes	18.3 (8.5–29.7)	11.9 (8.1–16.9)
Metamyelocytes	23.3 (14.0–34.2)	18.0 (9.8–25.3)
Bands	?	11.0 (8.5–20.8)
Segmented	12.9 (4.5–29.0)	10.7 (8.0–16.0)
Lymphocytes	16.0 (4.8–35.8)	15.8 (10.8–22.7)
Monocytes		1.8 (0.2–2.8)
Eosinophils	3.6 (1.0–9.0)	3.2 (1.2–6.2)
Basophils	0.06 (0.0–0.8)	<0.1 (0.0–0.2)
Plasma cells	0.4 (0.2–0.6)	1.8 (0.2–2.2)
M:E ratio	2.9 (1.2–5.2)	2.5 (1.2–5.0)

M : E, myeloid : erythroid.
Data from Mauer AM. *Pediatric hematology*. New York: McGraw-Hill, 1969; and Jandl JH. *Blood—textbook of hematology*. Boston: Little, Brown, 1987.

Other cell types and lesions may be identified within the bone marrow and must be assessed including the presence or absence of granuloma, necrosis, lymphoma, metastatic neoplasia, bone marrow fibrosis, lipid storage disease, benign lymphoid aggregates, and bone abnormalities such as osteosclerosis and osteoporosis. Also, a Prussian blue iron stain of the bone marrow aspirate and/or tissue section can provide useful information concerning a patient's iron stores (Fig. 40.13). Iron stains of a bone marrow trephine biopsy must be interpreted with caution. A decalcification may cause a false-negative interpretation; evaluation of the clot section for iron stores is probably superior to the trephine biopsy. In addition to iron stores, iron stains of the bone marrow aspirate smear need to be evaluated for the presence or absence of sideroblasts and, in particular, ringed sidero-blasts (Fig. 40.13); detection of sideroblastic iron cannot be reliably assessed on tissue section stains.

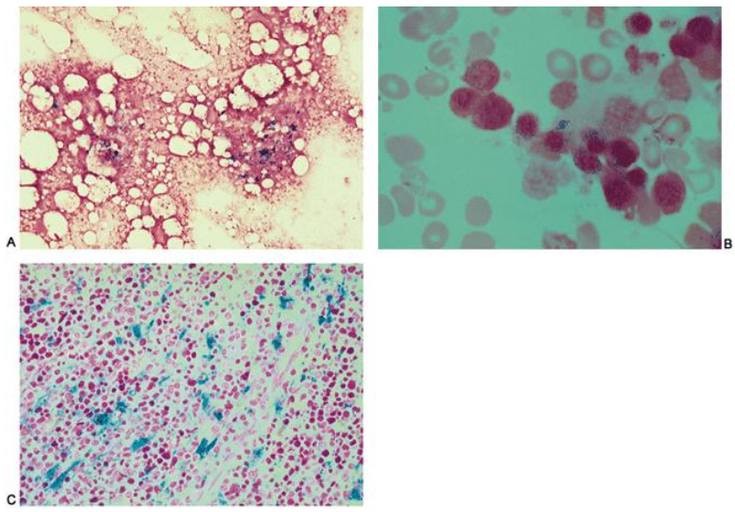


FIGURE 40.13. Iron stains of bone marrow. **A**, Bone marrow aspirate spicule shows iron stores; **B**, bone marrow aspirate smear shows sideroblastic iron (note the small green granules present in the cytoplasm); **C**, bone marrow aspirate clot section showing iron stores.

REACTIVE DISORDERS

Part of "40 - Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders"

Disorders of white blood cells include both malignant and nonmalignant diseases. The nonmalignant disorders of white blood cells may consist of quantitative abnormalities leading to an increase or decrease of a particular WBC component, benign morphologic

changes that have little if any effect on cellular function, and true functional disorders of a white blood cell type that may lead to significant patient morbidity and mortality. Although the peripheral blood is obviously the easiest vehicle by which to evaluate these nonmalignant abnormalities, the effects of these changes are truly reflected in the tissues in which they circulate, e.g., lymph nodes, spleen, or any other organ.

Neutrophilia

Neutrophils play an important role in the inflammatory response and are the primary cells involved in phagocytosis of foreign organisms. The number of circulating neutrophils is determined by multiple factors, including the rate of bone marrow production, the speed at which cells leave the marrow and enter the peripheral circulation, and changes in vascular margination. Clinical findings associated with neutrophilia are listed in Table 40.13. The most commonly encountered reason for neutrophilia is an acute bacterial infection. Other frequent causes include an inflammatory response to necrotic tissue (e.g., tumor necrosis), drugs (e.g., corticosteroids), and acute hemorrhagic episodes.

TABLE 40.13. DISORDERS ASSOCIATED WITH NEUTROPHILIA

Physiological neutrophilia
Neonates
Third-trimester pregnancy
Labor and delivery
Emotional stress
Exercise
Extreme cold or heat
Nausea and vomiting
Seizures
Infection
Bacterial
Fungal
Rickettsia
Parasite (rare)
Virus (rare/first 1 to 2 days)
Inflammation/tissue necrosis
Myocardial infarction
Tumor necrosis
Trauma
Surgery
Tissue infarction
Burns
Collagen vascular disorders
Dermatitis
Drugs/chemicals
Corticosteroids
Epinephrine
Digitalis
Etiocholanolone
Heparin
Lithium
Histamine
Endotoxin
Metabolic changes
Diabetic acidosis
Gout
Hyperthyroidism
Uremia
Eclampsia
Hematologic disorders
Acute hemorrhage
Hemolysis
Myeloproliferative disorders (see Chapter 45)
Cyclic neutrophilia
Hereditary/miscellaneous Idiopathic neutrophilia

Morphologic changes in neutrophils associated with acute infection include a shift toward the release of immature granulocytes (i.e., shift to the left), toxic granulation, Döhle bodies, and cytoplasmic vacuolation (Fig. 40.14). The left shift seen in patients with an acute infection consists of primarily neutrophil bands and metamyelocytes and occasionally more immature granulocytic precursors, such as myelocytes and promyelocytes; blast forms are only rarely seen. The quantitative assessment of neutrophil bands is a time-honored test within the hematology laboratory as an indicator of an acute infection. However, there are many problems in accurately determining neutrophil band counts, which mandates that some caution is necessary when interpreting band counts in isolation. Clinicians frequently use elevated band counts as evidence of acute infection and frequently may look at sequential changes in band counts as evidence of response or nonresponse to antiinfection therapy. However, this is fraught with inaccuracies for several reasons. The first is the poorly defined normal range of neutrophil bands in the peripheral blood. As gleaned from several reference laboratories and textbooks, the upper limit of normal for band counts in peripheral blood can range from 0.5×10^9/L to 1.8×10^9/L. This variation between laboratories is much greater than with any other peripheral blood component, such as total neutrophils, lymphocytes, and eosinophils.

The next problem is the inability to define a band accurately. Some have defined a band as any cell not having a well-defined filament between lobes, whereas others have required that a thick filament be present before a neutrophil band is identified (Fig. 40.15). In other words, there is a spectrum of cell appearances between the metamyelocyte and segmented neutrophil that prevents the band from being well defined. This inability to define a band accurately has led to extreme variability between technologists performing differential counts. Every laboratory has known “high banders,” “intermediate banders,” and “low banders.” National comparison studies of neutrophil band identification have shown coefficients of variations of greater than 100% between laboratories performing band counts. This lack of reproducibility also emphasizes the inaccuracy of differential counts performed by counting only 100 or 200 cells. The classic 95% competence limit table by Rümke et al. (5) (Table 40.7) demonstrates that at least 1,000 cells must be classified to get a reasonably accurate differential count. Thus, the actual range of what the true count may be of a low-percentage cell such as a neutrophil band is quite wide and leads to significant inaccuracy (Table 40.7). Any sequential use of band counts to evaluate therapeutic response is therefore inaccurate and wrought with clinical witchcraft.

There is also a significant amount of biological variability that will affect band counts. For example, it is well known that neutrophil counts may increase or decrease by several thousand, remain the same, or be quite variable from morning to evening among individuals. Again, this variability limits the accuracy of using sequential band counts in a meaningful fashion. There are conflicting reports in the literature as to the utility of band

counts in predicting acute infection. Some studies have shown that a “bandemia” is significantly associated with proven infection, whereas others have shown no correlation of band count to fever, infection, or tissue inflammation. In general, total leukocyte count and total neutrophil count provide more meaningful and consistent information in predicting an acute infection than the presence or absence of neutrophil bands. It is likely that a band count, which is carefully enumerated using clearly defined uniform criteria, is a useful indicator of acute inflammation. However, routine band counts performed by a clinical laboratory with multiple technologists during a 7-day workweek have little value in the evaluation of acute infection and should not be considered a reliable test.

Toxic granulation is another classic sign of an infectious neutrophilia (Fig. 40.14). This is characterized by coarse cytoplasmic granules that represent prominent primary granules still present from a “rush” through the normal maturation sequences. Toxic granulation may be difficult to identify with certainty unless

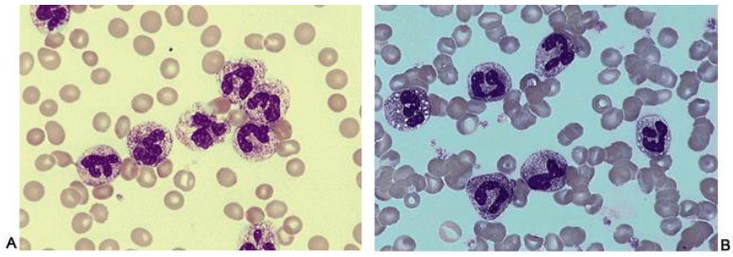


FIGURE 40.14. A, Toxic neutrophils with prominent azurophilic granules and cytoplasmic vacuolation; **B**, toxic neutrophils with cytoplasmic vacuolation.

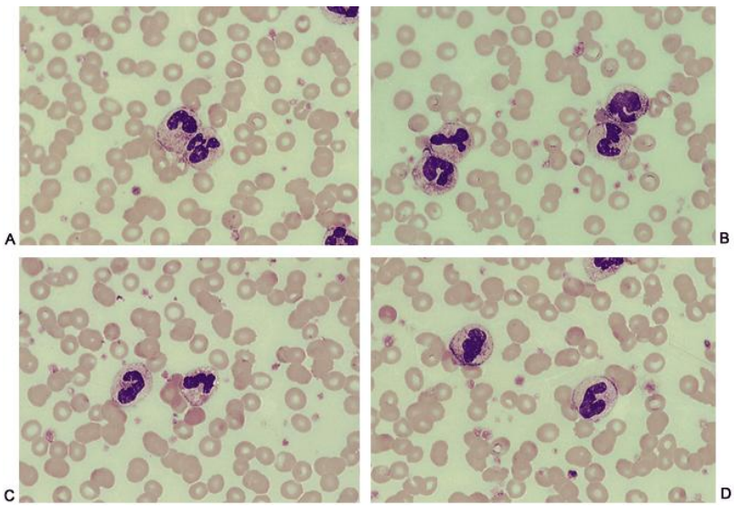


FIGURE 40.15. A–D: Examples of band neutrophils. Note the variation in nuclear appearance, complicating the exact definition of this cell type.

TABLE 40.14. DIFFERENTIAL DIAGNOSIS: LEUKEMOID REACTION VERSUS CHRONIC MYELOGENOUS LEUKEMIA

	Leukemoid Reaction	CML
WBC	10.0–100.0 × 10^9/L	30.0–500.0 × 10^9/L
Eosinophilia/Basophilia	None	Present
M:E ratio	5–10:1	>20:1
LAP	>100	<10
Organomegaly	Absent	Present

CML, chronic myelogenous leukemia; WBC, white blood cell count; M:E, myeloid : erythroid; LAP, leukocyte alkaline phosphatase.

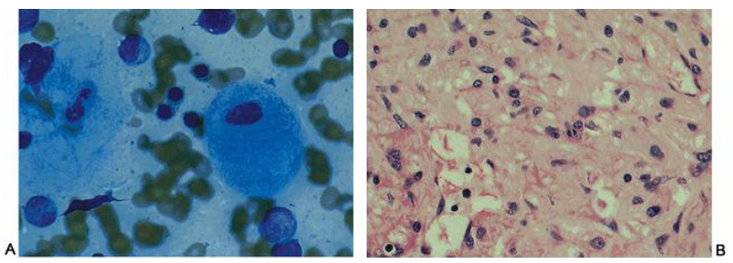


FIGURE 40.16. Gaucher's disease. **A**, Bone marrow aspirate shows macrophage with prominent blue cytoplasmic fibrils; **B**, bone marrow biopsy section shows classic accumulation of Gaucher cells having a fibular cytoplasm (tissue paper appearance).

appropriate peripheral blood smear staining is performed. A weak or faint Wright's stain may not accentuate toxic granulation; at the other extreme, laboratories not familiar with a Wright-Giemsa stain may have the tendency to overcall toxic granulation owing to the prominence of the granules seen with the Giemsa portion of the stain. Thus, laboratorians need to be familiar and confident with the stain that they are observing before diagnosing toxic granulation. Döhle bodies are a faint, pale blue cytoplasmic area near the periphery of a neutrophil that becomes prominent during an infectious episode (Fig. 40.14). These small bodies represent endoplasmic reticulum and are a very reliable sign of an infectious process. Again, Döhle bodies may be easily overlooked if appropriate staining is not performed. Cytoplasmic vacuolation is a less common feature of acute infection and is thought to represent cellular areas in which lysosomal ingredients have been released to engulf and destroy bacteria (Fig. 40.14). Cytoplasmic vacuolation caused by EDTA artifact can also occur and must be distinguished from vacuolation resulting from infectious process. The EDTA-associated vacuolation appears to be a result of the time spent in anticoagulant.

In most situations, the degree of neutrophilia in an acute infection seldom exceeds 50.0×10^9/L and is more typically in the 15.0 to 30.0×10^9/L range. In cases in which the neutrophilia exceeds 50.0×10^9/L, the term *leukemoid reaction* has been used. The obvious differential problem that arises in cases of leukemoid reaction is differentiation from chronic myelogenous leukemia (CML). Several features that will aid in the distinction of a leukemoid reaction from CML are listed in Table 40.14. Leukemoid reactions are typically characterized by morphologic signs of infections, such as a left shift, toxic granulation, Döhle bodies, and cytoplasmic vacuolation. The left shift in a leukemoid reaction is a gradation with greater prominence of segmented neutrophils, band forms, and metamyelocytes, whereas CML shows a prominence of both mature neutrophils and myelocytes. Leukemoid reactions lack dysplastic changes and basophilia and should only show rare blast forms at best. Laboratory tests that may be useful in this distinction include leukocyte alkaline phosphatase and cytogenetic karyotyping. The leukocyte alkaline phosphatase will be increased, and no Philadelphia chromosome will be found in a leukemoid reaction. Obviously, the clinical history and physical examination (e.g., presence or absence of splenomegaly) will also provide important diagnostic information.

Neutropenia

Neutropenia is defined as having counts less than 1.5 to 2.0×10^9/L. Neutropenia may result from either decreased marrow production or increased peripheral destruction of neutrophils. Causes of decreased neutrophil production in the bone marrow are vast and may include replacement of normal bone marrow elements by hematologic or metastatic malignancies, marrow fibrosis, or marrow necrosis. Chemotherapeutic or other toxic agents may also suppress normal marrow activity, including granulopoiesis; this group of toxins also includes alcohol and a variety of drugs and antibiotics (Table 40.15). Aplastic anemia, bone marrow irradiation, megaloblastic anemia (folate/vitamin

TABLE 40.15. NEUTROPENIA

Decreased marrow production
Agents that lead to bone marrow suppression
Chemotherapeutic agents
Radiation
Benzene
Chloroform
Alcohol
Arsenic
Drugs (nonchemotherapeutic)
Chloroamphenicol
Semisynthetic penicillins
Sulfonamides
Nitrofurantoin
Tricyclic antidepressants
Antithyroid drugs
Diuretics (thiazides)
Hypoglycemic agents
Quinidine, procanimide, alapurinol, and antihistamines
Bone marrow replacement
Hematologic malignancies
Metastatic malignancies
Myelofibrosis
Bone marrow necrosis
Storage disorders
Hematologic disorders with suppressed marrow production
Aplastic anemia
Paroxysmal nocturnal hemoglobinuria
Vitamin B_{12}/folate deficiency
Myelodysplastic syndromes
Chédiak-Higashi syndrome
Hereditary disorders
Fanconi's anemia
Familial cyclic neutropenias
Kostmann syndrome
Increased peripheral destruction/utilization
Hypersplenism
Collagen vascular diseases
Felty's syndrome
Cirrhosis
Immune-mediated
Antineutrophil antibody
Drug-associated antibody
Miscellaneous
Overwhelming infection (may also cause decreased production)
Especially in elderly, newborns, or patients with limited marrow reserve
Some viral infections
Miscellaneous
Pump-oxygenator in open heart surgery Hemodialysis

B_{12} deficiency), myelodysplastic syndromes, and congenital abnormalities of granulopoiesis may all be associated with neutropenia. The latter inherited disorders include Fanconi's anemia and familial benign and cyclic neutropenia (also known as Kostmann syndrome).

Increased destruction or utilization of neutrophils in the peripheral blood or tissue is common and may result from hypersplenism associated with several collagen vascular disorders, liver abnormalities, among others. Immune-mediated neutropenia owing to the presence of specific antineutrophil antibodies may also occur. Many viral infections, including human immunodeficiency virus (HIV), have been associated with neutropenia. Other particular causes of neutropenia are listed in Table 40.15. The neutropenia associated with viral disorders is thought to be the result of direct marrow damage by the infecting virus, which causes a disruption in normal myelopoiesis.

Usually in acute bacterial infection, significant neutrophilia will result. However, if the particular infection is overwhelming or if it occurs in an individual with borderline marrow reserves, the prolonged infection may deplete the marrow reserves and create more demand than what the marrow can respond to. This occurs primarily in the elderly, newborns, or any patient with an underlying myelosuppressive process. The resulting neutropenia will prevent an adequate response to the tissue infection and may lead to an overwhelming infection with subsequent death.

Monocytosis

Monocytosis is defined as an increase in peripheral blood monocytes greater than 0.8×10^9/L and is associated primarily with inflammatory and immune disorders. It is difficult to correlate monocytosis directly with specific disease states, as opposed to neutrophilia (Table 40.16). One of the classic associations of monocytosis is with tuberculosis infections. However, other infections such as subacute bacterial endocarditis, salmonellosis, listeriosis, syphilis, leprosy, and brucellosis can also be associated with a monocytosis. Monocytosis often appears during the recovery phase of an acute infection or after bone marrow suppression. This relative or absolute monocytosis heralds the marrow recovery and usually precedes the return of granulocytes. Monocytoses may also be associated with hematologic malignancies, such as myelodysplastic syndromes and some types of acute and chronic leukemia. Interestingly, as many as 25% of patients with Hodgkin's disease have been reported to have peripheral blood monocytosis. Monocytosis may also accompany many nonhematopoietic malignancies as well as some of the collagen vascular disorders. Thus, the finding of a patient with monocytosis is relatively nonspecific and can be associated with a variety of benign and malignant hematologic and nonhematologic disorders.

TABLE 40.16. DISORDERS ASSOCIATED WITH MONOCYTOSIS

Normal newborn infections
Tuberculosis
Syphilis
Leprosy
Salmonella
Brucellosis
Rickettsia
Subacute bacterial endocarditis
Parasites (some)
Marrow recovery phase
Acute infections
Neutrophil suppression
Hematologic disorders
Chronic myeloproliferative disorders
Myelodysplastic syndromes
Acute leukemias with monocytic component
Hodgkin's disease
Agranulocytosis
Collagen vascular disease
Gastrointestinal
Ulcerative colitis
Regional anuritis
Miscellaneous
Corticosteroids Lipid storage disorders

The mononuclear phagocyte system is also the primary cell affected in several storage disorders. These storage disorders are generally hereditary abnormalities or deficiencies of enzymes that are necessary for lipid storage and processing. The macrophages become the most obvious morphologic abnormality in these disorders as they become packed full of lipid material that cannot be further digested. Hematologic abnormalities, such as cytopenias, may result from bone marrows that are replaced with accumulated macrophages. Splenomegaly and hypersplenism may also result. Table 40.17 lists some of the hereditary disorders associated with storage diseases and the accompanying blood and bone marrow findings.

TABLE 40.17. STORAGE DISORDERS/BLOOD AND BONE MARROW FINDINGS

Blood	Bone Marrow	Storage Disorders
—	Gaucher cell	Gaucher's disease
Vacuolated lymphocyte		Tay-Sachs disease Batten-Spielmeyer-Vogt Glycogen storage diseases
Vacuolated lymphocyte	Foam cells	Niemann-Pick disease
Metachromatic inclusion	Metachromatic inclusion	Mucopolysaccharidosis Alder-Reilly anomaly

Gaucher's disease is a hereditary deficiency of glucocerebrosidase that results in an intracellular accumulation of the sphingoglycolipid glucocerebroside. Gaucher cells appear as enlarged macrophages within the bone marrow and are engorged with distinctive cytoplasmic fibriles that are meshed in an irregular pattern (Fig. 40.16). In bone marrow sections, Gaucher cells have been described as having a cytoplasmic appearance of wrinkled tissue paper. Pseudo-Gaucher cells have been commonly observed in patients with chronic myelogenous leukemia, representing increased glucocerebroside turnover from the massive leukocyte population in CML.

Pseudo-Gaucher cells have also been observed in patients with thalassemia and multiple myeloma. The pseudo-Gaucher cells of CML are thus different from true Gaucher cells in that they represent normally functioning macrophages overloaded with lipid byproducts, although they are morphologically very similar.

Niemann-Pick disease is a rare autosomal recessive disorder caused by deficiency of sphingomyelinase and results in accumulation of sphingomyelin in the mononuclear phagocytes. Numerous large, foamy macrophages are identified in the bone marrow of patients with Niemann-Pick disease (Fig. 40.17). These foamy cells are not specific for Niemann-Pick disease because such cells can be identified in other lipid disorders and various hematologic diseases.

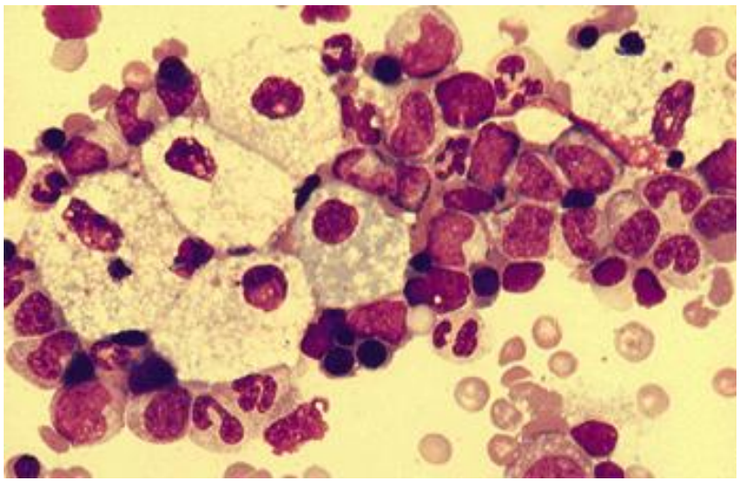


FIGURE 40.17. Niemann-Pick disease. Bone marrow aspirate smear shows large, foamy macrophages. These macrophages are not specific for Niemann-Pick disease and can be found in other diseases.

Eosinophilia

Eosinophils are primarily associated with anaphylactic responses and hypersensitivity reactions. Bacterial killing is not a major function of eosinophils, but eosinophils probably play a minimal role in antigen recognition. The cause of an eosinophilia in a patient can be quite difficult to determine. Benign eosinophilia is associated with a variety of drugs, parasitic infections, allergic reactions, and some collagen vascular diseases (Table 40.18). These must be differentiated from malignant disorders of eosinophils, such as hypereosinophilic syndrome.

TABLE 40.18. DISORDERS ASSOCIATED WITH EOSINOPHILIA

Drug reactions
Parasitic infestations
Trichinosis
Toxocariasis
Filariasis
Echinococcosis
Pneumocystis carinii
Aspergillosis
Coccidiomycosis
Strongyloidiasis
Ascaris
Schistosomiasis
Allergic reactions
Asthma
Dermatitis
Rhinitis
Graft rejection
Graft versus host disease
Pemphigus/pemphigoid
Farmer's lung
Collagen vascular diseases
Rheumatoid arthritis
Periarteritis nodosa
Hematologic disorders
Hodgkin's disease
Systemic mastocytosis
Chronic myelogenous leukemia
Pulmonary disease
Löffler's syndrome
Idiopathic/hypereosinophilic syndrome
Hypereosinophilic syndrome

Hypereosinophilic syndrome is a chronic myeloproliferative disorder characterized by a proliferation of mature eosinophils. The eosinophils frequently appear agranular and atypical. This disorder is characterized by splenomegaly, leukostatic problems and complications caused by the release of eosinophilic products in susceptible tissues (such as heart and brain).

Basophilia

Basophils are the least common of the peripheral blood cell components and typically number less than 0.1×10^9/L. CML is the most common and significant cause of basophilia. Benign

basophilia may be associated with hypersensitivity reactions, renal disease, myxedema, and some inflammatory responses. Other causes of basophilia are listed in Table 40.19.

TABLE 40.19. DISORDERS ASSOCIATED WITH BASOPHILIA

Hematologic malignancies
Chronic myeloproliferative diseases
Myelodysplastic syndromes
Mastocytosis
Acute basophilic leukemia (rare)
Hypersensitivity reactions
Hypothyroidism/myxedema
Ulcerative colitis
Radiation
Infections
Varicella
Smallpox
Miscellaneous
Renal disease (rare)

Monocytopenia/Eosinopenia/Basopenia

The determination of monocytopenia, eosinopenia, and basopenia can be made only if large numbers of cells are directly counted (Table 40.20); manual counts of 100 or 200 total cells are inaccurate for these low numbers. Eosinopenia and basopenia have both been reported to occur in situations of acute stress or acute infection, with corticosteroids of exogenous source, or with Cushing's syndrome. Monocytopenia has also been reported to occur after treatment with corticosteroids. Monocytopenia is also a well-described phenomenon in hairy cell leukemia.

TABLE 40.20. DISORDERS ASSOCIATED WITH MONOCYTOPENIA/EOSINOPENIA/BASOPENIA

Monocytopenia
Corticosteroids
Hairy cell leukemia
Eosinopenia
Acute stress
Acute infection
Corticosteroids
Cushing's syndrome
Basopenia
Acute stress
Acute infection
Corticosteroids
Cushing's syndrome

Reactive Lymphocytosis

Reactive lymphocytoses are polyclonal expansions of T and/or B lymphocytes. Table 40.21 lists some of the common causes associated with peripheral blood lymphocytoses. An acute viral infection is the classic disease associated with a lymphocytosis. These can include infectious mononucleosis (Epstein-Barr virus), cytomegalovirus, and a variety of other viral agents. It must be emphasized that an absolute lymphocyte count must be determined in the evaluation of a lymphocytosis, so that absolute lymphocytosis can be distinguished from relative lymphocytosis owing to absolute neutropenia. It must also be remembered that the normal range age of lymphocyte counts varies with age. Children younger than 4 years of age have a much higher absolute lymphocyte count than older children and adults (Table 40.4).

The morphologic hallmark of an acute viral infection is the so-called atypical or reactive lymphocyte (Fig. 40.18). This refers to lymphocytes that are transformed and are larger than the small lymphocyte with scant cytoplasm. These transformed lymphocytes have a slightly finer chromatin than a small lymphocyte and usually have a distinct nucleus and abundant cytoplasm. The cytoplasm may have a blue to plasmacytoid appearance. Occasional azurophilic granules (large granular lymphocytes) may also be identified. These reactive lymphocytes have also been called Downey cells. The atypical lymphocytes often comprise more than 20% of the total lymphocyte count. Some diseases, such as cytomegalovirus infection, hepatitis, and toxoplasmosis, are more commonly associated with a lymphocytosis than with an increase in plasmacytoid-appearing lymphocytes and plasma cells. The lymphocytosis associated with *Bordetella pertussis* may appear as deeply clefted cells that can mimic the clefted cells of a peripheralized, small cleaved cell lymphoma (Fig. 40.18). Obviously, a clinical history of lymphoma or lymphadenopathy and the patient's age will aid in this differential diagnosis. The lymphocytes in the so-called infectious lymphocytosis disorder are usually small, with scant cytoplasm. This vague disorder can be associated with WBCs greater than 100.0×10^9/L and is not associated with any known etiology.

Lymphopenia

Lymphopenia is defined as an absolute decrease in the lymphocyte count to less than 1.5×10^9/L in adults and less than 3.0 ×

10^9/L in children. This decrease in lymphocytes may be the result of decreased production or increased destruction or loss of peripheral lymphocytes. The decreased production of lymphocytes may be associated with a variety of inherited or acquired immunologic deficiencies. Rare disorders associated with abnormalities of the lymphatic system have also been associated with lymphopenias. Acquired immunodeficiency syndrome that is associated with HIV has become a frequent cause of lymphopenia because of the lymphocytotoxic effect of the HIV virus. Other abnormalities associated with lymphocytopenias include Hodgkin's disease, irradiation, chemotherapy, corticosteroids, Cushing's disease, and other diseases (Table 40.22).

Morphologic and Functional Leukocyte Disorders

Several qualitative disorders of neutrophils may be manifested as either morphologic or functional disorders (Table 40.23). Some of these abnormalities are little more than laboratory curiosities with little clinical impact on the patient, whereas other disorders are clinically quite significant and can lead to potentially life-threatening complications.

May-Hegglin Anomaly. The May-Hegglin anomaly is an uncommon autosomal dominant disorder characterized by large, abnormal Döhle bodies in neutrophils and monocytes, giant platelets, and a variable degree of thrombocytopenia (Fig.

TABLE 40.21. DISORDERS ASSOCIATED WITH LYMPHOCYTOSIS

Physiologic (first week to 4 years)
Infectious
Infectious mononucleosis (Epstein-Barr virus)
Infectious lymphocytosis
Cytomegalovirus
Infectious hepatitis
Pertussis
Brucellosis
Toxoplasmosis
Mycoplasma
Chronic infections
Syphilis
Drug sensitivity
Dilantin
Para-aminosalicylic acid
Miscellaneous
Autoimmune disorders
Hyperthyroid toxicosis
Addison's disease
Graft rejection
Hematologic malignancies
Acute lymphoblastic leukemia
Chronic lymphocytic leukemia
Peripheralized lymphomas
Chronic lymphoproliferative diseases

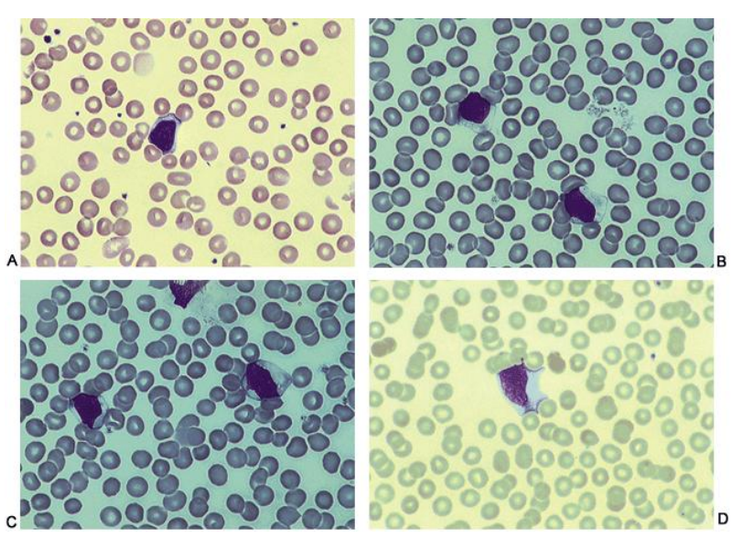


FIGURE 40.18. A–D: Various types of reactive, or atypical, lymphocytes.

40.19). Although the majority of patients with May-Hegglin anomaly are asymptomatic, some of these patients may have problems with abnormal bleeding from an unknown abnormality. Although it would be easy to surmise that a platelet defect is responsible for this abnormality, platelet function studies in the laboratory have not shown any abnormality. The Döhle bodies in May-Hegglin anomaly are quite prominent and are typically larger and bolder than the Döhle bodies associated with an acute infection. As with acute infections, the Döhle bodies in May-Hegglin anomaly consist of rough endoplasmic reticulum. The giant platelets measure anywhere from 4 to 8 μm in diameter and, when present with Döhle bodies, are diagnostic of the May-Hegglin anomaly.

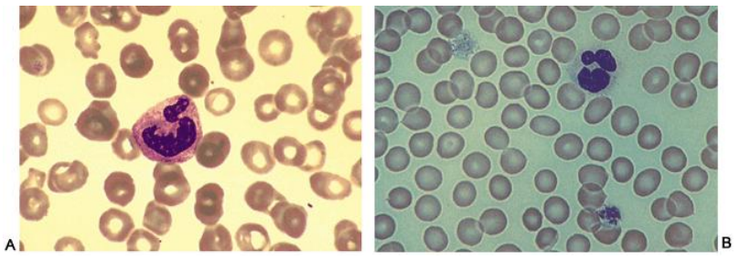


FIGURE 40.19. A, B: May-Hegglin anomaly shows large platelets and prominent Döhle bodies in the cytoplasm.

Pelger-Huët Anomaly. The Pelger-Huët anomaly is a benign autosomal dominant disorder characterized by the inability of the neutrophils to undergo proper segmentation. This results in a neutrophil that has a bilobed or dumbbell-shaped nucleus with coarsely clumped chromatin (Fig. 40.20). More than 75% of the neutrophils will show this bilobed segmentation. Patients who are homozygous for the Pelger-Huët anomaly are incredibly rare and have been reported to have neutrophils with a single, round nucleus without segmentation and with coarsely condensed chromatin. Both the homozygote and heterozygote states are associated with normal granulocytic function and are not thought to be associated with any clinical problems. Pseudo–Pelger-Huët neutrophils found in the myelodysplastic syndromes are morphologically identical to this hereditary type and are a characteristic finding in granulocytic dysplasia.

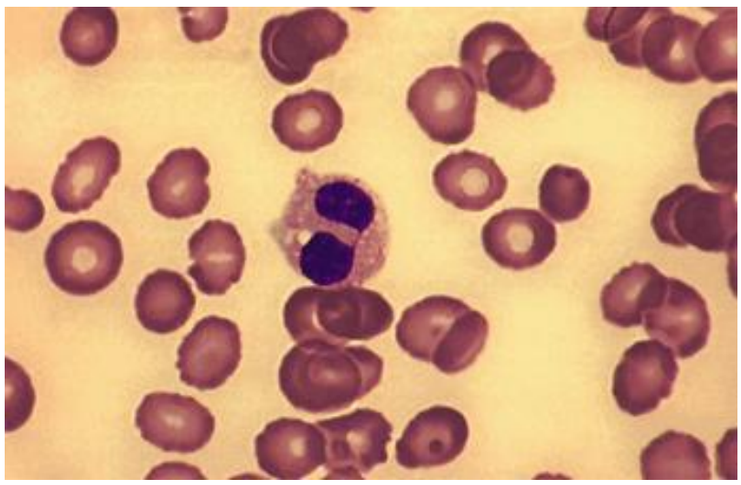


FIGURE 40.20. Pelger-Huët anomaly showing abnormal, bilobed segmentation of mature neutrophils.

Hereditary Hypersegmentation in Neutrophils. This rare autosomal dominant disorder results in giant neutrophils having five or more segments. Neutrophil precursor cells, such as the promyelocyte and myelocyte, also show nuclear indentation and segmentation. No oval macrocytes or other megaloblastic changes are seen. These neutrophils are reported to possess normal function.

Alder-Reilly Anomaly. This anomaly relates to the dense, large, azurophilic granules that can be found in neutrophils, eosinophils, basophils, and occasionally lymphocytes and monocytes from patients with a variety of metabolic disorders, including many of the mucopolysaccharidoses, such as Hurler's, Hunter's, or Maroteaux-Lamy syndromes (Fig. 40.21). This heavy granulation is easily confused with toxic granulation but is not associated with any infectious process and is a “permanent” feature of the circulating cells. In these syndromes, the neutrophils

become constipated with granular mucopolysaccharide deposits, which stain metachromatically with toluidine blue stains. This morphologic characteristic has also been reported to be found occasionally in otherwise healthy individuals.

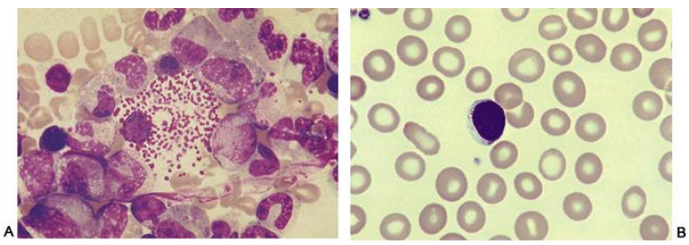


FIGURE 40.21. Alder-Reilly anomaly shows prominent cytoplasmic inclusions surrounded by a clear halo. **A**, Macrophage; **B**, lymphocyte.

Inclusions within the lymphocytes are less numerous than in the neutrophils or monocytes in patients with the Alder-Reilly anomaly. These basophilic inclusions are surrounded by a clear halo and stain metachromatically. The evaluation of the bone marrow will reveal these same inclusions in the mononuclear phagocytes, lymphocytes, and granulocytic cells.

Chédiak-Higashi Syndrome. The Chédiak-Higashi syndrome is a rare autosomal recessive disorder characterized by giant, cytoplasmic lysosomal inclusions in all types of circulating leukocytes (Fig. 40.22). The primary granules within neutrophils fuse into large cytoplasmic structures that are deficient in the enzymes required for normal neutrophil phagocytosis and chemotaxis. The granules found in Chédiak-Higashi syndrome are myeloperoxidase positive. Patients with Chédiak-Higashi syndrome also exhibit oculocutaneous albinism, recurrent infections, and an increased bleeding tendency. These patients are

TABLE 40.22. DISORDERS ASSOCIATED WITH LYMPHOPENIA

Immunodeficiency syndromes: inherited
DiGeorge syndrome
Bruton-type agammaglobulinemia
Severe combined immunodeficiency (Swiss-type)
Common variable hypogammaglobulinemia
Wiskott-Aldrich syndrome
Ataxia telangectasia
Immunodeficiency syndromes: acquired
Acquired immunodeficiency syndrome (human immunodeficiency virus)
Increased destruction
Corticosteroids
Cushing's syndrome
Radiation
Chemotherapy
Intestinal lymphocyte loss
Lymphangiectasia
Whipple's disease
Malabsorption syndromes
Malignancies
Hodgkin's disease
Terminal carcinoma
Miscellaneous
Bone marrow aplasia
Collagen vascular diseases
Renal failure
Sarcoidosis
Tuberculosis

TABLE 40.23. MORPHOLOGIC AND FUNCTIONAL LEUKOCYTE DISORDERS

May-Hegglin anomaly
Pelger-Hu't nuclear anomaly
Hereditary hypersegmentation in neutrophils
Alder-Reilly anomaly
Chédiak-Higashi syndrome
Chronic granulomatous disease
Myeloperoxidase deficiency
Miscellaneous
CD11/CD18 (CR3) deficiency
Specific neutrophil granule deficiency
Lazy leukocyte syndrome
Various immunoglobulin deficiencies
Autoimmune diseases
Chronic renal failure
Diabetes
Malnutrition
Burns
Various malignancies

functionally immunodeficient, and the majority, if they survive the infectious episodes, eventually develop high-grade lymphoproliferative processes.

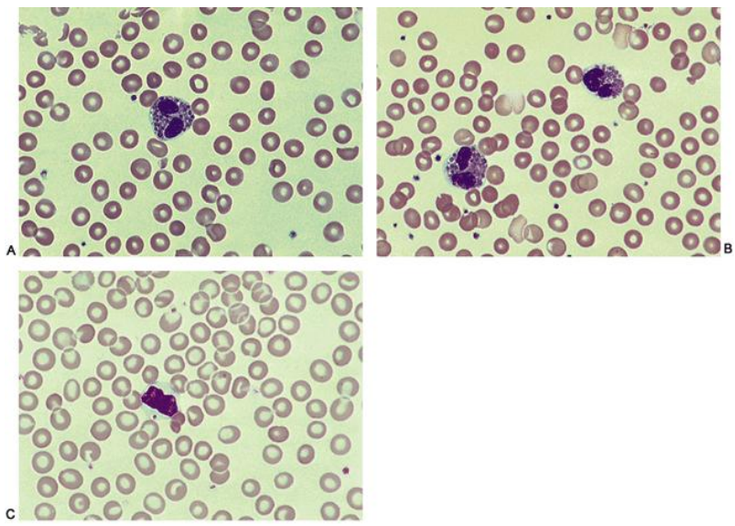


FIGURE 40.22. Chédiak-Higashi syndrome shows neutrophil **(a)**, eosinophil **(B)**, and lymphocyte **(C)**, with numerous large, fused granules within the cytoplasm.

Chronic Granulomatous Disease. Chronic granulomatous disease (CGD) is an X-linked (two thirds of cases) or autosomal recessive (one third of cases) disorder associated with a deficient functional ability of neutrophils to undergo an oxidative respiratory burst. This defect leads to an inability to succeed in the killing of bacterial organisms. Patients with CGD have recurrent and severe infections beginning early in life, typically with catalase-positive bacteria or fungi. The defect in at least the X-linked form is related to a deficiency in cytochrome b_{558}. There is also some association of CGD with the Macleod phenotype as well as Duchenne's muscular dystrophy in some patients. This relates to the close proximal location of the genes involved in these three disorders. The neutrophils and monocytes are morphologically unremarkable. Tests of phagocytic and chemotactic function are abnormal in patients with CGD. Specific tests that evaluate the oxidative metabolic pathway, such as the nitroblue tetrazolium dye test or oxidative burst analysis by flow cytometry, are used to evaluate the ability of the patient's neutrophils to undergo an intact oxidative burst and are used to confirm the diagnosis of CGD.

Myeloperoxidase Deficiency. Myeloperoxidase deficiency is a benign autosomal recessive trait associated with the absence of myeloperoxidase in neutrophils and monocytes. Paradoxically, eosinophilic myeloperoxidase is normal. Although neutrophils lack this enzyme, which mediates the oxidative burst, neutrophil function remains intact and is not associated with any significant clinical problems. Morphologically, the neutrophils and monocytes appear unremarkable. However, the myelo-peroxidase cytochemical stain will be negative in these pa-tients. Likewise, hematologic blood counters that rely on peroxidase cytochemical staining for determining differential counts will be unable to determine an accurate neutrophil count. Indeed, the majority of patients with this benign enzyme deficiency have been detected when evaluated on such hematologic analyzers.

Miscellaneous Defects. There are several other inherited or acquired defects that may interfere with any step of normal neutrophil function, including cellular adhesion to an endothelial surface, chemotaxis, opsonization, phagocytosis, degranulation, and the oxidative response. These may include specific defects such as CD11/CD18 (CR3) deficiency, specific neutrophil granule deficiency, lazy leukocyte syndrome, or various complement or immunoglobulin deficiencies. Several systemic diseases may also affect neutrophil function such as autoimmune diseases, chronic renal failure, diabetes, malnutrition, burns, and various malignancies.

SUMMARY

Part of "40 - Peripheral Blood and Bone Marrow: Morphology, Counts and Differentials, and Reactive Disorders"

This chapter outlined much of the foundation of clinical laboratory hematology. As in no other clinical laboratory, the hematologic evaluation of a patient specimen frequently requires a combination of morphologic features, instrument analysis, and clinical findings. A strong footing in morphologic skills is essential and the basis for understanding both basic hematopoiesis and clinical hematology. Modern hematology analyzers allow increasing automation within the laboratory and provide a wealth of information in a relatively rapid fashion. The frequency of hematologic abnormalities in both benign and malignant disease has underscored the importance of this field in clinical laboratory medicine. Important changes in the future will likely include increasing automation and robotics, computer-aided interpretation and control of specimen utilization, development of specific practice parameters related to the utilization of CBCs and differential counts, and the impact that modern therapeutic approaches, such as transplantation and cytokine/growth factors, will have on hematologic studies. The hematology laboratory is steeped in tradition, yet remains at the forefront of modern laboratory and clinical medicine.

REFERENCES

1. Baker J, Cornbleet PJ. Erythrocyte disorders. In: Howanitz JH, Howanitz PJ, eds. *Laboratory medicine—test selection and interpretation*. New York: Churchill-Livingstone, 1991:447–498.

2. Bessman JD, Gilmer PR Jr, Gardner FH. Improved classification of anemias by MCV and RDW. *Am J Clin Pathol* 1983;80:322–326.

3. Nelson DA, Davey FR. Hematopoiesis. In: Henry JB, ed. *Clinical diagnosis and management by laboratory methods*, 18th ed. Philadelphia: Saunders, 1991:599–601.

4. Wintrobe MM, Landsberg JW. A standardized technique for the blood sedimentation test. *Am J Med Sci* 1935;189:102–107.

5. Rümke CL, Bezemer PD, Kuik OJ. Normal values and least significant differences for differential leukocyte counts. *J Chron Dis* 1975;28:661–668.

6. Cornbleet J. Spurious results from automated hematology cell counters. *Lab Med* 1983;14:509–514.

7. van Assendelft OW. *Spectrophotometry of haemoglobin derivatives*. Assen, The Netherlands: Royal Van Gorcum, Ltd., 1970.

8. Mauer AM. *Pediatric hematology*. New York: McGraw-Hill, 1969.

SUGGESTED READINGS

General

Dacie JV, Lewis SM. *Practical haematology*, 5th ed. Edinburgh: Churchill-Livingstone, 1975.

Douglas AS, Dacie JV. The incidence and significance of iron-containing granules in human erythrocytes and their precursors. *J Clin Pathol* 1953;26:770–772.

Hyun BH, Ashton JK, Dolan K. *Practical hematology*. Philadelphia: Saunders, 1975.

Jandl JH. *Blood—textbook of hematology*. Boston: Little, Brown, 1987.

Koepke JA. *Practical laboratory hematology*. New York: Churchill-Livingstone, 1991.

Lewis SM. The constituents of normal blood and bone marrow. In: Hardisty RM, Weatherall DJ, eds. *Blood and its disorders*, 2nd ed. Oxford: Blackwell Scientific, 1982:3–56.

Miale JB. *Laboratory medicine—hematology*, 6th ed. St. Louis: Mosby, 1982.

Schumacher HR, Garvin DF, Triplett DA. *Introduction to laboratory hematology*. New York: Alan R Liss, 1984.

Sun NCJ. *Hematology—an atlas and diagnostic guide*. Philadelphia: Saunders, 1983.

Wintrobe MM, Lee GR, Boggs DR, et al. *Clinical hematology*. Philadelphia: Lea & Febiger, 1981.

Blood Cell Counting/Differential

Bentley SA. Alternatives to the neutrophil band count. *Arch Pathol Lab Med* 1988;112:883–884.

Bessman D. Microcytosis caused by RBC fragmentation. *JAMA* 1978;239:2475–2476.

Bessman JD. Heterogeneity of red cell volume: quantitation, clinical correlations, and possible mechanisms. *Johns Hopkins Med J* 1980;146:226–230.

Bessman JD. *Automated blood counts and differentials*. Baltimore: Johns Hopkins University Press, 1986.

Bessman JD, Banks D. Spurious macrocytosis, a common clue to erythrocyte cold agglutinins. *Am J Clin Pathol* 1980;74:797–800.

Bessman JD, Feinstein DI. Quantitative anisocytosis as a discriminant between iron deficiency and thalassemia minor. *Blood* 1979;53:288–293.

Bessman JD, Williams LJ, Gilmer PR. Mean platelet volume. *Am J Clin Pathol* 1981;76:289–293.

Bollinger PB, Drewinko B, Brailas CD, et al. The Technicon H*1™—an automated hematology analyzer for today and tomorrow. *Am J Clin Pathol* 1987;87:71–78.

Chien S, Dellenback J, Usami S, et al. Plasma trapping in hematocrit determination. Difference among animal species. *Proc Soc Exp Biol Med* 1965;119:1155–1158.

Coulter WH. High speed automatic blood cell counter and cell size analyzer. *Proc Natl Elect Conf* 1956;12:1034–1040.

Cox CJ, Habermann TM, Payne BA, et al. Evaluation of the Coulter Counter Model S-Plus IV. *Am J Clin Pathol* 1985;84:297–306.

Coyne R. The white blood cell differential count, a comparison of manual spread films and semi-automated spun films. *Can J Med Technol* 1976;38:18–24.

England JM. Prospects for automated differential leucocyte counting in the routine laboratory. *Clin Lab Haematol* 1979;1:263–273.

England JM, Walford DM, Waters DAW. Re-assessment of the reliability of the haematocrit. *Br J Haematol* 1972;23:247–256.

England JM, Ward SM, Down MC. Microcytosis, anisocytosis and the red cell indices in iron deficiency. *Br J Haematol* 1976;35:589–597.

Fisher SL, Fischer SP. Mean corpuscular volume. *Arch Intern Med* 1983;143:282–283.

Fossat C, David M, Harle JR, et al. New parameters in erythrocyte counting. Value of histograms. *Arch Pathol Lab Med* 1987;111:1150–1154.

Fulwyler MJ. Electronic separation of biologic cells by volume. *Science* 1965;150:910–911.

Giles C. The platelet count and mean platelet volume. *Br J Haematol* 1981;48:31–37.

Goldner FM, Mann WN. Statistical error of differential white counts. *Guys Hosp Rep* 1938;88:54–56.

Griswold DJ, Champagne VD. Evaluation of the Coulter S-PLUS IV. Three-part differential in an acute care hospital. *Am J Clin Pathol* 1985;84:49–57.

Hillman RS, Finch CA. *Red cell manual*, 5th ed. Philadelphia: Davis, 1985.

Hughes-Jones NC, Norley I, Young JMS, et al. Differential white cell counts by frequency distribution analysis of cell volumes. *J Clin Pathol* 1974;27:623–625.

Klee GG. Performance goals for internal quality control of multichannel hematology analyzers. *Clin Lab Haematol* 1990;12:65–74.

Koepke JA. *Differential leukocyte counting*. Skokie, IL: College of American Pathologists, 1978.

Koepke JA, Dotson NA, Shifman MA. A critical evaluation of the manual/visual differential leukocyte counting method. *Blood Cells* 1985;11:173–186.

Krause JR. Automated differentials in the hematology laboratory. *Am J Clin Pathol* 1990;93(Suppl 1):S11–S16.

Mansberg HP, Saunders AM, Groner W. The Hemalog D white cell differential system. *J Histochem Cytochem* 1974;22:711–724.

O'Sullivan MB. Quality control of multichannel haematology analyzers: critique of current methods and the need for performance goals. *Clin Lab Haematol* 1990;12:3–12.

Pierre RV, Payne BA, Lee WK, et al. Comparison of four leukocyte differential methods with the National Committee for Clinical Laboratory Standards (NCCLS) reference method. *Am J Clin Pathol* 1987;87:201–209.

Rümke CL. The statistically expected variability in differential leukocyte counting. In: Koepke JA, ed. *Differential leukocyte counting*. Skokie, IL: College of American Pathologists, 1978.

Saunders AM. Development of automation of differential leukocyte counts by the use of cytochemistry. *Clin Chem* 1972;18:783–788.

Shapiro MF, Greenfield S. The complete blood count and leukocyte differential count. An approach to their rational application. *Ann Intern Med* 1987;106:65–74.

Stiene-Martin EA. Causes for poor leukocyte distribution in manual spreader-slide blood films. *Am J Med Technol* 1980;46:624–632.

van Assendelft OW. Reference values for the total and differential leukocyte count. *Blood Cells* 1985;11:77–96.

Weick JK, Hagedorn AB, Linman JW. Leukoerythroblastosis: diagnosis and prognostic significance. *Mayo Clin Proc* 1974;49:110–113.

Wintrobe MM. Erythrocyte in man. *Medicine (Baltimore)* 1930;9:195–217.

Miscellaneous Tests

Cline MJ, Berlin NI. The reticulocyte count as an indicator of the rate of erythropoiesis. *Am J Clin Pathol* 1963;39:121–128.

Sox HC, Liang MH. Erythrocytes and the sedimentation rate. Guidelines for rational use. *Ann Intern Med* 1986:104:515–523.

Bone Marrow

Gulati GL, Ashton JK, Hyun BH. Structure and function of the bone marrow and hematopoiesis. *Hematol Oncol Clin North Am* 1988;2:495–511.

Reactive Disorders

Cline MJ, Lahrer RI, Territo M, et al. Monocytes and macrophages: functions and diseases. *Ann Intern Med* 1978;88:78–88.

Hyun BH, Gulati GL. Lymphocytosis and lymphocytopenia. *Lab Med* 1984;15:319–324.

Maladonado JE, Hanlon DG. Monocytosis: a current appraisal. *Mayo Clin Proc* 1965;40:246–251.

Wood TA, Frankel EP. The atypical lymphocyte. *Am J Med* 1967;42:923–936.

41

Red Blood Cell Disorders

Michele D. Raible

Brian D. Kueck

Serhan Alkan

- ANEMIA
- POSTHEMORRHAGIC ANEMIA
- IMPAIRED ERYTHROCYTE PRODUCTION

ANEMIA

Part of "41 - Red Blood Cell Disorders"

It is helpful to consider the circulating red blood cell and its progenitors as a single functional unit, the erythron. Many of the diseases afflicting the erythron may then be seen to result from the same basic processes that affect other organ systems: genetic disorders, disorders of immunity, neoplasia, infectious diseases, nutritional disease, and environmental disorders. The result of the majority of these disease processes is anemia.

Anemia is defined as a reduction in total red cell mass. As total red cell mass is not easily measured, anemia is considered to be present if the hemoglobin concentration or the hematocrit is below the normal range. The range of normal values is subject to variables imposed by lifestyles, geographic altitude, age, and gender.

Clinical Manifestations of Anemia

Anemia itself is not a disease, but rather an objective sign of an underlying process. Hence, the clinical manifestations of anemia, which result from the diminished delivery of oxygen to the tissues, are often similar regardless of the etiology. The rate of onset, the degree of reduction in red cell and plasma volume, and the adequacy of circulatory and respiratory compensation influence the signs and symptoms of anemia.

Physiologic adjustments of the cardiovascular system in response to anemia include increases in cardiac output, heart rate, and stroke volume. This compensatory high-output state may result in palpitations, tinnitus, dizziness, or syncope. High-output cardiac failure may ensue with prolonged anemia, advanced age, or underlying heart disease.

As the red cell mass decreases, blood flow is redistributed to maintain adequate visceral and cerebral oxygenation. Shunting of blood flow away from the skin results in the characteristic pallor of anemia, which is best observed by inspection of the oral mucosa, conjunctivae, or nail beds. To further facilitate oxygen delivery to critical tissues, the concentration of red cell 2,3-diphosphoglycerate increases shortly after the onset of anemic hypoxia.

At rest, few changes in respiration are noticeable. Most patients, however, exhibit signs of dyspnea on exertion. Respiratory failure is typically a sign of cardiac decompensation and is usually accompanied by pulmonary edema. Transfusion, in an attempt to improve the oxygen-carrying capacity of the blood, is tempting in such cases but must be approached with extreme caution because the risk of precipitating volume overload and further cardiovascular compromise is imminent.

A Morphologic Approach to Anemia

Although anemia may first be suspected from a patient's presenting symptoms, the diagnosis is usually based on laboratory findings. The laboratory data provided by an automated hematology analyzer supply a framework upon which further investigation can be built. Utilizing red cell indices, a morphologic classification to anemia can be structured (Table 41.1). The red cell indices only describe the average red cell, however. Important clues may be missed if evaluation does not include examination of the peripheral blood smear. A balanced population of microcytic and macrocytic red cells may otherwise go unnoticed in a normal mean corpuscular volume (MCV). Shape changes in the erythrocytes (poikilocytes) often offer clues to a number of specific disease states. Rouleaux, agglutination, or evidence of increased erythrocyte regeneration are also important findings that may direct further investigation. A measure of anisocytosis, or variation in red cell size, is often offered as part of an automated hematology profile and is referred to as the red cell distribution width (RDW) or the red cell morphology index (RCMI). Adding this information may help further stratify the morphologic classification of anemias.

TABLE 41.1. MORPHOLOGIC CLASSIFICATION OF ANEMIA

Microcytic hypochromic anemias
Iron deficiency anemia
Anemia of chronic disease
Thalassemias
Sideroblastic anemia
Normocytic normochromic anemias
Anemia of chronic disease
Aplastic anemia
Myelophthisic processes
Some hemolytic anemias
Some hemoglobinopathies
Macrocytic anemias
Vitamin B_{12} deficiency
Folic acid deficiency
Aplastic anemia
Liver disease
Myelodysplastic syndromes
Some hemolytic anemias

A Pathophysiologic Approach to Anemia

In addition to the information provided by the automated hematology analyzer, the history and physical examination may also provide clues directing further study. The history and review of the prior medical and laboratory data should help to establish the onset and duration of the anemia. Attention to cardiovascular complaints will also provide similar information. Blood loss may be documented by a history of menorrhagia or melena. Dietary inadequacies, current medicinal use, or history of toxin exposure may be elicited in questioning. A family history may provide evidence of an inherited red cell abnormality. Systemic symptoms such as fever or weight loss might point toward neoplasia or infection.

The physical examination should document the signs of anemia.

Signs of hyperbilirubinemia, which may suggest increased red cell destruction, should be sought. Lymphadenopathy or hepatosplenomegaly may provide clues to an underlying infection, lymphoproliferative disorder, or myeloproliferative state. Changes detected in the neurologic examination may reflect not only nonspecific signs of global cerebral hypoxia but may also provide clues indicative of vitamin B_{12} deficiency or hypothyroidism.

The information obtained through the history and physical examination is helpful in understanding the physiologic mechanisms underlying a decreased red cell mass. A pathophysiologic approach to the anemias can be utilized as an alternative to the morphologic classification scheme provided by the initial laboratory data (Table 41.2). Subsequent discussion of the red cell disorders will follow such a pathophysiologic approach. It is important, however, to understand both classification schemes and how they relate, as evaluation of the anemic patient requires a thorough and careful history, physical examination, laboratory investigation, and examination of the peripheral blood smear.

TABLE 41.2. A PATHOPHYSIOLOGIC APPROACH TO ANEMIA

1. Blood loss
2. Impaired red cell production
   a. Disorders of heme synthesis
   b. Disorders of DNA synthesis
   c. Inadequate erythroid precursors
3. Accelerated red cell destruction
   a. Intrinsic/inherited red cell abnormalities
   b. Extrinsic/acquired red cell abnormalities
   c. Paroxysmal nocturnal hemoglobinuria

POSTHEMORRHAGIC ANEMIA

Part of "41 - Red Blood Cell Disorders"

Acute Blood Loss

The clinical manifestations of acute blood loss are largely related to the loss of blood volume. Healthy individuals may tolerate a loss of as much as 20% of blood volume without clinical manifestations. The loss of larger quantities of blood results in cardiovascular distress; losses exceeding 50% of total blood volume typically result in death. The following discussion presumes a single episode of acute blood loss unattended by therapeutic intervention, an infrequent clinical scenario.

Immediately after acute blood loss, a proportional decline of plasma volume and red cell mass occurs, thus, the hemoglobin and hematocrit remain normal. In an attempt to maintain an adequate blood volume, albumin and fluid are shifted from the extravascular space to the vascular compartment. As this is a slow process, the maximum fall in hematocrit is reached approximately 3 days after blood loss when plasma volume is restored; before this time the extent of blood loss is underestimated by the hematocrit. The resulting anemia is initially normochromic and normocytic.

A compensatory increase in red cell production is stimulated by the release of erythropoietin. A lag phase of 3 to 5 days occurs as the erythrocytes mature in the marrow. Subsequently, evidence of increased erythrocyte regeneration is seen in the peripheral blood reflected by increased numbers of polychromatophilic macrocytes or by an increased reticulocyte count. Maximal reticulocytosis is evident approximately 10 days after blood loss; at this time, an increased MCV may be noted.

The platelet and leukocyte counts are also affected by acute blood loss. An early transient thrombocytopenia is typically followed within hours by a thrombocytosis that may reach $1{,}000 \times 10^9$/L. The leukocyte count may rise as high as 35×10^9/L, in part responding to an epinephrine-induced shift of the marginating granulocytes to the circulating granulocytic pool. Leukocytosis, which may be accompanied by a mild left shift in granulocytic maturation, occurs within several hours of blood loss and typically resolves within 2 to 4 days.

Such changes in the peripheral blood are readily recognizable as secondary to acute blood loss when hemorrhage is external. Acute blood loss into a body space or cavity, however, may initially go undetected clinically. In such instances, anemia, coupled with evidence of increased erythrocyte regeneration, might suggest a hemolytic anemia. Hyperbilirubinemia, typical of hemolytic anemias, may also be evident as internal blood loss leads to erythrocyte breakdown and heme metabolism. Other laboratory findings often seen in conjunction with hemolytic anemia, such as erythrocyte shape abnormalities, decreased serum haptoglobin, hemoglobinemia, hemoglobinuria, and hemosiderinuria are absent, however.

Chronic Blood Loss

The previously described changes of acute blood loss are lacking if the loss of blood occurs slowly over an extended period of time. Chronic blood loss results in anemia only when the rate of blood loss exceeds the regenerative capacity of the bone marrow or when the body's iron stores are depleted. In the latter circumstance, the resultant blood picture is one of iron deficiency.

IMPAIRED ERYTHROCYTE PRODUCTION

Part of "41 - Red Blood Cell Disorders"

The anemias discussed in this section each result from the bone marrow's inability to produce adequate numbers of functional

erythrocytes. This may arise from an inadequate number of erythroid precursors or from an insufficient quantity of substances essential for nuclear or cytoplasmic maturation. A deficiency of such factors vital to cell maturation often affects nonhematopoietic tissue as well; the following discussion, however, focuses only on the consequences to the hematopoietic system.

Disorders of Heme Synthesis

Cytoplasmic maturation is a reflection of hemoglobin synthesis. The normal biosynthesis of hemoglobin requires that iron, protoporphyrin, and globin are present in optimal amounts within the developing erythrocyte. Therefore, abnormalities in synthesis

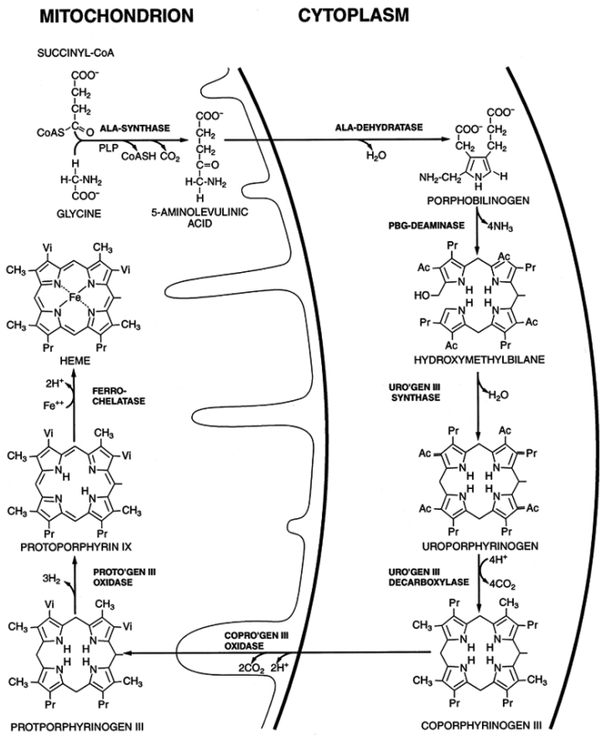


FIGURE 41.1. Heme biosynthesis. M, methyl (-CH_3); P, propionate (-CH_2 -CH_2 -COOH); V, vinyl (-CH=CH_2); ALA, (-aminolevulinic acid; PBG, porphobilinogen; Uro I synthase, uroporphyrinogen I synthase; Uro III cosyn, uroporphyrinogen III cosynthase; Urodecarb, uroporphyrinogen decarboxylase. Uro'gen, uroporphyrinogen; Copro'gen, coproporphyrinogen; Proto'gen, protoporphyrinogen. (From Wintrobe MM, et al. Erythropoiesis. In: *Clinical hematology*, 8th ed. Philadelphia: Lea & Febiger, 1981, with permission.)

or integration of any of these three constituents result in an anemia characterized by deficient hemoglobin synthesis, morphologically manifest as hypochromic erythrocytes. Because mitosis is terminated when a critical intracellular concentration of hemoglobin is reached, erythrocytes deficient in hemoglobin typically undergo an additional cell division, resulting in the production of small or microcytic red blood cells. Abnormalities of globin synthesis are discussed in Chapter 42, and, therefore, the following discussion is limited to abnormalities of heme synthesis (Fig. 41.1).

Under the influence of δ-aminolevulinic acid (ALA) synthetase and pyridoxal 5′-phosphate as a cofactor, succinyl coenzyme A (CoA) and glycine combine to form ALA. This initial step is the major rate-limiting step in heme synthesis, largely controlled as an inverse function of the intracellular concentration of heme. Two molecules of ALA then combine to form porphobilinogen, with subsequent condensation of four porphobilinogen molecules to form the tetrapyrrole ring structure of uroporphyrinogen. Uroporphyrinogen undergoes a number of changes, each under the influence of specific enzymes, eventually resulting in an oxidized molecule of protoporphyrin IX. The final step in heme synthesis occurs in the mitochondria as ferrous iron is incorporated into protoporphyrin IX.

Diminished heme synthesis will result from abnormalities at any point in this multistep pathway. Examples include pyridoxine deficiency, or the diminished conversion of pyridoxine to pyridoxal 5′-phosphate seen in alcoholics; the toxic effects of lead, which inhibits ALA dehydrase and inhibits iron delivery to protoporphyrin IX; and deficiencies of enzymes integral to the tetrapyrrole conversions of uroporphyrinogen that result in the rare, typically inherited group of disorders known as the porphyrias.

Normally, protoporphyrin IX is produced in slight relative excess to the amount of iron available for heme synthesis and can be measured clinically as free erythrocyte protoporphyrin (FEP). FEP is actually zinc protoporphyrin, a naturally fluorescent moiety. The action of heme synthetase is not specific for iron and zinc, the second most plentiful transitional metal within the erythrocyte, is readily incorporated into protoporphyrin IX when iron is unavailable. If incorporation of iron is reduced, owing to either an enzymatic block or a deficiency of available iron, an excess of protoporphyrin IX results (increased FEP). Increased FEP is perhaps the most sensitive laboratory indicator available for iron-deficient erythropoiesis.

Iron-Deficiency Anemia

If one excludes the clinically silent thalassemia syndromes, iron deficiency is the primary cause of anemia worldwide. To understand the multiple causes of iron deficiency and the consequent biochemical changes that are used clinically as aids in establishing a diagnosis of iron-deficiency anemia, one must first understand iron metabolism.

Iron Metabolism

Iron is supplied exogenously through the diet in two forms: heme iron and nonheme iron. Heme iron is derived from hemoglobin, myoglobin, or other heme proteins in foods of animal origin. Nonheme iron must be converted to or maintained in the ferrous state to facilitate absorption, a process enhanced by the acid environment of the stomach. Iron is absorbed primarily in the duodenum or upper jejunum by means of specific mucosal receptors. The absorption of iron is regulated at the level of the mucosal cell, and hence, the tight balance of iron is controlled (Fig. 41.2). Under normal circumstances, 5% to 10% of ingested iron is absorbed; in states of iron deficiency, absorption may increase severalfold. Exactly how this process occurs and how it is controlled remain unclear.

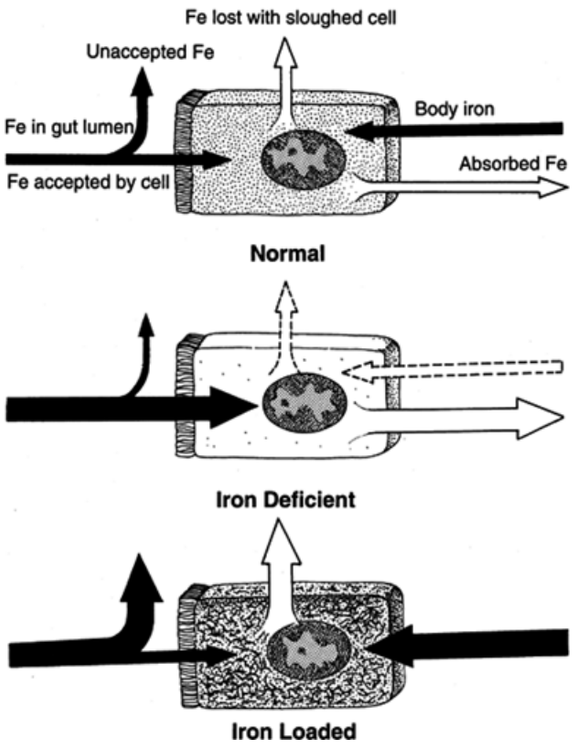


FIGURE 41.2. The concentration of iron within the intestinal mucosal cell likely parallels body iron stores and serves to regulate the amount of iron absorbed. Iron accepted into the mucosal cell may pass through the cell and enter the circulation, a fate that is promoted by decreased intracellular iron concentration. Alternatively, iron may be retained within the mucosal cell and subsequently lost from the body as the cell is sloughed with normal epithelial turnover. Increased intracellular iron promotes retention of iron as well as diminishing acceptance of iron into the cell. (Modified from Crosby WH. The control of iron balance by the intestinal mucosa. *Blood* 1963;22:441.)

Once absorbed, iron may be sequestered within the mucosal epithelial cell to be subsequently excreted with cell death, or it may pass through the cell to enter the circulation, where it is quickly complexed with the plasma carrier protein transferrin. The synthesis of transferrin, occurring principally in the liver, is regulated by feedback inhibition in inverse relation to the body's iron stores. Under normal circumstances, approximately one third of the binding sites of transferrin are saturated with iron.

Transferrin provides purposeful transport of iron by directing

the transferrin-iron complex to transferrin receptors located on the surface of all cells requiring iron. The critical role of transferrin in iron transport is evident by the rare congenital disorder atransferrinemia. The absence of transferrin in this disease results in severe transfusion-dependent anemia as well as hemosiderosis owing to random iron deposition throughout the body.

When bound to the surface of the developing red cell, the transferrin-iron complex is internalized by a process of pinocytosis. Iron is then liberated from transferrin with subsequent return of the intact transferrin molecule to the plasma, where it resumes its transport function.

The majority of the iron delivered to the developing erythrocyte is incorporated into heme. A small amount of residual nonheme iron is accounted for as ferritin. These small ferritin granules can be visualized with a Prussian blue stain of a bone marrow aspirate. Normally, one to five small randomly distributed Prussian blue–positive granules are seen in the cytoplasm of 20% to 60% of normoblasts; these nonheme iron-containing erythroblasts are referred to as sideroblasts (Fig. 41.3). The number of sideroblasts present reflects both the availability of iron and the incorporation of iron into heme. A good correlation exists between the number of sideroblasts and the percentage of saturation of transferrin, defined as serum iron × 100/total iron binding capacity or serum iron × 100/transferrin. These normal cytoplasmic ferritin deposits are removed by solubilization and extrusion of ferritin into the surrounding media, a process that requires active cellular oxidative metabolism.

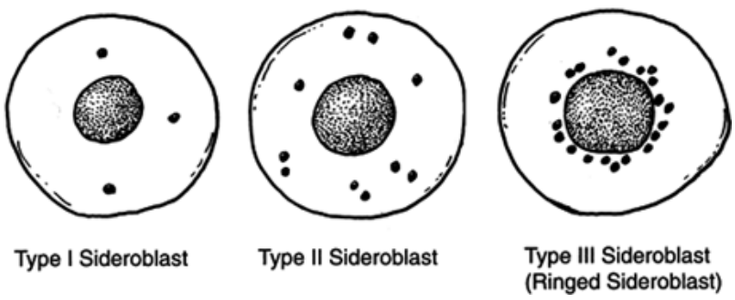


FIGURE 41.3. Small particles of nonheme iron may be seen in the cytoplasm of developing erythroblasts utilizing iron stains such as Prussian blue. Type I sideroblasts contain one to five small, randomly distributed iron particles and account for 20% to 60% of erythrocytes. Type II sideroblasts contain five or more coarse, randomly distributed iron granules and are often seen in states of excess reticuloendothelial iron and/or abnormal iron utilization. Type III sideroblasts, or ringed sideroblasts, are the morphologic hallmark of the sideroblastic anemias and are characterized by a ring or arc of iron encircling two thirds or more of the erythroblast nucleus.

Iron incorporated into heme remains in the erythrocyte throughout its life span. Erythroid senescence results in removal of the red cell from circulation, primarily by the splenic macrophages. Within the macrophage, hemoglobin is catabolized with ultimate liberation of iron, protoporphyrin, and globin. Iron is then stored in the reticuloendothelial system of the spleen, bone marrow, and liver. From these storage sites, iron is liberated to transferrin to be directed to a newly developing erythrocyte and to begin the iron cycle once again.

Iron is stored in two forms: hemosiderin and ferritin. Ferritin is well characterized and is the major physiologic storage form of iron. Hemosiderin, considered to represent an agglomerate of degraded ferritin and debris from lysosomal vacuoles, is distinguished from ferritin by its insolubility in water and its intense stainability with Prussian blue. The latter characteristic is employed in assessing iron stores in tissue sections or bone marrow aspirate smears.

Iron stores may also be assessed by quantitation of serum ferritin. Although ferritin is typically sequestered within the reticuloendothelial system, a trace amount normally leaks into the circulation. Serum ferritin generally correlates with the body's iron stores. A low serum ferritin is usually a valid indication of decreased iron stores. An elevated serum ferritin, however, does not necessarily indicate adequate or excess body iron stores because ferritin shows an acute phase response, which may persist for several days to weeks. Consequently, serum ferritin may not be a valid indication of iron stores when measured in the face of underlying inflammatory processes, neoplasia, or liver disease. Interestingly, immunohistochemical assessment of ferritin in bone marrow biopsy specimens has been reported to be a very sensitive tool in the evaluation of body iron stores.

Etiology

From the foregoing discussion, it is evident that iron is highly conserved by the body. How then does iron-deficiency anemia develop? In general, iron deficiency may arise either from an inadequate intake of iron or from a depletion of the body's iron stores. Dietary inadequacy is typically not a problem in the United States, except during periods of increased metabolic need. Rapid periods of growth, as found during the first year of life and early in adolescence, require additional dietary intake. Consequently, infants between the ages of 6 and 20 months are vulnerable, particularly those maintained solely on an unsupplemented milk diet. Increased demands for iron are also noted during pregnancy and lactation.

A decreased intake of iron may arise in the presence of an adequate diet if the body is incapable of absorbing iron. Malabsorption of iron may be seen with achlorhydria, after partial or total gastrectomy, or as a complication of diffuse enteritis (e.g., celiac disease). Pica, the ingestion of unusual substances, may be seen both as a cause and as a manifestation of iron deficiency. When habitually ingested, substances such as clay or starch can not only interfere with iron absorption but also serve as an iron-poor caloric substitute.

In the adult population of the United States, depletion of the body's iron stores through chronic blood loss is a much more common cause of iron-deficiency anemia than is an insufficient intake of iron. If an average diet is consumed, a steady loss of as little as 3 to 4 mL of blood per day can result in negative iron balance. Gastrointestinal bleeding is the most frequent cause of iron deficiency in men and postmenopausal women; menstrual blood loss must be considered first in premenopausal women. Common gastrointestinal disorders resulting in blood loss are presented in Table 41.3. Unusual causes of chronic blood loss include hemoglobinuria and pulmonary hemosiderosis.

TABLE 41.3. COMMON CAUSES OF GASTROINTESTINAL BLOOD LOSS

Esophagus
Varices
Stomach
Ulcer
Gastritis
Small intestine
Ulcer
Crohn's disease
Colon
Ulcerative colitis
Diverticulitis
Carcinoma
Rectum
Hemorrhoids

Clinical Symptoms

Iron-deficiency anemia is typically insidious in onset and gradual in progression. Consequently, patients adapt well and medical attention is often delayed. Many patients present with signs and symptoms common to anemia. Others may present with symptoms related to the cause of the iron deficiency (e.g., ulcer disease), whereas in some patients, the anemia may be discovered incidentally.

Because iron is vital to the growth and development of all cells, it is not surprising that a deficiency of iron leads to manifestations beyond the hematopoietic system. The clinical changes appear as the body becomes incapable of replacing cells at a rate equivalent to cell loss or exfoliation. Abnormalities of the nails, tongue, and upper gastrointestinal tract are common. Brittle, breakable, or ridged fingernails may be seen, eventually becoming concave or spoon-shaped (koilonychia). Absence or flattening of the papillae of the tongue may be accompanied by soreness or burning. Angular stomatitis, characterized by fissures or ulcers at the corners of the mouth, may also be noted. Upper esophageal webs or strictures may develop; the combination of dysphasia, stomatitis, lingual abnormalities and hypochromic anemia is referred to as Paterson-Kelly or Plummer-Vinson syndrome.

Laboratory Findings

Serum Chemistry. The appearance of a microcytic hypochromic anemia is actually a finding late in the course of iron deficiency. Several laboratory changes reflective of the body's declining iron stores are observed before the appearance of microcytic hypochromic red cells. Decreased iron stores, reflected by decreased serum ferritin, are found early in the course of iron deficiency, as stored iron is liberated to transferrin to maintain adequate delivery of iron to the developing erythroblast. As iron stores are depleted, serum iron falls and a compensatory increase in transferrin synthesis begins. These two events result in decreased transferrin saturation; by definition, iron-deficient erythropoiesis is present when the transferrin saturation falls below 16%. Elevation of FEP becomes evident when insufficient iron is available for heme synthesis. As previously mentioned, although low serum ferritin is usually a very good indicator of decreased iron stores, elevated serum ferritin may not indicate adequate or excess body iron stores, particularly in the presence of underlying inflammatory processes and other disorders. Recently, a quantitative test measuring the soluble fragment of transferrin receptor in serum (sTfR) has become available. Elevated levels of sTfR are present in iron-deficient subjects, as well as in individuals with malignancies or conditions with ineffective hematopoiesis. Levels do not appear to be increased in anemia of chronic disease without iron deficiency. Although further studies are necessary, measurement of sTfR may prove to be useful in distinguishing iron deficiency from anemia of chronic disease.

Blood. Early in the course of iron deficiency, the anemia appears normocytic and normochromic. As heme synthesis is impaired, hypochromic erythrocytes become evident. Microcytosis usually develops in tandem. Consequently, the MCV, mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCHC) are typically proportionally reduced. Examination of the peripheral blood smear reveals an erythroid population heterogeneous in both size and shape, features that become quite striking in severe cases (Fig. 41.4). As iron deficiency becomes pronounced, many erythrocytes exhibit only a thin peripheral rim of hemoglobin. Poikilocytes that are often seen include ovalocytes, pencil cells, teardrop forms, fragments, and target cells.

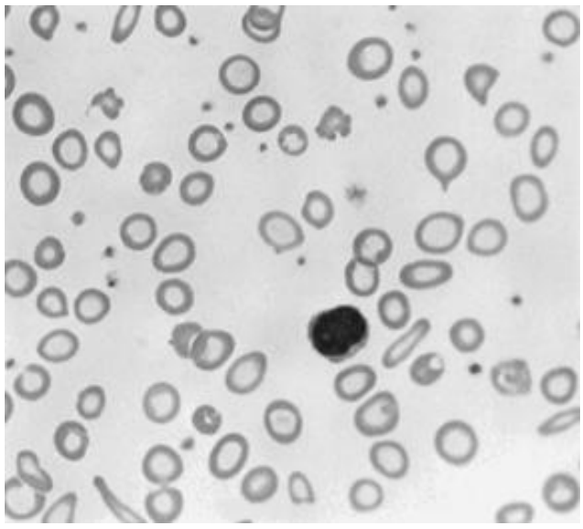


FIGURE 41.4. The peripheral blood smear in fully developed iron-deficiency anemia exhibits erythrocytes rich in central pallor. Although virtually all erythrocytes are microcytic, note the prominent variation in both size and shape, typical of iron-deficiency anemia.

The leukocyte count is typically normal, although slight granulocytopenia may be seen. The platelet count is often elevated, commonly as much as twice normal. Values reaching $1{,}000 \times 10^9$/L may rarely be seen, mimicking essential thrombocythemia.

Bone Marrow. The bone marrow often demonstrates mild erythroid hyperplasia. The developing erythroblasts appear small with scanty cytoplasm and exhibit frayed or ragged cell borders (Fig. 41.5). Nuclear budding or fragmentation may be noted. Periodic acid-Schiff-positive erythroblasts indicative of abnormal cytoplasmic maturation may also be seen. Many of these small abnormal erythroblasts die within the marrow, resulting in

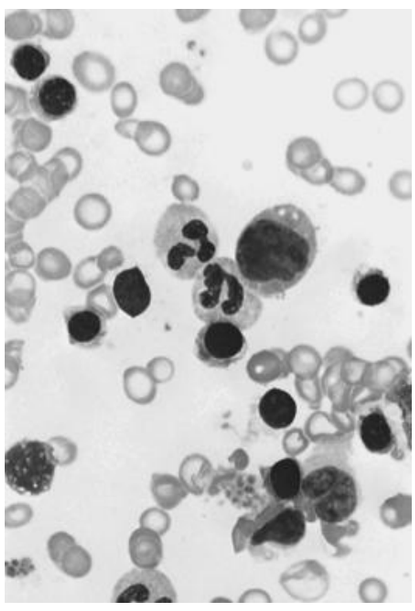


FIGURE 41.5. The developing erythroblasts in iron deficiency are small with scanty cytoplasm, often having a frayed, ragged, or moth-eaten appearance.

ineffective erythropoiesis. A Prussian blue stain demonstrates an absence of stainable iron. Sideroblasts are decreased or absent.

Therapy

It cannot be overemphasized that iron-deficiency anemia is not a disease but merely a sign of disease. Consequently, patients presenting with iron-deficiency anemia must be thoroughly evaluated to establish the cause of the deficiency. After the cause has been identified and appropriate therapy instituted, iron stores must be replenished. This is usually best accomplished by oral administration of ferrous iron salts, usually ferrous sulfate; although several oral iron preparations are available, cost is generally increased and substantial benefits compared with ferrous sulfate are often lacking. Parenteral administration of iron is infrequently necessary but may be indicated in unusual circumstances such as malabsorption, intolerance to or inability to complete a course of oral therapy, or a continued rapid rate of blood loss (so that oral therapy cannot compensate).

Signs of iron deficiency usually begin to abate within days of iron therapy. Hematologic response can be monitored with the reticulocyte count, which should reach a maximum 7 to 10 days after the institution of therapy. The reticulocyte count seldom exceeds 10% and slowly falls as the hemoglobin level rises. Normalization of the hemoglobin should be complete within 8 weeks, but continued iron therapy should be continued to replenish storage iron, generally 3 to 6 months.

Sideroblastic Anemia

Sideroblastic anemias are a group of disorders characterized by impaired utilization of iron resulting in diminished heme synthesis. The uptake of iron by the developing erythrocyte is partly controlled by the concentration of intracellular heme, which acts as a feedback inhibitor. The diminished heme synthesis resulting from impaired utilization of iron therefore results in a continued stimulus for iron absorption despite an adequate or increased level of intracellular iron. Excess iron is deposited in the mitochondria of the erythrocytes, the site of incorporation into protoporphyrin IX. The excessive deposition of iron is seen as rings or arcs around the nucleus of the erythroblast with a Prussian blue stain; these ringed sideroblasts are the morphologic hallmark of this group of disorders (Fig. 41.6). Ringed sideroblasts often succumb to the toxic effects of iron, leading to intramedullary destruction or ineffective erythropoiesis.

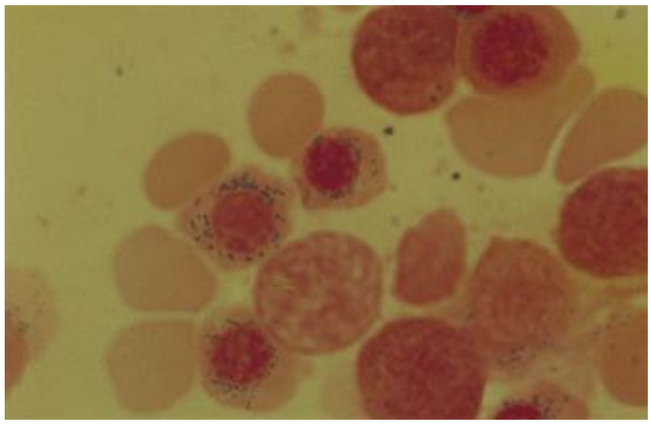


FIGURE 41.6. Ringed sideroblasts (Prussian blue stain). Numerous erythroblasts demonstrate deposition of granular iron nearly encircling the nucleus.

Diminished heme synthesis results in microcytic hypochromic erythrocytes. The degree of microcytosis and hypochromia varies among the different forms of sideroblastic anemia. Often, a dimorphic red cell population is present consisting of a distinct population of microcytic hypochromic erythrocytes among normocytic or even macrocytic red cells (Fig. 41.7).

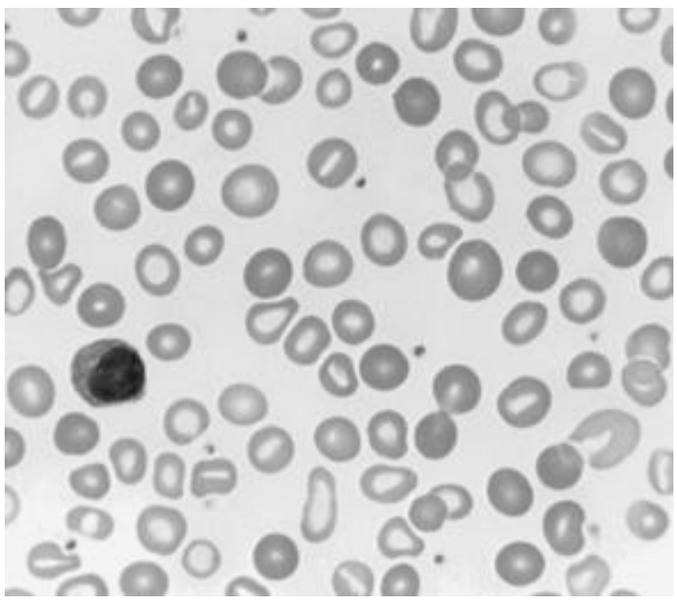


FIGURE 41.7. The peripheral blood smear of refractory anemia with ringed sideroblasts is characterized by a dimorphic red cell population. A distinct population of hypochromic, often microcytic erythrocytes can be found admixed with normochromic, normocytic, or macrocytic erythrocytes.

Iron absorption by the body continues despite the inability of the erythron to utilize iron properly in heme synthesis. Consequently, plasma (serum) iron rises, transferrin becomes saturated, and transferrin synthesis is suppressed. With time, the reticuloendothelial stores become overloaded, and iron is deposited in the parenchymal cells of various tissues. Excessive iron deposition in the heart, pancreas, or liver can result in organ dysfunction and sequelae analogous to the genetic disorder hemochromatosis.

These three features, ringed sideroblasts, a population of microcytic hypochromic erythrocytes, and hyperferremia, characterize the sideroblastic anemias. As a group, these disorders may be divided into hereditary and acquired forms.

Hereditary Sideroblastic Anemia

Hereditary forms of sideroblastic anemia are distinctly less common than acquired sideroblastic anemia. Most hereditary forms have been traced to an X-linked pattern of inheritance. Typically, it is only the male hemizygote who exhibits manifestations of the disease; rarely, female carriers may be moderately anemic.

The severity of the anemia is variable. Often severe microcytic hypochromic anemia is detected in infancy or early childhood, but mild forms of the disease may go undetected until early adulthood. The red cell population often exhibits marked anisopoikilocytosis, obscuring the classic dimorphic picture ascribed to sideroblastic anemia. Coarse basophilic stippling is present, and, when found in microcytic hypochromic cells, is a clue to the diagnosis. Ringed sideroblasts constitute 10% to 40% of erythroblasts and are typically seen only in late stages of erythroid development. Serum iron is high, and transferrin and reticuloendothelial stores are saturated. Over the course of years, parenchymal deposition may lead to cirrhosis, cardiomyopathy, and/or diabetes. Consequently, the removal of excess iron should be initiated. This may not only delay or prevent the damage of parenchymal iron deposition, but it may also improve the hematologic status of the patient by diminishing the toxic effects of excess iron on the erythron.

Variable defects of ALA synthetase have been implicated as the cause of X-linked sideroblastic anemia. Some forms respond to pharmacologic doses of pyridoxine; although not truly pyridoxine deficient, these patients possess ALA synthetase variants that either have an increased Michaelis constant (K_m) for pyridoxal 5′-phosphate or are unstable and abnormally sensitive to mitochondrial proteases. Clinical responses to pyridoxine are variable. Subtotal correction of the hemoglobin is common with persistence of microcytic hypochromic erythrocytes.

A few cases of hereditary sideroblastic anemias have been described that, although clinically indistinguishable from the X-linked variants, are indeterminate in their pattern of genetic transmission. An autosomal recessive pattern of inheritance has been suggested.

Acquired Sideroblastic Anemia

Acquired sideroblastic anemia is typically a disorder of adulthood, occurring as an idiopathic condition or as a secondary reversible phenomenon during administration of certain medicinal agents or after exposure to various myelotoxic chemicals (Table 41.4).

TABLE 41.4. DRUGS IMPLICATED IN SECONDARY SIDEROBLASTIC ANEMIA

Antituberculous agents
Isonicotinic acid hydrazine
Pyrazinamide
Cycloserine
Ethionamide
Chloramphenicol
Chemotherapeutic agents
Alkylating agents
Antimetabolites
Lead
Alcohol

The antituberculous drugs and alcohol interfere with pyridoxine metabolism or the conversion of pyridoxine to the active pyridoxine 5′-phosphate, respectively. Ringed sideroblasts may be transiently associated with the administration of chemotherapeutic drugs or may appear years after therapy as part of a secondary or therapy-related myelodysplastic syndrome or leukemia. Lead is toxic to the erythron in several ways, including inhibitory effects at multiple sites in the heme biosynthetic pathway. Regardless of the etiology, the number of ringed sideroblasts and the degree of anemia and microcytosis tend to be less prominent than in either the inherited or idiopathic forms of sideroblastic anemia.

Ringed sideroblasts have also been described in association with a variety of other hematologic and nonhematologic conditions, including myeloproliferative disorders, multiple myeloma, lymphomas, megaloblastic anemia, collagen vascular disorders and myxedema. In some circumstances, this finding may be secondary to the administration of therapeutic agents, as previously described.

Known causes of sideroblastic anemia must be excluded before rendering the diagnosis of idiopathic sideroblastic anemia, which falls within the diagnostic category of myelodysplastic syndromes as refractory anemia with ringed sideroblasts. The myelodysplastic syndromes, including refractory anemia with ringed sideroblasts, are discussed in Chapter 43.

Iron Overload

Iron is highly conserved by the body. Under normal circumstances, the 1 to 2 mg of iron lost daily through turnover of the upper gastrointestinal epithelium is balanced by absorption of ingested iron. Prolonged administration of iron or enhanced absorption of iron cannot be matched by increased iron excretion because there is not a physiologic pathway for iron excretion in humans. A gradual progressive buildup of iron eventually leads to accumulation outside the reticuloendothelial system, causing injury to parenchymal cells. It is postulated that the toxic effects of iron may be related to the formation of free radicals, causing lipid peroxidation, or to the liberation of lysosomal enzymes. The clinical manifestations of parenchymal damage are referred to as hemochromatosis. Organs principally involved include the liver (cirrhosis), the heart (cardiomyopathy or arrhythmias), and the pancreas (glucose intolerance). Additionally, later in the disease, patients may exhibit bronze pigmentation of the skin, resulting from a combination of iron deposition and increased

melanin. Arthritis, resembling rheumatoid arthritis, and manifestations of hypogonadism may also be present. Hemochromatosis has been divided into hereditary (primary) and secondary forms. Common causes of hereditary and secondary forms of hemochromatosis are shown in Table 41.5.

TABLE 41.5. HEREDITARY (PRIMARY) AND SECONDARY CAUSES OF HEMOCHROMATOSIS

Hereditary hemochromatosis
Genetic abnormalities involving *HFE* gene
C282Y homozygosity
C282Y/H63D compound heterozygosity
Non-*HFE* related
Juvenile hemochromatosis
Autosomal dominant hemochromatosis
Secondary hemochromatosis
Iron-loading anemias:
Thalassemia, sideroblastic anemia, chronic hemolytic anemias
Dietary or parenteral iron overload
Chronic liver disease
Dysmetabolic iron overload
Post-portacaval shunting

Hereditary Hemochromatosis

Increased body iron stores with primary parenchymal cell deposition of iron owing to inappropriate iron absorption is the hallmark of hereditary hemochromatosis (HHC). The diagnosis of HHC requires a high index of clinical suspicion and careful clinicopathologic correlation. HHC should be considered in any patient with unexplained hepatomegaly, abnormal skin pigmentation, cardiomyopathy, diabetes, arthritis, or hypogonadism. Previous criteria for the diagnosis of HHC have included stainable hepatic iron of grade 3 or 4 (of four grades), hepatic iron concentration greater than 80 μmol/g (dry weight), and hepatic iron index greater than 1.9. Clinical tests such as elevated serum ferritin or increased transferrin saturation are also currently used.

Recent advances in the understanding of HHC have enhanced the diagnostic capabilities in this disorder. Most patients with HHC carry a mutation involving the hemochromatosis gene (*HFE*) that encodes a novel nonclassic major histocompatibility complex (MHC) class I-like molecule. There are two common mutations involving the *HFE* gene. The more common mutation involves a point mutation at nucleotide 845 (G to A), resulting in a substitution of tyrosine for cysteine at amino acid 282 (referred to as C282Y mutation). The other mutation, in which a point mutation occurs at nucleotide 187 (C to G), results in a substitution of aspartate for histidine at amino acid 63 (H63D). The majority of the HHC patients (85% to 90%) have homozygosity for the C282Y mutation. Some recent studies reported mutations involving non-*HFE*–related genes. For example, the juvenile hemochromatosis locus has been mapped to chromosome 1q, which is genetically distinct from adult *HFE*-associated hemochromatosis.

HHC is a common inherited disorder, with a prevalence of one in 300 to 400 U.S. whites. As many as one in 10 white Americans have been estimated to carry at least one allele with the C282Y mutation. The prevalence in other ethnic populations appears to be lower. Although clinically significant iron overload usually develops in individuals who are homozygous for C282Y, the phenotype shows considerable variation, suggesting that other factors affect the overall expression of the disease. Occasional C282Y homozygotes do not have evidence of iron overload. Individuals heterozygous for the C282Y mutation may have increased serum iron indices and a small increase in hepatic iron concentration but usually do not develop progressive iron accumulation or significant liver disease in the absence of cofactors or additional mutations.

H63D homozygotes and compound heterozygotes (C282Y and H63D) patients have a higher risk of hemochromatosis compared with normal population.

HFE encodes for a 343 amino acid nonclassic MHC class I molecule. This protein contains a short cytoplasmic tail, a membrane-spanning domain, and three extracellular domains analogous to the α1, α2 and α3 domains of MHC class I molecules. The α1 and α2 domains form the ligand-binding cleft in the MHC class I molecules. The α3 domain is the most highly conserved within the class I molecules. This region interacts with $ß_2$-microglobulin, which appears to be essential for the correct cell surface presentation of the class I MHC molecules.

HFE is believed to play a general role in the cellular uptake of iron. The HFE–$ß_2$-microglobulin complex is expressed on the basolateral surface of the epithelial cells in the stomach, colon, and biliary tract, as well as on the sinusoidal lining cells of the liver. In the crypt cells of the small intestine, HFE has a unique subcellular localization in proximity to probable sites of iron absorption.

In HHC, there is an inappropriately high level of iron absorption, with increased transfer of iron from intestinal cells. Absorbed iron becomes tightly bound to transferrin in the bloodstream and enters cells after endocytosis of diferric transferrin bound to transferrin receptors on the cell surfaces. The internalized endosome becomes acidified and causes the release of iron from transferrin. It is believed that HFE interacts with the transferrin receptor and participates, along with transferrin, in the receptor-mediated endocytosis. The carboxy region of the α1 domain and the adjacent loop of HFE are the binding sites. The transferrin receptor is necessary for the assembly, transport and expression of the cell surface of HFE. This interaction likely decreases the number of transferrin binding sites, blocks the internalization of the transferrin receptor, decreases the affinity of the transferrin receptor for transferrin, and reduces the amount of iron taken by reducing the amount released from transferrin in the endosomes.

The C282Y disrupts the disulfide mutation bond in the α1 and α3 domains of HFE. This prevents the binding of $ß_2$-microglobulin, which is essential for the correct intracellular processing and transport of HFE to the plasma membrane. The mutant protein is retained in the endoplasmic reticulum and fails to undergo late Golgi processing. It is not expressed on the cell surface and is unable to participate in its normal function. The H63D mutation is located in the α1 domain and, unlike the C282Y mutation, does not interfere with $ß_2$-microglobulin binding. H63D HFE does not appear to have any significant difference from the wild type protein in its intracellular processing, transportation, and cell surface expression. The functional significance of this mutation is not clear. H63D HFE can form a complex with the transferrin receptor; however, it does not reduce

the affinity of HFE to transferrin to the same extent as wild type HFE.

Genotypic analysis (discussed in Chapter 11) of hemochromatosis is currently recommended for individuals with first-degree family members who have been identified with a mutation in the *HFE* gene, as well as in patients with evidence of iron overload, such as increased serum transferrin saturation, high serum ferritin levels, or excess iron staining on liver biopsy. Screening of the normal population by genetic analysis is not currently recommended. Transferrin saturation and unbound iron binding capacity are currently recommended for initial screening of individuals. Those individuals with normal values should be reassured and retested later if there is any clinical suspicion of hemochromatosis. *HFE* testing should include molecular analysis for both C282Y and H63D mutations as discussed in the molecular section.

Secondary Hemochromatosis

Secondary causes of iron overload principally result in deposition of iron within the reticuloendothelial system (also referred to as hemosiderosis). Damage to parenchymal tissues indistinguishable from inherited hemochromatosis, however, occurs with time as iron becomes deposited outside the saturated reticuloendothelial stores.

Increased intake of dietary iron may rarely result in iron overloading; more frequently, iron overloading is a complication of the anemias characterized by lifelong ineffective erythropoiesis. Examples include thalassemia major, inherited sideroblastic anemia, and some forms of congenital dyserythropoietic anemia. All display hyperabsorption of iron in response to an increased rate of erythropoiesis. A program of maintenance transfusion therapy may be necessary in these disorders, further complicating the picture of iron overload through the associated parenteral administration of iron in blood products.

Anemia of Chronic Disease

Anemia of chronic disease is the term used to describe the anemic state that is often seen in conjunction with a number of chronic nonhematologic disorders (Table 41.6). Considering the array of conditions with which anemia of chronic disease has been associated, it is not surprising to find that it follows closely behind iron deficiency as a cause of anemia.

TABLE 41.6. CONDITIONS ASSOCIATED WITH THE ANEMIA OF CHRONIC DISEASE

Chronic infections
Tuberculosis
Chronic fungal infections
Subacute bacterial endocarditis
Osteomyelitis
Pyelonephritis
Pelvic inflammatory disease
Chronic inflammatory disorders
Rheumatoid arthritis
Systemic lupus erythematosus
Sarcoidosis
Rheumatic fever
Neoplasms
Carcinoma
Malignant lymphoma

Beyond the descriptive association with chronic disease, anemia of chronic disease may also be characterized as anemia of impaired iron utilization. As such, it is defined by a reduction in both serum iron and serum transferrin. Hypoferremia typically outpaces the hypotransferrinemia, resulting in decreased transferrin saturation. These changes occur in the presence of adequate or increased iron stores, indicating sequestration of iron in the reticuloendothelial system.

Iron trapping is only one factor in the pathogenesis of anemia of chronic disease. Additionally, there is diminished erythrocyte survival time, which is further compounded by the inability of the bone marrow to compensate by increasing the rate of erythropoiesis. Each of these factors, in part, appears to be related to the sustained release of interleukin-1 (IL-1) found during chronic inflammation and neoplasia.

IL-1 triggers the release of lactoferrin, a glycoprotein found in the secondary granules of neutrophils, which competes with transferrin for the binding of iron. Unlike iron bound to transferrin, which is transported to tissues in need, lactoferrin-bound iron is delivered to the macrophage, where it is internalized and effectively trapped, preventing delivery to the developing erythron.

Decreased erythrocyte survival may, in part, be related to fever induced by IL-1 or to enhanced ingestion of erythrocytes by macrophages activated by IL-1. IL-1 activation of macrophages and T lymphocytes may also alter the microenvironment of the bone marrow, creating conditions unfavorable for a compensatory erythroid hyperplasia.

Clinical and Laboratory Features

Anemia of chronic disease typically develops within 1 to 2 months after the onset of a chronic illness and remains stable throughout the disease course. It is typically mild to moderate in severity, with the hematocrit ranging from 0.30 to 0.40. As such, the signs and symptoms of the underlying disease tend to overshadow the effects of the anemia.

The anemia is frequently normocytic and normochromic; however, microcytic hypochromic anemia may occasionally be encountered. Anisopoikilocytosis is usually minimal, in contrast to the findings in iron-deficiency and sideroblastic anemias. The reticulocyte count, when corrected for the degree of anemia, is normal or low. The bone marrow shows essentially normal numbers of erythroid precursors with no significant morphologic abnormalities. Iron stores appear adequate or increased in the marrow; however, the number of sideroblasts is decreased (less than 20%); this combination of findings is virtually diagnostic of anemia of chronic disease.

Anemia of chronic disease and iron-deficiency anemia have in common a diminished delivery of iron to the erythron; hence, both are characterized by decreased serum iron. The two disorders may be distinguished by quantitation of serum transferrin and ferritin. Transferrin is normal or decreased in anemia of chronic disease, in contrast to the elevated values found in iron deficiency. Transferrin saturation is typically decreased in both disorders; however, values are often less than 10% in iron deficiency, a level seldom reached in anemia of chronic disease. Ferritin

is normal or increased in anemia of chronic disease, reflecting adequate iron stores, whereas ferritin is decreased in iron deficiency. Caution must be exercised, however, in interpreting serum ferritin in the presence of inflammation or neoplasia because it is an acute phase reactant.

Summary: A Laboratory Approach to the Disorders of Heme Synthesis

The thalassemias, iron-deficiency anemia, sideroblastic anemia, and anemia of chronic disease must each be considered in the differential diagnosis of a microcytic hypochromic anemia. From the preceding discussion, however, it can be seen that the morphologic classification of anemia is not perfect; this is evident in the frequent presentation of anemia of chronic disease as a normocytic anemia and sideroblastic anemia as a normocytic or macrocytic anemia. Iron-deficiency anemia, sideroblastic anemia, and anemia of chronic disease are related beyond a common morphologic presentation; they share a defect in heme synthesis related to the deficiency, utilization, or transport of iron. Taking advantage of the altered handling of iron exhibited by these disorders, differentiation can often be achieved by simple quantitation of serum iron, transferrin, and ferritin. Decreased serum iron or ferritin, increased transferrin or soluble transferrin receptor, or decreased/absent marrow iron stores are consistent with iron-deficiency anemia. Because the defects in hemoglobinopathies and thalassemias are unrelated to iron metabolism, normal iron studies in the presence of a microcytic hypochromic anemia may suggest hemoglobinopathy or thalassemia, but iron overload may occur with frequent transfusion support.

Disorders of DNA Synthesis: Megaloblastic Anemia

The megaloblastic anemias are a group of disorders characterized by a reduced rate of DNA synthesis. RNA synthesis, its processing, and ultimately protein synthesis are, however, unaffected. The resultant dyssynchrony in nuclear (DNA) and cytoplasmic (RNA) development is the morphologic hallmark of the megaloblastic anemias. Deficiencies of vitamin B_{12} or folic acid are by far the most frequent etiologic factors.

Vitamin B_{12}

Metabolism

Vitamin B_{12}, or cobalamin, is composed of a corrin ring linked to a ribonucleotide through one of the nitrogens of its base. The corrin ring resembles the tetrapyrrole ring structure of heme; cobalt occupies the central position of the corrin ring analogous to the position occupied by iron in heme. The unsatisfied valence of the cobalt ion is filled in humans largely by CN(cyanocobalamin) or OH^- (hydroxycobalamin). These cobalamin species each have nutritional properties of B_{12}, yet have no biochemical activity. Rapid tissue conversion to the active coenzymes methylcobalamin and adenosylcobalamin occurs through the action of coenzyme synthetases.

Methylcobalamin and adenosylcobalamin are vital cofactors for two independent metabolic reactions in humans. Methylcobalamin is a coenzyme needed for converting N^5-methyltetrahydrofolate to tetrahydrofolate (FH_4) (Fig. 41.8). FH_4 is an essential element in a number of biochemical reactions; most important is its role as an intermediary coenzyme in DNA synthesis. Conversion of FH_4 to the folate coenzyme $N^{5,10}$–methylene FH_4 is essential for the conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), ultimately allowing integration of thymidine [as deoxythymidine triphosphate (dTTP)] into DNA. Therefore, in large part, the impairment of DNA synthesis seen with vitamin B_{12} deficiency is related to abnormalities in folate metabolism. This explains the shared manifestations of vitamin B_{12} and folic acid deficiency.

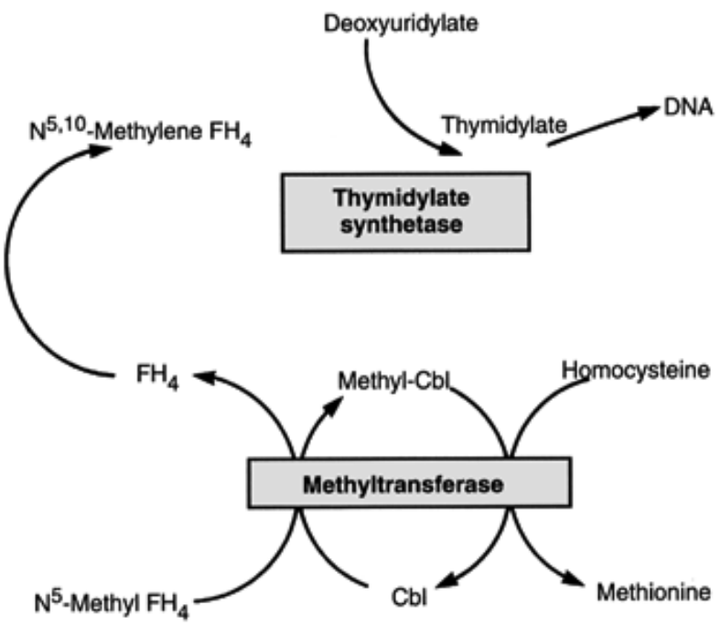


FIGURE 41.8. The role of methylcobalamin in methionine synthesis and its integral relationship to folate metabolism and DNA synthesis. FH_4, tetrahydrofolate; -Cbl, cobalamin. (From Williams W, et al. *Hematology*, 3rd ed. New York: McGraw-Hill, 1983:314, with permission.)

The inability to convert dUMP to dTMP, seen with both vitamin B_{12} and folic acid deficiency, alternatively results in the conversion of dUMP to deoxyuridine triphosphate (dUTP). The excess dUTP ultimately becomes incorporated into DNA in place of dTTP because DNA polymerase is unable to distinguish between these two nucleotides. The substitution of dUTP for dTTP results in errors in strand copying followed by attempts to edit the incorrectly inserted uridylate residue. Subsequent efforts to synthesize DNA correctly obviously fall short, as the requisite thymidylate is in short supply. Repeated cycles of fragmentation and attempted repair slow the synthesis of DNA and result in the characteristic nuclear chromatin pattern of megaloblastic anemia.

The other metabolically active form of cobalamin, adenosylcobalamin, is a coenzyme in humans essential to the conversion of methylmalonyl-CoA to succinyl-CoA, the basic fuel of the Krebs cycle (Fig. 41.9). Although adenosylcobalamin cannot be linked to DNA synthesis, it has been postulated, but not well established, that interruption of this reaction may play a role in the development of the neurologic changes seen in vitamin B_{12} deficiency. Inhibition of the conversion of methylmalonyl-CoA to

succinyl-CoA results in an accumulation of the methylmalonyl-CoA precursor propionyl. In high concentration, propionyl-CoA may replace acetyl-CoA as the usual primer for synthesis of even-chain fatty acids. Consequent errant production and insertion of odd-chained fatty acids into lipid membranes may result in lipid membrane abnormalities such as demyelination and subsequent neurologic impairment. An alternate hypothesis suggests that the impaired conversion of methionine owing to the deficiency of methylcobalamin results in a deficiency of *S*-adenosylmethionine (SAM). SAM is important for certain transmethylation reactions, including some that may be important in the maintenance of myelin.

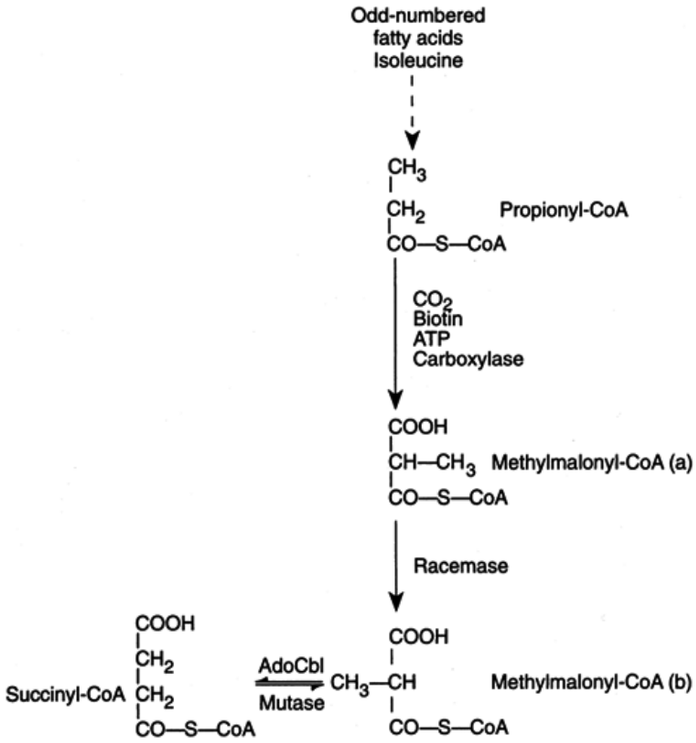


FIGURE 41.9. Adenosylcobalamin as a cofactor in the conversion of methylmalonyl-coenzyme A to succinyl-coenzyme A. AdoCbl, adenosylcobalamin. (From Williams W, et al. *Hematology*, 3rd ed. New York: McGraw-Hill, 1983:315, with permission.)

Vitamin B_{12} Absorption

Synthesized by many bacteria and some molds, vitamin B_{12} is found only in foods of animal origin such as meat, eggs, cheese, and milk. It has been estimated that the minimal daily requirement of vitamin B_{12} is 0.6 to 1.2 µg/day, an amount easily achieved in Western diets, which commonly contain 5 to 30 µg of vitamin B_{12}.

Vitamin B_{12} is released from food sources by the digestion of animal proteins. Free B_{12} is then competitively bound either to intrinsic factor (IF), a high-affinity glycoprotein produced by the parietal cells of the stomach, or to R proteins, a class of B_{12} binding proteins so designated because of their rapid electrophoretic mobility. R proteins are found in a variety of sites, including gastric juice and saliva, and serve to bind the majority of newly ingested B_{12}. Under normal circumstances, however, the R protein–B_{12} complex is short-lived; degradation of R protein by pancreatic proteases occurs in the proximal small intestine, liberating B_{12}, which then avidly binds to IF. IF greatly facilitates the absorption of B_{12}, providing direct and safe transport through the gastrointestinal tract to specific receptors located on the mucosa of the terminal ileum. Once the IF-B_{12} complex is bound, B_{12} is internalized by a poorly understood process.

Not to be overlooked is the enterohepatic circulation of B_{12}. More than 50% of the B_{12} entering the intestine each day arrives through this route. The majority is reabsorbed through IF-dependent mechanisms.

After absorption in the ileum, vitamin B_{12} is released to the circulation, where it is bound by transcobalamin (TC), the B_{12} transport protein. TC exists in three forms, designated TC I, TC II, and TC III; TC II is the major physiologic carrier of B_{12}, although normally binding only 10% to 25% of the plasma concentration of B_{12}. TC II has great capacity for rapid B_{12} transport to appropriate receptors at sites of need (dividing cells) or storage (liver). When bound by the receptor, the B_{12}–TC II complex is internalized by endocytosis, B_{12} is liberated, and TC II is degraded by lysosomal proteases.

TC I and TC III belong to the family of T proteins. Although most of the plasma B_{12} concentration is bound to TC I, this relatively stable complex has little capacity for transfer of B_{12} to tissue sites. TC III is largely synthesized in granulocytes and, although capable of binding B_{12} *in vitro*, is not engaged in the transport of endogenous B_{12} to tissue cells. Under normal circumstances, TC I is approximately 50% saturated with B_{12}, whereas less than 5% of TC II is saturated. The sum of the unsaturated TC I and TC II is referred to as the unsaturated B_{12} binding capacity.

Quantitative Assays

Serum levels of vitamin B_{12} are largely determined using isotope dilution techniques, although cobalamin-dependent microbial

assay systems may still be used. Universal use of IF as a binding protein in radioisotope assays eliminates binding of cobalamin analogs and yields values that compare favorably with those obtained with a microbiological system. Normal serum levels of B_{12} typically are in the range of 200 to 900 pg/mL. The serum B_{12} concentration is but a minute fraction of the total body stores of B_{12}, estimated at 1 to 5 mg, the liver being the primary site of storage.

Abnormalities of B_{12} binding proteins have been reported in a variety of disease states. Quantitation of both serum B_{12} and (unsaturated) B_{12} binding capacity are often cited as a tool in the evaluation of the chronic myeloproliferative disorders. Increases in one or both parameters largely reflect elevations of TC I and TC III liberated from the breakdown of increased numbers of granulocytes in these disease states.

Vitamin B_{12} Deficiency

Five basic mechanisms may lead to B_{12} deficiency: inadequate intake, increased requirement, defective absorption, defective transport, and defective enzyme activity. Defective transport owing to an abnormality of a TC carrier protein is rare, as are abnormalities of the coenzyme synthetases needed for conversion of B_{12} to the active forms of methylcobalamin or adenosylcobalamin. Also uncommon, given the body's reserves of B_{12}, are deficiencies occurring during transient periods of increased B_{12} requirement, e.g., during growth and development, pregnancy, or lactation. Similarly, an inadequate dietary intake of B_{12} is largely limited to strict vegetarians who abstain from all animal proteins, including milk and eggs. Assuming a previously healthy diet, strict avoidance of foods replete with B_{12} for a period of 3 to 5 years is required before achieving complete exhaustion of the body's B_{12} stores. For these reasons, deficiencies of B_{12} are most frequently encountered as a result of defective absorption.

Pernicious Anemia. Pernicious anemia (PA) is the most common cause of impaired B_{12} absorption and is the result of an acquired failure to secrete IF by the gastric parietal cells. The markedly diminished or absent secretion of intrinsic factor is typically associated with achlorhydria; both findings are related to the universal presence of gastric atrophy. PA is most commonly found in individuals of northern European ancestry, although no racial or ethnic group is immune. Typically, the disorder presents after the age of 40. The etiology of the disorder remains speculative. Circumstantial evidence suggests a hereditary basis for the disorder, although neither a mode of inheritance nor a link to any genetic marker is currently known. An autoimmune basis for pernicious anemia has also been proposed based on the frequent finding of autoantibodies in patients with PA and the histologic findings of a lymphoplasmacytic infiltrate in the gastric biopsies of patients with PA. Two types of autoantibodies are found in patients with PA. Approximately 90% of patients have antibodies to parietal cell cytoplasm in their serum. The exact role, if any, of these antibodies in the pathogenesis of PA is uncertain. Antiparietal cell antibodies have also been found in individuals with atrophic gastritis unassociated with PA, in persons with Hashimoto's thyroiditis, and in a small number of normal individuals. Serum antibodies to intrinsic factor have been detected in 55% of patients, whereas assays of gastric juice allow detection of anti-IF antibodies in 75% of patients with PA. Anti-IF antibodies are of two types: blocking antibodies, which block the attachment of B_{12} to IF, and binding antibodies, which prevent the absorption of B_{12} in the ileum. Although these antibodies have shown functional impairment of IF activity *in vivo*, their temporal sequence appears to postdate the onset of disease, thereby negating their role as a primary etiologic agent in this disease process.

Variant Forms of PA. Two rare variant forms of PA are found. Juvenile PA is distinguished from classic PA solely by its clinical appearance in adolescence or early adulthood. A rare congenital form of PA has also been described. Patients with congenital PA have either an absence of IF or a functionally aberrant form of IF. In contrast to other forms of PA, the defect in IF is selective; the appearance and function of the gastric mucosa in all other aspects is normal. The congenital form of PA is inherited as an autosomal recessive defect with hematologic manifestations appearing in infancy. Like classic PA, both juvenile and congenital PA respond to the administration of IF.

Gastrectomy. B_{12} deficiency typically develops 3 to 5 years after total gastrectomy as B_{12} absorption ceases and the body's B_{12} stores are gradually depleted. B_{12} deficiency is less predictable after subtotal gastrectomy; often appearing one to two decades after surgery, it may be related to the slow development of postoperative gastritis and gastric atrophy.

Pancreatic Insufficiency. Loss of the exocrine function of the pancreas, regardless of cause, may result in B_{12} deficiency. Pancreatic proteases are integral to the digestion of gastric R proteins, which compete with IF for B_{12}. Intact R proteins inhibit formation of the B_{12}-IF complex, thereby diminishing absorption in the terminal ileum.

Intestinal Malabsorption. Several small intestinal abnormalities may result in diminished B_{12} absorption. Surgical resection or diseases of the terminal ileum, such as Crohn's disease, effectively remove the site of B_{12} absorption. Malformations of the intestinal tract may predispose to the development of localized pockets of bacterial overgrowth. These proliferative microorganisms may effectively deprive the host of B_{12}. Similarly, competition for B_{12} may be encountered with intestinal infestation by the fish tapeworm *Diphyllobothrium latum*.

Laboratory Evaluation of B_{12} Deficiency. Several tests may be employed in the laboratory investigation of megaloblastic anemia. Serum or plasma quantitation of B_{12} is the usual method of detecting a vitamin B_{12}-deficient state. Less direct and less specific measures of a B_{12} deficiency state, such as urinary excretion of methylmalonic acid or the deoxyuridine suppression test, are seldom needed. After B_{12} deficiency is established, the etiology must be determined. The most frequent test employed to determine whether the patient lacks IF is the Schilling test. Shortly after oral administration of 0.5 to 2.0 µg of radiolabeled B_{12}, a large intramuscular flushing dose of nonlabeled B_{12} is administered, which will saturate B_{12} binding sites. When saturated, the orally ingested radiolabeled B_{12} absorbed by the ileum will be excreted

in the urine. A 24-hour urine collection normally reveals greater than 7% absorption of a 1-µg dose. If excretion is found to be diminished, the test is repeated 1 week later. The subsequent test, however, provides purified IF with the orally administered B_{12}. Improved absorption after administration of IF implies a deficiency of IF, whereas continued poor absorption signifies intestinal malfunction. Unfortunately, the Schilling test has two common pitfalls: inadequate urine collections and underlying renal dysfunction.

Antiparietal cell and anti-IF antibodies have been previously described and may be used as additional tests to help substantiate the diagnosis of PA.

Folic Acid

Metabolism

Folic acid, or pteroylglutamic acid, consists of three parts: pteridine, *p*-aminobenzoic acid, and glutamic acid. It serves as the parent for a large family of compounds having similar nutritional value to which the generic term folate is applied. Folates are widely distributed in a variety of foods, being synthesized by higher plants and microorganisms. Green leafy vegetables, fruits, and dairy products provide the greatest sources of folate in the usual Western diet. Unfortunately, folates are thermolabile, so a large portion of dietary folates may be lost in food preparation, particularly in boiling. A normal balanced diet, however, has no difficulty in meeting the recommended daily allowance for adults of 200 to 400 µg. In sharp contrast to vitamin B_{12}, the body's stores of folate are limited. After cessation of dietary intake, a decline in blood folate levels can be detected within a few weeks, and megaloblastic anemia may ensue in 3 to 5 months.

Most food folates are in the form of reduced methyl or formyl derivatives; approximately 90% are polyglutamates. Absorption occurs predominantly in the upper small intestine. When ingested in the monoglutamate form, folate is readily absorbed. The more prevalent polyglutamates, however, must first be converted to the monoglutamate form through the action of deconjugating enzymes found in the brush borders of intestinal epithelial cells as well as within intracellular lysosomes.

After assimilation by the intestinal epithelial cell, folate is converted to N^5-methyl TH_4, in part through the action of dihydrofolate reductase; it is in this form that folate is found in circulation. Cellular incorporation occurs as N^5-methyl TH_4 relinquishes its methyl group to homocysteine, forming methionine and TH_4, a reaction requiring methylcobalamin as a coenzyme (Fig. 41.8). Intracellular TH_4 may then proceed as a precursor to DNA synthesis or may be converted to an intracellular polyglutamate storage form through the action of an adenosine 5′-triphosphate (ATP)–dependent synthetase. The storage form may undergo later deconjugation ultimately to yield TH_4 at critical times in DNA synthesis. This interconversion of the mono- and polyglutamate forms appears to play a role in the regulation of DNA synthesis.

Folate Deficiency

Folate deficiency may arise from inadequate intake, impaired absorption, increased requirements, or defective folate metabolism. The most frequent cause of folate deficiency is a folate-poor diet. With limited folate reserves, the effects of an unbalanced diet are soon felt by chronic alcoholics, the impoverished, and the “tea and toast” elderly. Diets associated with the development of folate deficiency are characterized by a predominance of starches and grains with relatively little animal protein or fresh green vegetables. Megaloblastic anemia arising from folate deficiency in the setting of general malnutrition may also be complicated by diminished reserves of other hematinic agents.

Intestinal malabsorption is also a common cause of folate deficiency. Although folate absorption occurs principally in the proximal small intestine, receptor sites are found throughout the small bowel. Consequently, for megaloblastic anemia to occur secondary to impaired folate absorption, diffuse small intestinal disease is generally required. Tropical sprue and adult celiac disease (nontropical sprue or gluten-sensitive enteropathy) both bring about such a state of generalized malabsorption, resulting in a deficiency of a variety of nutrients, including folate. Widespread involvement of the small bowel by Crohn's disease may also result in folate deficiency.

Increased folate requirements are seen in several physiologic and pathologic conditions exhibiting an overall increase in cell proliferation and DNA synthesis. Folate deficiency typically occurs during the last trimester of pregnancy, when folate requirements increase approximately fivefold. After delivery, lactation brings further demands for additional folate. Infancy is marked by the greatest rate of growth, and hence increased folate intake is required. Pathologic hyperproliferative states also require additional folates and include a variety of chronic hemolytic anemias: sickle cell anemia, hereditary spherocytosis, thalassemia, and chronic autoimmune hemolytic anemias. This is not surprising if one considers that normal hematopoiesis is responsible for approximately half of total DNA synthesis. Rapidly dividing tumors such as leukemias, high-grade lymphomas, and small cell carcinoma may also usurp the body's folate supply.

Impaired folate metabolism is the desired effect of certain drugs. Chemotherapeutic agents such as methotrexate and aminopterin interfere with the conversion of folate to tetrahydrofolate. Consequently, administration of these agents causes megaloblastic changes in the marrow. More commonly prescribed, but much less frequently the cause of severe megaloblastic changes, are such antifolates as trimethoprim and triamterene. The induction of megaloblastic changes as a side effect of drug therapy is, however, not limited to the antifolates. Several other chemotherapeutic agents that interfere with DNA synthesis (antimetabolites, alkylating agents) are also known to induce similar changes (Table 41.7).

TABLE 41.7. DRUG-INDUCED SUPPRESSION OF DNA SYNTHESIS

Folate antagonists
Methotrexate
Aminopterin
Trimethoprin
Pyrimethamine
Triamterene
Antimetabolites
5-Fluorouracil
Hydroxyurea
Cytosine arabinoside
Alkylating agents
Cyclophosphamide
Nitrogen mustard
Chlorambucil
Busulfan
Melphalan
Nitrous oxide

Aside from these drug-induced defects in folate utilization, inborn errors of folate metabolism have been reported. These include dihydrofolate reductase deficiency and formiminotransferase deficiency, as well as congenital defects of folate absorption.

Laboratory Evaluation of Folic Acid Deficiency

Both microbiological and radioisotopic assays are available for the quantitation of serum folate. Serum levels of folate are, however, quite labile and sensitive to short-term variation in vitamin intake. Quantitation of red cell folate provides a better long-term measure of true tissue folate levels because folate metabolism

ceases in the developing red cell after nuclear extrusion. Thereafter, folate levels remain stable throughout the life span of the erythrocyte. Changes in red cell folate come about over time as red cells are released from a marrow environment of differing folate concentrations. Megaloblastic anemia caused by folate deficiency typically exhibits serum folate concentrations less than 3 ng/mL and red cell folate concentrations less than 100 ng/mL.

The following cautions must be exercised in interpreting the results of serum or red cell folate concentrations. Because serum folates are greatly influenced by short-term dietary variation, quantitation should be obtained before ingestion of a balanced hospital diet replete with folate. Similarly, red cell folate may be affected by the administration of transfusions, and assays should therefore be performed on samples obtained before the administration of red cell products. Finally, because B_{12} is a requisite cofactor for the uptake and utilization of folate by the cell, in its absence, folate is trapped in the serum. Therefore, in vitamin B_{12} deficiency, serum folate may be elevated and red cell folate decreased.

Clinical Features of Megaloblastic Anemia

Vitamin B_{12} deficiency is most frequently insidious in onset, likely reflecting the prolonged period of time required to deplete the body's B_{12} stores. Given our capacity to compensate over time, it is not surprising that patients often present with few signs or symptoms despite a frequently marked degree of anemia. Classically, patients with PA have been described as pale, but not wasted, with lemon-yellow skin reflecting a combination of anemia and mild hyperbilirubinemia. Blue eyes, a broad face and chest, and blond or prematurely gray hair complete the stereotype. Some of these features may reflect the Northern European ancestry of many patients with PA. It is important to remember, however, that no race or gender is immune from this disorder; consequently, many of these features may not be present.

A diagnostic triad of weakness, sore tongue, and paresthesias is cited as the classic symptom complex of vitamin B_{12} deficiency. Indeed, weakness reflects anemia. Glossitis, present in approximately half of patients, is the most visible indicator of generalized epithelial atrophy. Particularly vulnerable is the gastrointestinal tract, where rapid turnover of epithelial cells is a normal event. Gastric atrophy, occurring either as the *sine qua non* of PA or as a secondary complication of other forms of B_{12} deficiency, coupled with villous atrophy of the small bowel, may lead to symptoms of heartburn, bloating, loss of appetite, or diarrhea. Paresthesias are among many neurologic manifestations complicating B_{12} deficiency. Neurologic changes chiefly affect the white matter of the posterior and lateral columns of the spinal cord, leading to symmetrical numbness and tingling of the hands and feet, diminished vibratory or position sense, progressive weakness, and an unsteady or ataxic gait. The terms *subacute combined degeneration* or *combined system disease* are employed to encompass these abnormalities of multinerve pathways. The severity and rate of progression of the neurologic changes do not always correlate with the degree of anemia or other findings of B_{12} deficiency. Consequently, familiarity with this complication of vitamin B_{12} deficiency is mandatory if it is to be recognized and arrested before serious irreversible neurologic changes occur.

The manifestations of folic acid deficiency are often tainted by the myriad of health problems associated with chronic alcoholism and poverty. In fact, associated health problems may be the initiating reasons for seeking medical attention. Hematologic manifestations are similar to those found in megaloblastic anemia attributable to vitamin B_{12} deficiency. The neurologic changes typical of B_{12} deficiency, however, are absent.

Peripheral Blood and Bone Marrow Findings

The peripheral blood and bone marrow findings in folic acid and vitamin B_{12} deficiency are indistinguishable. Recognizing that DNA synthesis is critical to all developing cell lines, it follows that patients with full-blown megaloblastic anemia typically present with pancytopenia. Red cell indices classically reveal macrocytosis; a MCV greater than 120 fl is common. The MCH is elevated, whereas the MCHC is normal. The reticulocyte count is normal or decreased, reflecting the inability of the bone marrow to produce red cells. Unfortunately, however, not all patients present with pancytopenia and macrocytosis. Although anemia is a fairly universal finding, the presence and degree of leukopenia and thrombocytopenia are variable. Superimposed infection may lead to an elevated white blood count at presentation. Likewise, the MCV may appear normal with coexisting iron deficiency or a silent thalassemia trait.

Examination of the peripheral blood smear reveals a number of important clues to the diagnosis (Fig. 41.10). Macroovalocytes and hypersegmented neutrophils are invariably present. Anisopoikilocytosis of moderate to marked degree is present including ovalocytes, teardrop forms, schistocytes, and spherocytes. Basophilic stippling and Howell-Jolly bodies may be seen; two or more Howell-Jolly bodies may be found per cell, unlike in postsplenectomy states. With severe anemia, nucleated red blood cells may be present that exhibit characteristic megaloblastic nuclear changes.

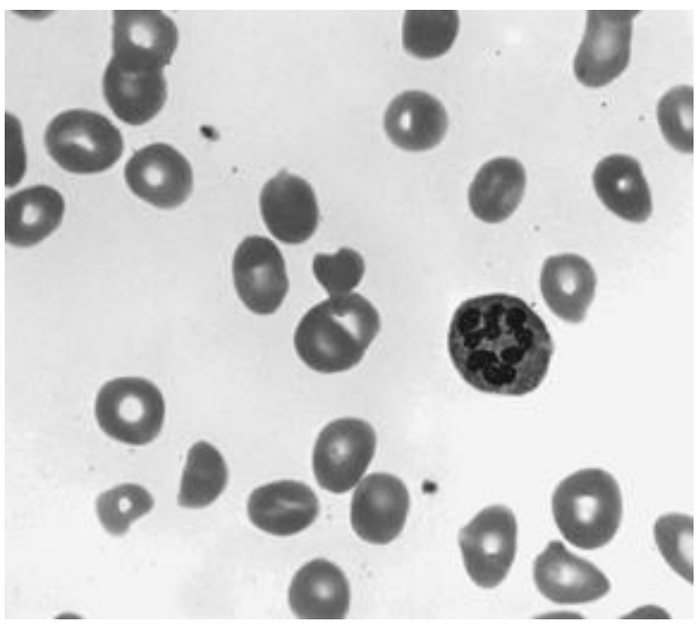


FIGURE 41.10. Megaloblastic anemia is characterized by hypersegmented neutrophils and the presence of well-hemoglobinized macrocytes and macroovalocytes.

Bone marrow aspirate smears reveal an intensely cellular marrow with increased numbers of erythroblasts (Fig. 41.11). Nuclear cytoplasmic dyssynchrony may be found in all myeloid cell lines. The changes in the red cell series, however, often appear

most striking. Changes are most evident in the later stages of erythroid maturation, when cellular enlargement and open fenestrated or sievelike nuclear chromatin contrasts sharply with the smaller size and shrunken, condensed nucleus of the late polychromatophilic and orthochromatic normoblasts. Intramedullary destruction is morphologically manifested as nuclear karyorrhexis or fragmentation. Bizarre nuclear budding or lobulation may also be seen. Although megaloblastic changes in developing neutrophils tend to be overshadowed by those of the erythroid series, they offer important clues to both early and partially treated megaloblastic anemia. Giant bands and metamyelocytes are frequent; as the name implies, these cells are large and often exhibit an increased nuclear:cytoplasmic ratio. Nuclear chromatin appears less mature than anticipated for the stage of maturation indicated by nuclear shape. Premature nuclear segmentation is also common as these giant bands exhibit nuclear folding and twisting as if to form segments of a mature neutrophil. Such changes appear before those of the erythroid series and persist even after normoblastic erythroid maturation returns after therapeutic intervention. Abnormalities may also be seen in the megakaryocytes. These include complex hypersegmentation and a similar open or fenestrated nuclear chromatin pattern.

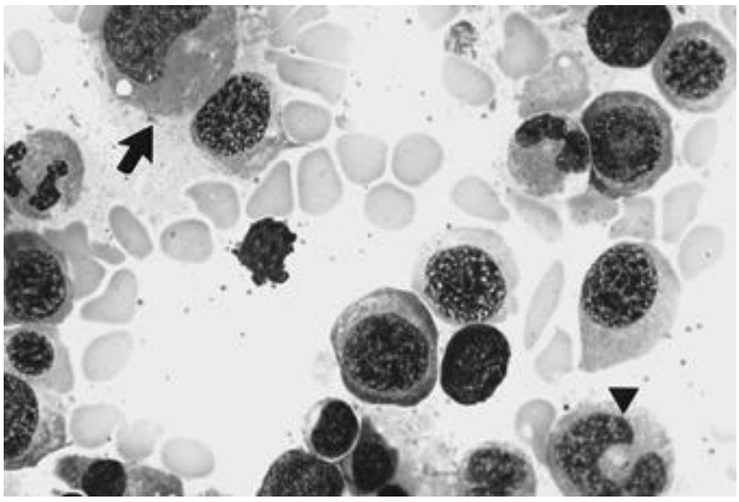


FIGURE 41.11. Bone marrow aspirate smear in megaloblastic anemia demonstrates numerous large erythroblasts having a fenestrated nuclear chromatin. The nuclear appearance has been likened to that of a piece of salami. Nuclear changes are also evident in the developing neutrophils: giant metamyelocytes (*arrow*) and giant bands (*arrowhead*).

Perhaps no other disorder better exemplifies the hazards of interpreting bone marrow trephine biopsies without the benefit of aspirate smears than does megaloblastic anemia. The trephine biopsy is virtually always hypercellular, in many cases approaching 100%. The lag in nuclear maturation results in an accumulation of erythroblasts in the early to middle stages of maturation. This, coupled with the immature appearance of the nuclear chromatin, imparts the picture of a marrow overrun with primitive hematopoietic cells. The temptation to regard this picture as one of acute leukemia can be overcome by careful examination, which demonstrates the presence of all three cell lines, evidence of cellular maturation, and characteristic giant bands.

Abnormal maturation leads to intramedullary destruction of all cell lines, resulting in the paradox of ineffective hematopoiesis: pancytopenia in the presence of a hypercellular bone marrow. Intramedullary red cell destruction results in mild hyperbilirubinemia, increased LDH, and increased serum iron. Red cells are rich in LDH_1 and LDH_2; LDH_1 predominates in immature red cells, which may result in a “flip” of the normal LDH_1:LDH_2 ratio. Iron is liberated during the catabolism of red cells at a rate greater than it can be utilized. Consequently, serum iron increases, transferrin saturation increases whereas synthesis decreases, and iron stores become replete. Marrow sideroblasts may increase in number, and types II and III (ringed) sideroblasts may be found.

Differential Diagnosis

The combined findings of pancytopenia, macrocytosis, and hypersegmented neutrophils are virtually diagnostic of megaloblastic anemia. Individually, however, these findings are not specific, each being associated with a host of disorders largely unrelated to B_{12} or folic acid metabolism. Pancytopenia generally implies diminished marrow precursors, ineffective hematopoiesis, or hypersplenism. Macrocytosis may occur in chronic liver disease, hypothyroidism, pregnancy, normal neonates, marrow failure, chronic hemolytic states with elevated reticulocyte counts, myelodysplastic syndromes, and after administration of a variety of cytotoxic drugs. Hypersegmented neutrophils, although uncommon outside of megaloblastic anemia, may be seen as a congenital anomaly and in iron deficiency, sepsis, renal disease, and chronic myeloproliferative disorders.

Megaloblastic changes in one or more cell lines of the bone marrow are frequently seen after the administration of a variety of chemotherapeutic agents. Most often these changes affect the developing neutrophils to the greatest degree and pose no diagnostic dilemma given the patient's history. Megaloblasts may be seen as one of many changes in the myelodysplastic syndromes or in erythroleukemia. The clinical history, the absence of B_{12} or folic acid deficiency, and the presence of dysgranulopoiesis, dysmegakaryopoiesis, and/or increased numbers of myeloblasts will all help separate these neoplastic conditions from megaloblastic

anemia. Several inborn errors may also lead to megaloblastic changes or frank megaloblastic anemia, including congenital dyserythropoietic anemia, hereditary orotic aciduria, and Lesch-Nyhan syndrome.

Management

Therapy for megaloblastic anemia is aimed both at replenishing the body's stores of vitamin B_{12} or folic acid and at documenting and correcting the underlying cause. Parenteral administration of B_{12}, typically on a monthly basis, is required therapy for PA and malabsorption states, which are not amenable to medical intervention. Folic acid deficiency, most often dietary in nature, is amenable to oral administration of folic acid. Folic acid preparations for parenteral administration are available if required.

Improvement in the patient's hematologic status is initially evident within 3 to 5 days by an increase in the reticulocyte count, which typically reaches a maximum at 10 days. A prompt decrease in serum iron is noted as effective red cell production begins. Amelioration of megaloblastic changes in the red cell series is largely complete within 24 to 48 hours, whereas changes in the neutrophil line persist in the peripheral blood and bone marrow for as long as 10 days. Within 1 to 2 months, all peripheral blood parameters should return to normal.

Neurologic deficits are arrested after the institution of appropriate therapy. Regression is, however, not always complete. Mild deficits of short duration are more apt to resolve than are chronic or severe impairments. It is important to diagnose the cause accurately and to administer the correct hematinic agents appropriately in patients with megaloblastic anemia. This is underscored by the progression of neurologic changes that occurs in B_{12} deficiency inappropriately treated with folic acid despite partial hematologic correction.

Bone Marrow Failure: Inadequate Erythroid Precursors

Included in this section are the anemias that result from an inadequate number of erythroid precursors occurring either as a selective phenomenon or, more frequently, as part of the trilineage suppression that results from a reduction in pluripotent hematopoietic stem cells.

Aplastic Anemia

Aplastic or hypoplastic anemia is typically an acquired defect in pluripotent stem cells, although defects in growth factors, immunoregulatory signals, or the marrow microenvironment have also been proposed as possible mechanisms. Regardless of the etiology, the net effect is a reduction in erythroid, granulocytic/ monocytic, and megakaryocytic cell lines in the bone marrow and their progeny in the peripheral blood. The incidence of aplastic anemia varies worldwide; in the Western nations, five to 10 new cases per million general population are estimated to occur annually. Cases are approximately equally divided into idiopathic and secondary forms, the latter composed largely of chemical or drug injury, radiation exposure, or infection-associated aplasia.

Etiology

Before labeling aplastic anemia as idiopathic, careful exclusion of known etiologic factors must be undertaken. This is particularly true in cases occurring in childhood, adolescence, or early adulthood, when a putative cause is found more frequently than in adulthood.

Drug-Related Injury. Chemical- or drug-related injury to the marrow may arise either as a dose-dependent or an idiosyncratic phenomenon. Marrow ablation is the desired dose-dependent effect of many chemotherapeutic agents. Similar dose-dependent marrow aplasia is encountered as an untoward side effect of exposure to chemicals such as benzene, its derivative trinitrotoluene, arsenic, or alcohol. Hematopoietic suppression associated with heavy alcohol intake is typically self-limited and reversible after cessation of drinking, unlike that found with benzene, trinitrotoluene, and arsenic, which often progress to fatal aplasia. Chronic benzene and arsenic exposure are both associated with the development of dysplasia, largely marked by bizarre megaloblastic changes of the erythroid series. In the case of benzene exposure, such dysplastic changes may progress and evolve to acute myelogenous leukemia (AML); not unexpectedly, the frequency of erythroleukemia is increased compared with that found in *de novo* AML.

Chloramphenicol is perhaps the most widely publicized example of an idiosyncratic drug reaction leading to marrow aplasia. This adverse side effect cannot be related to drug dose or route of administration. Likewise, its occurrence and the rapidity of onset cannot be predicted. This mainly fatal complication has no association with the mild, reversible, dose-dependent erythroid suppression known to occur in approximately half of individuals exposed to large doses of chloramphenicol. Similar idiosyncratic reactions are encountered with quinacrine, phenylbutazone, carbonic anhydrase inhibitors, gold, and a variety of anticonvulsants, including phenytoin, mephenytoin, trimethadione, methsuximide, and carbamazepine. The effect of these agents on the three hematopoietic cell lines is not always uniform, and in some cases selectively results in red cell aplasia, agranulocytosis, or thrombocytopenic purpura.

Ionizing Radiation. The hematopoietic system, like all body cells, is susceptible to damage by ionizing radiation. The untoward effects of radiation are dependent on the nature and quantity of radiant energy, the duration of exposure, and the mitotic rate of the exposed cell population. It is the high mitotic activity of the bone marrow that makes it one of the most vulnerable organ systems to the toxic effect of radiation. Aplastic anemia may ensue after a single exposure to large but sublethal doses of radiation, as well as repeated or long-term exposure to moderate doses.

Infection. Several viruses have been associated with myelosuppression. Most notable is the rare, unexplained, and often fatal aplasia that occurs in association with hepatitis. This unpredictable sequela is most frequently associated with non-A and non-B hepatitis and exhibits no relationship to the severity of liver disease. The time to onset is variable, with most cases becoming manifest approximately 6 weeks after clinically evident

hepatic disease. Human parvovirus (serotype B19) is now recognized as a cause of aplastic anemia. This virus, as well as several others, is responsible for myelosuppression, most frequently manifesting in individuals with an underlying chronic hemolytic anemia. Although granulocytes and platelets may be reduced, a profound exacerbation of the chronic anemia is the major cause of morbidity in these cases. Because of the shortened erythrocyte survival inherent in the underlying hemolytic state, affected individuals are exquisitely sensitive to the abrupt cessation of erythropoiesis caused by the selective cytotoxic invasion of erythroblasts by the parvovirus. Such temporary erythroblastopenia is typically preceded by or associated with an upper respiratory or flulike illness. Although short-term transfusional support may be necessary, these aplastic crises are generally self-limited.

Idiopathic. Aside from the direct toxic effect of various drugs, chemicals, and radiant energy, the pathogenesis of aplastic anemia is not fully understood. Evidence supporting an immunologic mechanism has emerged from a number of observations linking aplasia with disorders of immunoregulation. Support for this concept has been found *in vitro* through alterations in marrow lymphoid populations and *in vivo* by the success in certain individuals of antilymphocyte or antithymocyte globulin as a therapeutic modality.

Clinical Features

The onset of aplastic anemia is variable, in part dependent on etiologic factors. Signs and symptoms of generalized bone marrow failure are evident: pallor, weakness, fatigue, and dyspnea on exertion are associated with anemia; infections and fever are associated with granulocytopenia; and petechiae, ecchymoses, or overt hemorrhage may be associated with thrombocytopenia.

Peripheral Blood and Bone Marrow Findings

Pancytopenia is the rule in the peripheral blood; however, the decrease in erythrocytes, leukocytes, and platelets is often not proportional. In part, this may be related to the susceptibility of the marrow cell lines to the injurious agent. Erythrocytes are usually normochromic and normocytic with little variation in size and shape, although mild macrocytosis may be evident in some cases. Neutrophils are significantly decreased, and lymphocytes predominate on examination of the peripheral smear. Dysplastic features are not common in any cell line unless there has been exposure to toxins such as benzene or arsenic.

The bone marrow trephine biopsy shows a marked reduction in cellularity, largely replaced by fat. Foci of increased cellularity may be found scattered throughout the marrow, most often dominated by erythroblasts. Such hot spots of hematopoiesis may be misleading when aspirated or when only a small core of tissue is available for examination. This underscores the need for a generous sample of tissue in cases of suspected marrow aplasia.

Aspirate smears typically reflect the marked diminution in hematopoietic elements. Lymphocytes and plasma cells predominate; increased numbers of mast cells and histiocytes, the latter often laden with iron or particulate debris, are also frequently seen. Mild cytoatypia may be seen when foci of hematopoiesis are aspirated; however, this is never a striking feature unless exposure to toxins has occurred.

Treatment

The clinical course of patients with aplastic anemia depends on the etiology, the severity of the pancytopenia, the age of the patient, and the response to therapy. In some instances, marrow recovery may ensue after discontinuation of drugs known to act as marrow suppressants. Similarly, removal from the source of a known environmental toxin may initiate marrow recovery. Transient episodes of myelosuppression are amenable to short-term transfusional support. All too often, however, marrow aplasia is severe and irreversible, requiring long-term medical intervention.

Therapy with a number of drugs has met with varying success. Corticosteroids, androgens, lithium carbonate, antithymocyte globulin, and cyclophosphamide have been utilized as marrow stimulants or as immunosuppressive agents aimed at reversing putative immune-mediated aplasia. Repeated transfusions can obviously provide needed red cells and platelets but do little to offset granulocytopenia. Multiple transfusions are, however, complicated by the associated morbidity of iron overload and the risk of posttransfusion infectious complications. Additionally, in young patients with HLA-matched bone marrow donors, repeated transfusions have an adverse effect on bone marrow transplantation, the treatment of choice for this group of patients.

Pure Red Cell Aplasia

As the name implies, pure red cell aplasia is characterized by anemia secondary to a marked reduction in the number of red cell precursors in the marrow. Granulocytes and megakaryocytes are unaffected. To some degree, the aplastic crises known to complicate various chronic hemolytic states can be viewed as an acute form of pure red cell aplasia. Similarly, young children may experience a transient arrest of erythropoiesis after a viral illness (transient erythroblastopenia of childhood). Idiosyncratic drug reactions may also selectively involve the red cell series.

Many cases of pure red cell aplasia are associated with a thymoma; in conjunction, some patients also exhibit myasthenia gravis. Thymectomy will restore normal hematopoiesis in a subset of these patients. Similar to aplastic anemia, an immunologic mechanism has been proposed as a pathogenetic factor in pure red cell aplasia. Accordingly, therapy has included immunosuppressive agents as well as transfusional support.

Constitutional and Hereditary Disorders Associated with Aplastic Anemia

Several hereditary or constitutional syndromes are associated with aplastic anemia or the variant pure red cell aplasia. Manifestations are apparent in childhood and consist of hematologic abnormalities as well as a variety of other anomalies, including skeletal malformations, pigmentary changes of the skin, retarded mental and sexual development, and renal abnormalities.

Fanconi's anemia is inherited as an autosomal recessive disorder. Patients exhibit progressive pancytopenia and marrow aplasia during the first decade of life. Associated findings include increased levels of fetal hemoglobin and expression of i-antigen, reflecting reversion to fetal hematopoiesis. Like ataxia telangiectasia

and xeroderma pigmentosum, Fanconi's anemia is a disorder of defective DNA repair. Cytogenetic analysis reveals evidence of chromosome instability including chromosomal breaks and sister chromatid exchange. Not surprising, therefore, is termination in AML in approximately 10% of cases of Fanconi's anemia.

Diamond-Blackfan anemia predominantly affects the red cell line and may therefore be viewed as a constitutional form of pure red cell aplasia. Anemia is often present at birth and invariably develops within the first 2 years of life. Macrocytosis is common. The marrow shows features common to pure red cell aplasia. Like Fanconi's anemia, fetal hemoglobin and i-antigen expression are increased. Transfusional support is the mainstay of therapy, and a significant number of patients demonstrate improvement in their hematologic status with glucocorticoid therapy. Spontaneous remission has been reported in as many as 20% of cases.

Congenital Dyserythropoietic Anemias

The congenital dyserythropoietic anemias (CDAs) are a group of inherited disorders characterized by ineffective erythropoiesis, morphologic abnormalities involving the erythroid lineage, and mild to moderate anemia. Three main types of CDA have been described, as well as a number of variants. Classification of CDAs is mainly based on the morphologic findings in the bone marrow in conjunction with serologic studies.

Clinical findings include variable anemia, ranging from mild with no symptoms to severe requiring transfusion. Most commonly, however, the anemia is mild to moderate with hemoglobin concentrations between 8 and 11 g/dL. Red blood cells usually are normocytic or macrocytic and may show anisopoikilocytosis and punctate basophilic stippling; the stippling may be pronounced in some cases and serve as the first clue to an underlying CDA. Reticulocytes are usually normal or slightly increased. The red blood cell half-life is decreased to a variable extent in these disorders. The bone marrow shows marked erythroid hyperplasia, inversed myeloid to erythroid ratio, and variable degrees of multinuclearity in the red blood cell precursors. Intramedullary hemolysis occurs due to ineffective erythropoiesis. Myeloid and megakaryocytic lineages are generally unremarkable, although a few cases describing morphologic abnormalities involving these lineages have been reported. Specific chromosomal findings linked to CDAs were recently reported; analysis of erythroid precursors usually shows a hyperdiploid population indicating ineffective cell division that leads to an increased complement of chromosomes.

Hyperabsorption of iron from the gastrointestinal system, as well as frequent transfusion support, frequently leads to secondary hemochromatosis. Transferrin saturation, serum ferritin levels, and stainable iron in the bone marrow are usually increased, but there are no ringed sideroblasts. The intramedullary hemolysis results in elevation of serum LDH, slight hyperbilirubinemia, and suppression of serum haptoglobin, most consistently in CDA II.

In routine clinical practice, the ineffective erythropoiesis is reflected by a suboptimal reticulocyte response for the degree of anemia in the presence of bone marrow erythroid hyperplasia. The diagnosis of CDA is usually made after the exclusion of other causes of congenital and acquired dyserythropoiesis. The prevalence of these disorders may be more common than appreciated because asymptomatic cases with little or no anemia may be missed.

Although ineffective erythropoiesis is generally associated with morphologic evidence of dyserythropoiesis, these findings are not always morphologically apparent. Hence, the diagnosis may rest on clinical suspicion in conjunction with physical findings and additional laboratory studies as discussed below.

Based on morphologic and serologic criteria, CDAs are subgrouped into three types: I, II, and III. CDA I is an autosomal recessive disorder presenting with splenomegaly and mild to moderate macrocytic anemia. The electron microscopic examination of CDA I shows degenerative changes in the nucleus of the erythroid cells; characteristic morphologic features include uneven condensation of chromatin, leading to a spongy nuclear configuration, and interchromatin bridges between nuclei of two separate erythroblasts.

CDA II, the most common of the three CDAs, is inherited as an autosomal recessive disorder. CDA II is distinct from the other types because acidified normal sera (from some, but not all, normal people) lyse CDA II red blood cells. Immunoglobulin (Ig) M antibodies against abnormal cell surface components on CDA II red blood cells mediate this process. Therefore, CDA II is also commonly known as hereditary erythroblastic multinuclearity with a positive acidified serum test. In contrast to patients with paroxysmal nocturnal hemoglobinuria, CDA II red cells are not lysed by the patient's own acidified serum or by isotonic sucrose. An additional feature of CDA II is that these red cells show enhanced agglutination and lysis to anti-i and anti-I.

Splenomegaly and jaundice are usually found in the majority of CDA II patients. Several other diseases have been associated with CDA II, including Sweet's disease, von Willebrand's disease, and Dubin-Johnson syndrome. However, it appears that these are most likely coincidental occurrences rather than true associations. Bone marrow examination shows a variable degree of polychromatophilic and orthochromatic erythroblasts and binuclear or multinucleated erythroid cells. Karyorrhexis is commonly seen.

CDA III is a very rare form of CDA and is also known as familial erythroid multinuclearity and hereditary benign erythroreticulosis. This disorder is inherited as an autosomal dominant trait; although some sporadic cases have been reported, these cases more likely represent spontaneous dominant mutations. Bone marrow examination reveals markedly dysplastic erythroid cells showing giant erythroblasts with up to 12 nuclei, which is the most distinctive finding in CDA III. The most important features differentiating CDA III from CDA I are the presence of splenomegaly and milder erythroid abnormalities in CDA I. In contrast to CDA II, acidified serum lysis test is negative and agglutination and lysis of erythrocytes to anti-i and anti-I are variable.

The defect in CDAs is believed to be intrinsic to all erythroid precursors rather than a subpopulation of cells. Previous studies showed that the abnormality underlying CDA II involves aberration in membrane proteins and lipid composition. Glycoproteins on CDA II erythrocytes show abnormal carbohydrate

structures, leading to the abnormal reactivity with anti-i and anti-I sera. Based on studies of carbohydrate structure, abnormalities in the glycosylation pathway are most likely involved in the etiology of CDA II.

As mentioned above, the diagnosis of CDA requires exclusion of other causes of anemias. Red blood cell folate and serum vitamin B_{12} levels should be used to exclude megaloblastic anemia attributable to these deficiencies. Morphologic abnormalities involving the granulocytic or megakaryocytic lineages, or the presence of ringed sideroblasts, suggest a myelodysplastic syndrome. In such cases, cytogenetic studies may help to establish a precise diagnosis. Hemoglobin electrophoresis may be useful to investigate congenital anemia related to a hemoglobinopathy. Hereditary spherocytosis or other membrane or enzyme disorders may also need to be excluded by specific tests such as osmotic fragility test, membrane protein analysis assays, and enzyme analysis.

Myelophthisic Anemias

Encompassed under the term myelophthisic anemia are the peripheral cytopenias that arise as the result of bone marrow replacement by a variety of disorders. Although most frequently encountered in conjunction with malignant neoplasms metastasizing to or arising within the bone marrow cavity, nonneoplastic disorders may similarly result in marrow replacement. Examples of the former include metastatic carcinoma, leukemia, Hodgkin's and non-Hodgkin's lymphomas, and multiple myeloma. Marrow replacement occurring in the setting of nonneoplastic disease is typically the result of histiocytic proliferation seen either with disseminated granulomatous inflammation such as tuberculosis or in association with a variety of storage disorders.

Marrow infiltration results in varying degrees of anemia, leukopenia, and thrombocytopenia, which may occur singly or in combination. In addition, immature neutrophils and erythroblasts are often present in the circulation, a finding referred to as a leukoerythroblastic reaction. It is easy to conceptualize the reduction in peripheral counts as a crowding-out phenomenon, i.e., normal hematopoietic tissue being replaced by foreign elements. In a similar fashion, immature white and red cells may be squeezed out of the marrow by advancing tumor. It is likely, however, that other mechanisms also come into play, such as an altered and inhospitable marrow microenvironment leading to impaired development, destruction, or premature release of immature hematopoietic elements.

Other changes may also be encountered in the peripheral blood. Reticulocytes are often increased disproportionally to the degree of anemia. This finding may reflect premature release from the marrow. Teardrop-shaped erythrocytes may also be found. Although this shape change has been linked to the aforementioned infiltrative marrow disorders, it is typically a more striking finding in idiopathic myelofibrosis.

Both the diagnosis and etiology of a myelophthisic process rely on bone marrow examination. In many circumstances, the disease process does not diffusely involve the bone marrow. Chances of detection, therefore, are increased by bilateral bone marrow biopsies. Biopsies directed at sites of clinical or radiographically evident disease may also provide an increased yield. Often the aspirate smears may be unrewarding, as many focal lesions are associated with reticulin deposition and are therefore poorly aspirable. Necrosis, when encountered in the aspirate or trephine biopsy, is suggestive of marrow infiltration requiring further investigation.

Accelerated Erythrocyte Destruction

Normal Erythroid Senescence

It is well established that the average natural life span of the human erythrocyte is 120 days. The exact mechanism or terminal events that allow these senescent cells to be recognized and removed from circulation is, however, not well understood. Erythrocytes undergo a number of age-related changes, including diminished intracellular concentrations of glycolytic enzymes, water, and solutes; altered membrane composition; and reduced surface area and volume. These factors undoubtedly have detrimental effects on deformability and functions vital to cellular integrity. However, where and how red blood cells are actually removed from circulation are elusive, as observations to date have elucidated neither the progressive morphologic signs of aging nor the sites where active red cells are lost from circulation each second.

Hemoglobin Catabolism

Senescent red cells liberate hemoglobin, the most of which are catabolized within the reticuloendothelial system of the liver, spleen, and bone marrow. A small amount of hemoglobin, however, is released to circulation where, after dissociation to a-P dimer form, it is bound by haptoglobin, an α_2-glycoprotein synthesized by the liver. The haptoglobin-hemoglobin complex is rapidly cleared by the liver, where catabolism of both hemoglobin and haptoglobin occurs.

Human nature is to conserve, and red cell catabolism is no exception. Within the reticuloendothelial system, hemoglobin is dissociated into its three main building blocks: globin, iron, and protoporphyrin. Globin chains are rapidly degraded with return of their constituent amino acids to the plasma for future protein synthesis, iron is conserved to begin its role again in hemoglobin synthesis, and protoporphyrin is degraded.

The protoporphyrin ring is initially cleaved by heme oxygenase, yielding biliverdin and carbon monoxide. This reaction provides the only source of endogenous carbon monoxide in the body. Biliverdin is reduced to bilirubin and bilirubin is then transported to the liver after solubilization by albumin, where the bilirubin-albumin complex (indirect bilirubin) is rapidly cleared by the liver (Fig. 41.12). Within the hepatocyte, bilirubin is conjugated and excreted in the bile mainly as bilirubin diglucuronide (direct bilirubin). Subsequent hydrolysis by bacterial enzymes present in the ileum and colon frees bilirubin, which is then reduced by anaerobic flora to a family of compounds collectively referred to as urobilinogens. Urobilinogens are largely excreted in the stool. Approximately 10% to 20% of urobilinogen is absorbed from the terminal ileum and redirected to the liver, where it is returned to the bile for excretion. A small amount of urobilinogen escapes this enterohepatic circulation and is excreted by the kidneys.

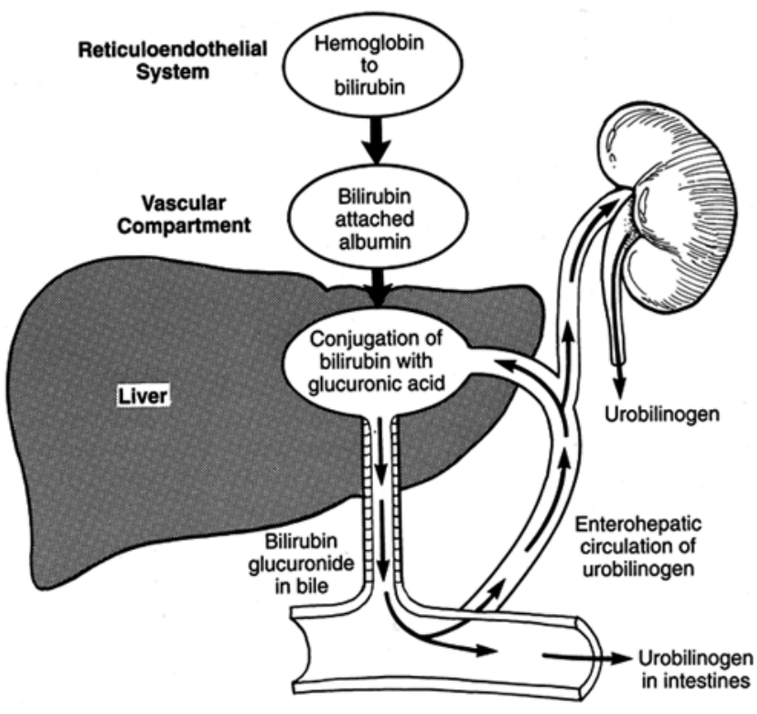


FIGURE 41.12. Hemoglobin liberated from senescent erythrocytes is catabolized in the reticuloendothelial system, and the protoporphyrin ring is converted to bilirubin, which is transported to the liver loosely bound to albumin. In the liver, bilirubin is conjugated and excreted in the bile as bilirubin diglucuronide. Enzymatic hydrolysis occurs in the intestine, with subsequent reduction to a family of compounds, the urobilinogens. Ten percent to 20% of the urobilinogen is reabsorbed and largely returned to the bile; a small amount of urobilinogen escapes the enterohepatic circulation and is excreted into the urine. Elevated levels of unconjugated hemoglobin and urinary urobilinogen are present in the hemolytic anemias. (From Petz LD, Garratty G. The diagnosis of hemolytic anemia. In: Petz LD, Garratty G, eds. *Acquired immune hemolytic anemias*. New York: Churchill-Livingstone, 1980:4, with permission.)

The hemolytic anemias are anemias resulting primarily from increased red cell destruction. By definition, many anemias secondarily associated with a slight reduction in red cell life span are excluded. Disorders associated with defective erythropoiesis leading to marked intramedullary red cell destruction are, by convention, also excluded. The hemolytic anemias are therefore the disorders in which premature extramedullary red cell destruction is the primary cause of the anemia.

Often the terms intravascular and extravascular are used to denote sites of red cell destruction. The term extravascular, however, is somewhat misleading because, in fact, the distinction is between lysis within the systemic circulation and destruction within the sinusoids of the reticuloendothelial system.

Laboratory Findings

Regardless of the mechanism or site of hemolysis, the hemolytic anemias share manifestations of increased red cell catabolism. Although increases in endogenous carbon monoxide production and fecal urobilinogen are measures of hemolysis, the most clinically useful laboratory parameter of increased red cell destruction is serum bilirubin. Elevated bilirubin levels, however, are nonspecific and may be found as a sign of hepatocellular or biliary disease. Elevated levels of indirect bilirubin are viewed as a better measure of hemolysis, provided that disorders of bilirubin conjugation are excluded.

Decreased levels of serum haptoglobin are also found in hemolysis. Although often suggested as an indicator of intravascular red cell destruction, diminished levels of haptoglobin are also found with extravascularly or reticuloendothelial-mediated red cell destruction. It is important to note that haptoglobin is an acute phase reactant; normal values may therefore be encountered in hemolytic disorders complicated by infectious or inflammatory states.

Once haptoglobin is depleted through hepatic degradation of the haptoglobin-hemoglobin complex, free hemoglobin emerges in the circulation (Fig. 41.13). Quantitation of plasma (or serum) free hemoglobin is possible but is fraught with the difficulty of obtaining a blood sample free of the effects of a traumatic venipuncture. In circulation, the free dimeric form of hemoglobin is readily filtered by the glomerulus and absorbed by the proximal convoluted tubules of the kidney. Here hemoglobin may undergo further degradation. Liberated iron is stored in the proximal tubular epithelial cells as ferritin and hemosiderin. When sloughed in the urine, ferritin containing tubular epithelial cells can be detected by using a Prussian blue stain. This measure of increased red cell destruction is most useful in chronic hemolytic states because of the lag phase that occurs between the onset of hemolysis and the detection of hemosiderinuria. The absorptive capacity of the proximal tubular cells for hemoglobin can be overcome, allowing free hemoglobin to pass into the urine. Hemoglobinuria, however, is not detectable until the absorptive capacity of both haptoglobin and the tubular epithelial cells is surpassed.

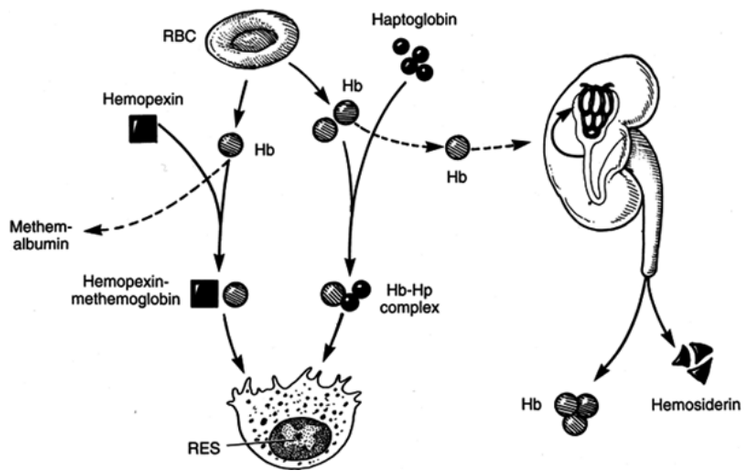


FIGURE 41.13. Intravascular destruction of erythrocytes liberates hemoglobin, which is bound in circulation by haptoglobin. The hemoglobin-haptoglobin complex is cleared by the liver; in the process haptoglobin is consumed. Depletion of haptoglobin results in free dimeric hemoglobin, which is readily filtered and partially excreted by the kidney. Partial reabsorption of hemoglobin occurs in the tubular epithelial cells. In chronic hemolysis, hemosiderin may be detected in the urine with turnover and sloughing of tubular epithelium. Free hemoglobin in the circulation is also oxidized to methemoglobin, which is bound by hemopexin and cleared by the liver analogous to the hemoglobin-haptoglobin complex. With depletion of hemopexin, methemoglobin is bound by albumin to form a methemalbumin. Intravascular hemolysis classically results in decreased haptoglobin, hemoglobinuria, decreased hemopexin, methemalbuminemia, and hemosiderinuria. (From Petz LD, Garratty G. The diagnosis of hemolytic anemia. In: *Acquired immune hemolytic anemias*. New York: Churchill-Livingstone, 1980:3, with permission.)

Plasma hemoglobin not bound to haptoglobin nor removed by the kidney is oxidized to methemoglobin. Subsequently, the oxidized heme moiety undergoes rapid dissociation and is then bound by hemopexin. This complex is cleared by the liver in a manner similar to the haptoglobin-mediated clearance of hemoglobin. After depletion of hemopexin, oxidized heme binds to albumin, forming methemalbumin. Both the depletion of hemopexin and the presence of methemalbumin are two additional

measures of hemolysis, although they offer little practical clinical utility.

Another intracellular constituent of erythrocytes that is readily measurable in the laboratory is LDH. Because LDH is not restricted to erythrocytes, serum elevations may be encountered in a variety of disorders. Isoenzyme fractionation is somewhat more specific, as fractions LDH_1 and LDH_2 predominate in erythrocytes.

In the vast majority of uncomplicated hemolytic anemias, increased red cell destruction provides a stimulus for erythropoiesis. Invariably then, evaluation of suspected hemolytic states should include not only measures of increased red cell destruction but also parameters of accelerated erythropoiesis. These changes are readily apparent in a peripheral blood smear as an increased number of polychromatophilic erythrocytes. A more objective measure is obtained with a reticulocyte count. Brisk reticulocytosis may elevate the mean corpuscular volume. Circulating erythroblasts may also be seen. Bone marrow examination typically reveals an erythroid hyperplasia, which in some cases may be quite striking, as the bone marrow is capable of increasing erythropoiesis six- to eightfold.

Pathophysiologic Classification

A pathophysiologic classification of disorders resulting in shortened red cell survival is presented in Table 41.8. Broad separation is based on the nature of the abnormality: is the defect intrinsic to the red cell, or do extrinsic forces result in hemolysis? With one exception, intrinsic disorders are inherited defects, whereas extrinsic disorders are acquired. Paroxysmal nocturnal hemoglobinuria, the exception, is an acquired intrinsic abnormality.

TABLE 41.8. CLASSIFICATION OF HEMOLYTIC DISORDERS

Inherited hemolytic disorders
Defects in the erythrocyte membrane
Hereditary elliptocytosis
Hereditary pyropoikilocytosis
Stomatocytosis
Abetalipoproteinemia (acanthocytosis)
Enzyme deficiencies of the pentose phosphate pathway
Glucose-6-phosphate dehydrogenase
Defects in globin structure and synthesis
Hemoglobinopathies
Thalassemias
Unstable hemoglobin disease
Enzyme deficiencies of the glycolytic pathway
Pyruvate kinase
Hexokinase
Glucose-phosphate isomerase
Phosphofructokinase
Aldolase
Other
Defects in nucleotide metabolism
Pyrimidine 5′nucleotidase
Acquired hemolytic disorders
Immune-mediated erythrocyte destruction
Transfusion of incompatible blood
Hemolytic disease of the newborn
Autoimmune
Infectious agents
Protozoans
Bacteria
Paroxysmal nocturnal hemoglobinuria
Traumatic erythrocyte destruction
Macrovascular
Microvascular
Chemicals/drugs/venoms
Physical agents

Intrinsic and Inherited Red Cell Abnormalities

Intrinsic erythrocyte abnormalities are commonly grouped as membrane, metabolic, or hemoglobin defects. In keeping with convention, therefore, a discussion of the hemoglobinopathies is not undertaken here, but may be found in Chapter 42.

As inherited defects, these disorders may present a number of clinical features not typically seen with acquired red cell abnormalities. In large part, this reflects the chronic nature of these disorders as well as their capacity to afflict newborns in severe cases. Persistent hyperbilirubinemia often leads to cholelithiasis, in turn complicated by cholecystitis. Hyperbilirubinemia in neonates may lead to the debilitating neurologic disorder kernicterus.

Chronic hemolytic disorders may be complicated by periods or crises of profound anemia. As previously described, individuals may experience an aplastic crisis. Somewhat of a misnomer, the term aplastic crisis refers to a transient, self-limited exacerbation of the underlying anemia resulting from a cessation of erythropoiesis

often seen in association with a recent viral illness. A hemolytic crisis may arise as a complication of a disorder causing splenic enlargement. A sudden increase in spleen size will exacerbate the underlying anemia because the spleen is often the site of red cell sequestration and destruction. A megaloblastic crisis may complicate chronic hemolysis as folate stores are depleted through the increased requirements of compensatory erythroid hyperplasia. Both aplastic and megaloblastic crises are characterized by reticulocytopenia. The abrupt onset and marked erythroid hypoplasia of an aplastic crisis, however, contrast with the slow development and megaloblastic appearance of folate deficiency.

Structural Membrane Defects

The shape and reversible deformability of the normal erythrocyte are properties determined by the membrane skeleton. Lying directly beneath the external lipid bilayer membrane is a complex network of proteins that provide structural support (Fig. 41.14). Spectrin is the major component of the red cell cytoskeleton and consists of two intercoiled, nonidentical filamentous subunits that form heterodimers. The chain heads of each dimer pair bind with opposite subunit heads of another dimer pair to form tetramers. The tails of spectrin tetramers bind with a protein cluster of short actin protofilaments. This interaction is markedly enhanced by protein 4.1. The spectrin-actin-protein 4.1 assembly forms a two-dimensional web that is secured to the overlying lipid bilayer through ankyrin, which anchors spectrin to the cytoplasmic domain of the anion transporter. Additional linkage occurs through the binding of protein 4.1 to glycophorin.

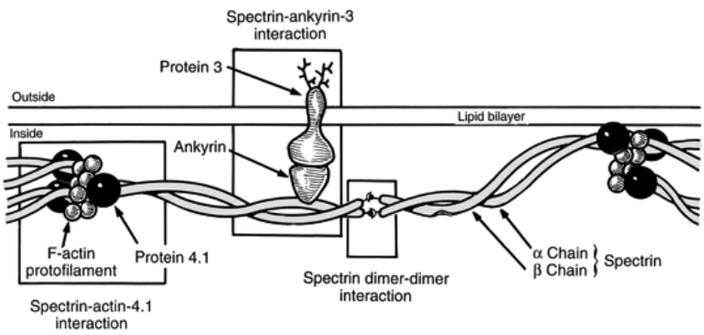


FIGURE 41.14. Schematic illustration of the erythrocyte cytoskeleton. (From Palek J, Lux SE. The red cell membrane skeletal defects in hereditary and acquired hematolytic anemias. *Semin Hematol* 1983;20:189, with permission.)

Defects in the red cell cytoskeleton, arising either from a deficiency of constituent proteins or through defective protein interaction, are a major cause of hereditary hemolytic anemias. Regardless of the mechanism or associated morphologic changes, red cells in this group of disorders exhibit decreased deformability. Consequently, they sluggishly traverse the splenic cords, where they are subjected to an environment hostile to metabolic activity. The metabolic stress further weakens an inherently defective cell, and after a variable number of passages, the cells succumb to reticuloendothelial destruction.

Hereditary Spherocytosis

Hereditary spherocytosis is the descriptive name given to a group of inherited disorders with characteristic spheroidal red cell morphology. It is the most common hereditary hemolytic anemia among people of Northern European ancestry. Its frequency led to the practice of separating anemias into spherocytic and nonspherocytic varieties, nomenclature that, in part, is still utilized today.

At least two distinct forms of spherocytosis are recognized. Approximately 75% of individuals exhibit an autosomal dominant mode of inheritance and cytoskeletal defects, which have been linked either to defective interaction of spectrin with protein 4.1 or to defects in ankyrin. In 25% of cases, an affected parent cannot be identified, negating an autosomal dominant inheritance pattern. This subset of individuals may be heterogeneous, including cases of spontaneous mutation as well as cases of autosomal recessive inheritance. A quantitative deficiency of spectrin has been identified in this group of patients, the magnitude of which directly correlates with clinical severity.

Clinical Features. The severity of the anemia varies, not only between patients, but also within a given patient over the course of time. Approximately 20% of patients have a mild form of the disease. In these individuals, hemolysis is compensated for by accelerated erythropoiesis, resulting in subclinical disease. The majority of patients with hereditary spherocytosis have a moderate lifelong anemia associated with cholelithiasis, intermittent jaundice, and splenomegaly. The hemoglobin level fluctuates, reflecting the precarious balance between the hyperplastic erythron and splenic activity. In approximately 10% of cases, anemia is severe. Transfusional support is often required in infancy and childhood until splenectomy is logistically feasible.

Hematologic Findings. The morphologic hallmark of hereditary spherocytosis is the spherocyte, a generally smaller, spheroidal red cell lacking central pallor and consequently appearing densely hemoglobinized (Fig. 41.15). The spherocytes contrast sharply with the increased number of large polychromatophilic reticulocytes indicative of compensatory erythroid hyperplasia. The interplay between spherocyte and reticulocyte numbers results in a widely variable MCV. Whereas the MCH is normal, the MCHC is increased, a combination of findings generally

unique to hereditary spherocytosis. The number of spherocytes present in the peripheral blood varies but in most cases is quite significant. In the absence of other significant poikilocytosis, the presence of spherocytes is strongly suggestive of either hereditary spherocytosis or a warm autoimmune hemolytic anemia. These two conditions may be distinguished by a direct antiglobulin test.

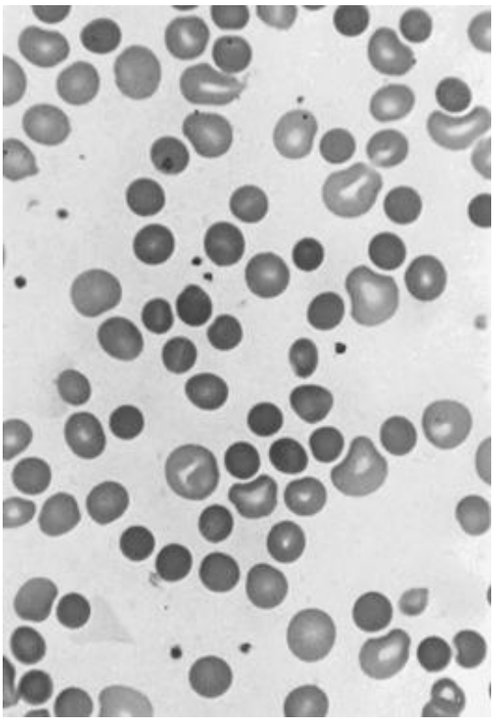


FIGURE 41.15. The peripheral blood smear in hereditary spherocytosis demonstrates a population of small, densely hemoglobinized, spheroidal erythrocytes that sharply contrast with the large polychromatophilic reticulocytes indicative of increased erythropoiesis.

In cases in which the diagnosis is suspected but the peripheral blood findings are inconclusive, an osmotic fragility test may be helpful. This procedure measures the red cell's capacity to withstand hypotonic stress. Because spherocytes exhibit a decreased surface area to volume ratio, their capacity to absorb water in hypotonic salt solutions is impaired. Increased hemolysis usually begins at sodium concentrations less than 0.5%. Exacerbation of this phenomenon may be induced by incubation for 24 hours, during which time the red cells additionally become metabolically stressed.

Treatment. In symptomatic patients, splenectomy is beneficial in removing the site of red cell destruction. After splenectomy, patients still exhibit spherocytes as well as the morphologic changes of a postsplenectomy state. Failure to improve after splenectomy suggests an alternative diagnosis; later relapses suggest hyperplasia of accessory splenic tissue.

Hereditary Elliptocytosis

Hereditary elliptocytosis (HE) is a clinically, genetically, and morphologically heterogeneous group of disorders that are linked through the morphologic finding of more than 20% elliptocytes in the peripheral blood (Fig. 41.16). Based on clinical findings and red cell morphology, three major categories of HE have emerged: the common form, the spherocytic form, and the stomatocytic form.

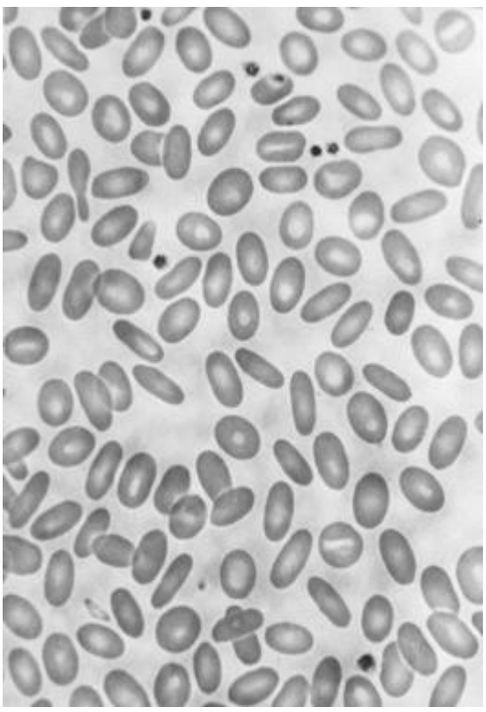


FIGURE 41.16. Numerous elliptocytes characterize the common form of hereditary elliptocytosis.

Clinical Findings. In the common form of HE, the majority of individuals exhibit numerous elliptical red cells with little or no evidence of hemolysis. Several exceptions may occur, leading to clinically symptomatic hemolysis. Moderate to severe hemolysis may occur during the first year of life in some black children, coinciding with morphologic findings in the peripheral blood similar to hereditary pyropoikilocytosis (HPP). The hematologic and clinical picture changes to one typical of common HE over the first 12 months. Unexplainably, a small subset of individuals with common HE will exhibit significant and chronic hemolytic anemia with clinical manifestations similar to hereditary spherocytosis. Additionally, the rare homozygous form of common HE results in severe transfusion-dependent hemolytic anemia. Red cell morphology in this unusual condition found in the offspring

of related HE parents resembles that of HPP. Spherocytic HE represents the dual inheritance of two nonallelic genes: one for mild HE and one for mild hereditary spherocytosis. As might be expected, the peripheral blood shows varying numbers of spherocytes and elliptocytes, either of which may predominate. Clinically mild to moderate hemolytic anemia results.

Stomatocytic HE is an unusual variant limited to Melanesia and surrounding islands. It is associated with changes in red cell antigen expression that appear to serve a protective function against malarial infection.

Several defects in the red cell cytoskeleton have been discovered in HE. The most frequent variant, occurring in common HE, is defective spectrin dimer–dimer interaction owing to an abnormal a chain of spectrin. Quantitative deficiencies of protein 4.1 are associated with the spherocytic form of HE. Heterozygous and homozygous patterns of inheritance are directly related to the magnitude of protein deficiency and clinical severity. Defective binding of ankyrin to the anion transporter has also been described in HE.

Differential Diagnosis. In the typical form of common HE, the diagnosis is easily established by the finding of numerous, fairly uniform elliptocytes in the peripheral blood. Other poikilocytes are few. The reticulocyte count is usually normal or mildly elevated, reflecting the asymptomatic or well-compensated nature of this disorder. The distinction between this form of HE and other disorders in which elliptocytes may be found is easily made on the basis of two observations: the frequency of elliptocytes and the presence of other size and shape changes that may give a clue to an unrelated hematologic condition. The peripheral smears of other forms of HE are not pathognomonic and may show features suggestive of hereditary spherocytosis or HPP. The diagnosis in these cases requires study of family members and may necessitate analysis of cytoskeletal protein composition.

Therapy. In the common form of HE, patients are asymptomatic or very well compensated and require no specific therapy. Splenectomy is the treatment of choice in more symptomatic HE variants.

Hereditary Pyropoikilocytosis

HPP is a rare congenital hemolytic anemia occurring most commonly in blacks. Red cells in this disorder demonstrate increased susceptibility to thermal injury. In contrast to normal erythrocytes, which can withstand temperature to 49°C, red cells in HPP begin to fragment and disintegrate at 45 to 46°C.

The mode of inheritance and the exact cytoskeletal defect(s) are not completely understood. Similar to HE, a defect in spectrin dimer–dimer association has been described that is partly attributable to an abnormality of the α chain of spectrin. The similarity of this structural defect to that seen in HE may explain the partial overlapping in morphologic findings seen early in the course of some cases of common HE.

The peripheral blood findings in HPP are striking and nearly pathognomonic (Fig. 41.17). Poikilocytosis is unmatched by any other disorder. Erythrocytes appear as if they are disintegrating, with fragments, numerous microspherocytes, and wisps of red cell membrane. These changes all lead to a markedly reduced MCV. A similar peripheral blood picture may be found in cases of third-degree burns owing to thermal injury of normal red cells and in clostridial sepsis secondary to the action of a lecithinase

on red cell membranes. As previously noted, infants with common HE may also exhibit similar changes in the peripheral blood, which gradually revert to a picture typical of common HE by 1 year of age.

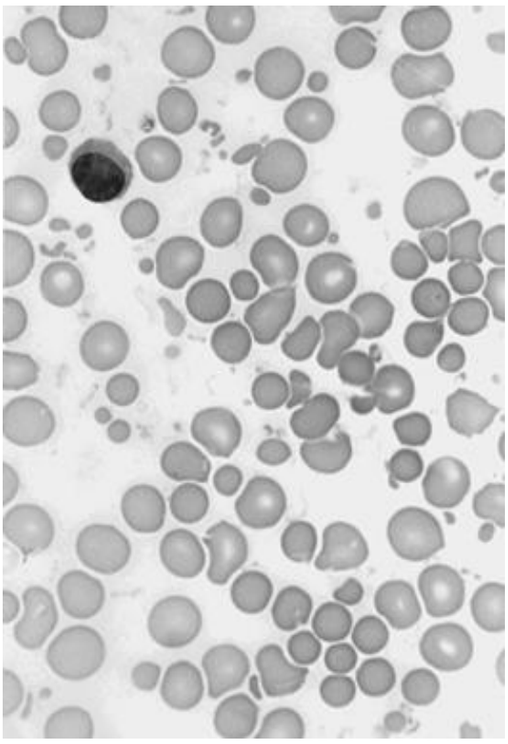


FIGURE 41.17. Striking poikilocytosis characterizes hereditary pyropoikilocytosis. Minute spheroidal erythrocytes result in a marked reduction in mean corpuscular volume. Note the irregular fragments of red cell membrane dusting the background of the peripheral smear.

HPP results in moderate to severe hemolytic anemia, necessitating splenectomy, which serves to diminish or abate hemolysis.

Hereditary Stomatocytosis

Although the precise structural defect is unknown, increased permeability to both sodium and potassium ions is a consistent feature of this rare disorder. Active Na^+-K^+ ATPase pumps are incapable of counterbalancing the flow of cations across the cell membrane. The net influx of sodium brings with it water, resulting in a swollen cell that is susceptible to osmotic and mechanical lysis.

Clinically, hereditary stomatocytosis varies in its expression. Most individuals are asymptomatic or exhibit only mild anemia because of brisk compensatory erythroid hyperplasia.

The morphologic hallmark of hereditary stomatocytosis is the stomatocyte, a cell with a slitlike or “fish-mouth” area of central pallor (Fig. 41.18). In suspension, these cells appear uniconcave or bowl shaped. A small number of similar-appearing cells may appear normally in air-dried smears or may be found in slightly increased numbers in association with a variety of disorders, making it difficult to establish conclusively the presence of hereditary stomatocytosis in its mild form. Evidence of hemolysis (more than 35% stomatocytes) and the absence of other poikilocytes, however, provide convincing evidence for this rare autosomal disorder. The MCV in these patients may be strikingly elevated, reflecting the reticulocytosis and the swollen nature of the stomatocytes.

Symptomatic patients may benefit from splenectomy, provided that splenic sequestration is documented.

Rh_{null} Disease

Rh_{null} disease is a rare hereditary disorder characterized by deletion of all Rh determinants, including the Landsteiner-Wiener antigen. The Rh locus, unlike many other blood group antigens, is restricted to red cells. It appears to be located in part within the red cell membrane; therefore, it is not surprising the deletion results in membrane malfunction.

Morphologically, this disorder is associated with stomatocytes and spherocytes. Clinically, a mild compensated hemolytic anemia is found, which requires no therapy.

Abetalipoproteinemia (Acanthocytosis)

Abetalipoproteinemia is a rare disorder of lipid metabolism. The absence of apolipoprotein B results in the inability to transport triglycerides in the blood. Clinically, the syndrome presents in infancy with steatorrhea; progressive development of acanthocytosis, ataxic neuropathy, and an atypical form of retinitis pigmentosa ensues.

The acanthocytes found in the peripheral blood are dense, contracted erythrocytes with multiple irregularly spaced, broad, spiny projections (Fig. 41.19). This is presumably related to an increase in the relative sphingomyelin content of the red cell membrane, resulting in a rigid cell. The mild shortening of red cell survival is clearly overshadowed by the nonhematologic complications of this disorder.

Acanthocytes may also be found as a complication of severe hepatocellular disease. Unlike abetalipoproteinemia, acanthocytosis in this setting is associated with a severe hemolytic anemia and heralds a poor and often fatal prognosis. The defect is again related to lipid alterations of the red cell membrane, in this case

excess cholesterol resulting from increased high-density lipoproteins in the peripheral blood. Similar acanthocytosis may be seen associated with the hepatocellular dysfunction accompanying neonatal hepatitis.

For unknown reasons, acanthocytes are also found in association with the McLeod blood group, an X-linked disorder in which red cell Kx antigen, the precursor substance for the Kell blood group system, is absent. The McLeod phenotype may be associated with chronic granulomatous disease because of the proximity of the genetic loci for these two disorders.

Congenital Nonspherocytic Anemias: Enzyme Deficiencies Resulting in Shortened Red Cell Survival

Lacking mitochondria, the mature erythrocyte relies on glycolysis for energy production. Approximately 90% of glycolysis occurs through the anaerobic Embden-Meyerhof (EM) pathway as glucose is converted to lactic acid (Fig. 41.20). The metabolically crucial byproducts of this multistep pathway are ATP and nicotinamide adenine dinucleotide (reduced form) (NADH). ATP, the main energy compound of the erythrocyte, is required for active

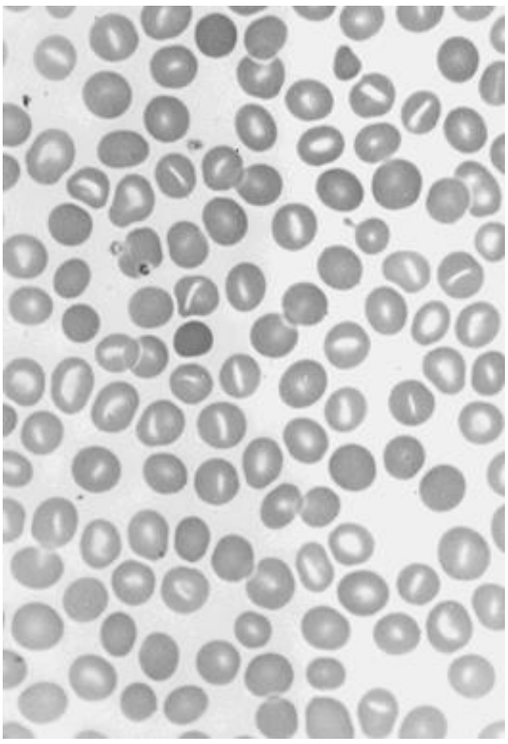


FIGURE 41.18. Hereditary stomatocytosis is characterized by numerous erythrocytes having slitlike central pallor.

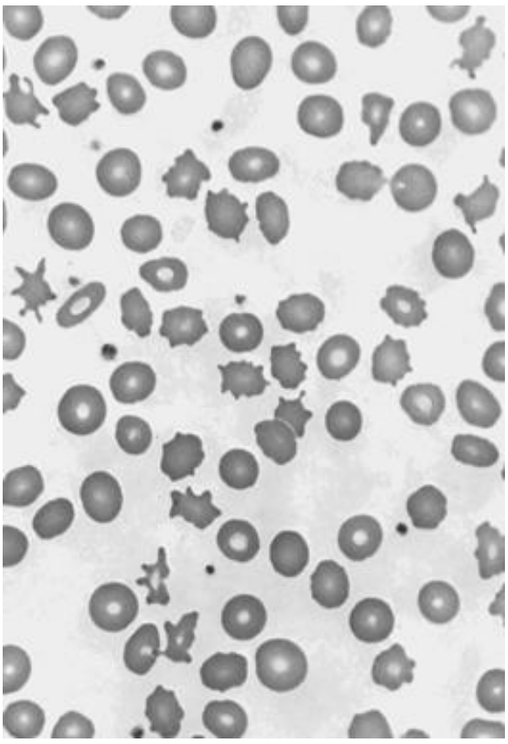


FIGURE 41.19. Acanthocytes are erythrocytes with irregularly placed, broad, blunted projections. They are typically found in significant numbers in abetalipoproteinemia and in end-stage liver disease. In both conditions, plasma lipid abnormalities result in alterations in the composition of the red cell membrane.

cation transport across the red cell membrane. NADH is an essential cofactor for reduction of the small amount of methemoglobin normally produced each day in the red cell.

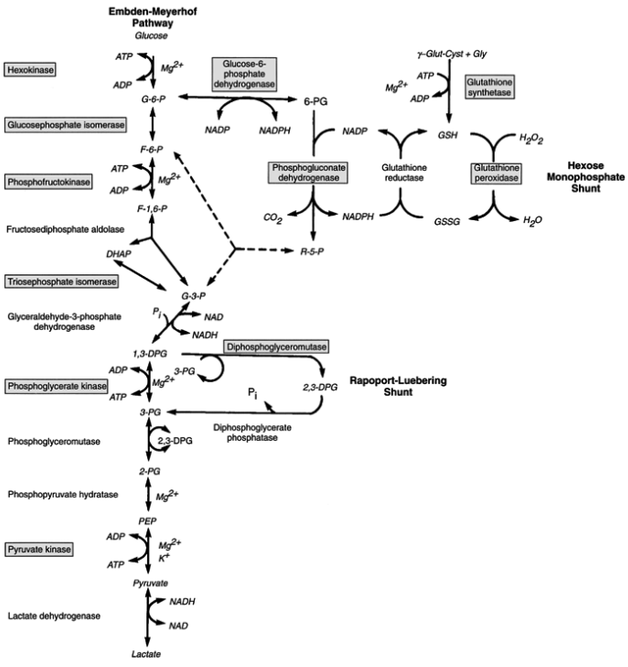


FIGURE 41.20. Glycolytic pathway with related hexose monophosphate shunt (glutathione metabolism) and Rapoport-Luebering shunt [2,3-diphosphoglycerate (2,3-DPG), metabolism].

Integrally related to the EM pathway is the Rapoport-Luebering shunt, which results in the production of 2,3-diphosphoglycerate (2,3-DPG). Activity of this shunt is related to the erythrocyte's ATP and 2,3-DPG requirements. 2,3-DPG plays an essential role as a regulator of oxygen delivery to the tissues by altering the affinity of hemoglobin for oxygen.

Approximately 10% of glycolysis occurs aerobically through the hexose monophosphate (HMP) shunt, a pathway that generates the nicotinamide adenine dinucleotide phosphate (reduced form) (NADPH) required for reduction of glutathione. Reduced glutathione is critical for the protection of hemoglobin from oxidant damage. Unlike the EM pathway, the HMP shunt can increase activity approximately 30-fold in times of oxidative stress.

Given the complexity and vital role of these pathways, it is not surprising that defects in any number of the intermediary enzymes can serve to uncouple the reaction, resulting in red cell injury. For the most part, red cell injury resulting from deficiencies of the HMP or EM pathways are not associated with significant morphologic changes and are broadly categorized as congenital or hereditary nonspherocytic hemolytic anemias.

Enzyme Deficiencies of the Hexose Monophosphate Pathway

NADPH generated by the HMP shunt is a necessary cofactor for maintaining glutathione in a reduced state. Reduced glutathione protects the red cell from naturally occurring and drug-induced oxidants that may potentially damage hemoglobin, membrane proteins, or other intracellular enzymes.

Oxidative damage to red cells is frequently manifested as Heinz bodies, small spheroidal inclusions of denatured hemoglobin visible by supravital staining. Heinz bodies exert their deleterious effects by attaching to the inner membrane of the red cell and causing deformation, alteration of cell membrane permeability, and reduction in red cell deformability. Attempted removal of these harmful inclusions occurs both through exocytosis and through active excision by the reticuloendothelial system of the spleen (Fig. 41.21). Within the spleen, many Heinz body–laden erythrocytes are trapped and destroyed as they lose the pliability needed to traverse the splenic red pulp. Others escape and return to the circulation, appearing morphologically deformed as bite cells or blister cells.

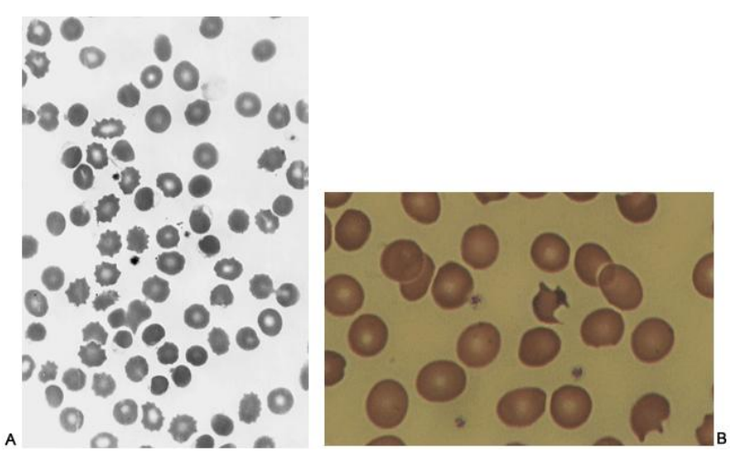


FIGURE 41.21. Blister cells **(a)** and bite cells **(B)** are characteristic of Heinz body–mediated hemolysis. Blister cells appear to have a partially raised or blistered membrane; bite cells appear to have one or more “bites” removed from the cell. Similar bite cells can be found in microangiopathic hemolytic anemia; however, they are then associated with irregular red cell fragments of various sizes and shapes, which are not seen in association with Heinz body–mediated hemolysis.

Glucose-6-Phosphate Dehydrogenase Deficiency

Glucose-6-phosphate dehydrogenase (G6PD), the initial and rate-limiting enzyme of the HMP pathway, accounts for more than 99% of all cases of hemolytic anemia attributable to enzyme deficiencies of this pathway. It is the most prevalent inborn metabolic disorder of red cells. The structure and synthesis of G6PD are controlled by genes on the X chromosome. Numerous variants of the enzyme are found, based on a variety of physiochemical properties. Normal G6PD, or G6PD B, is the most common form of the enzyme found in all population groups. G6PD A, separated based on electrophoretic mobility, exhibits normal enzyme activity and is found in approximately 20% of American black males. An unstable variant of G6PD A, designated G6PD A-, results in enzyme deficiency in aging red cells. G6PD A- is found in approximately 11% of American black males. Another variant, G6PD Mediterranean, found frequently in Sicilians, Greeks, Sephardic Jews, and Arabs, exhibits even greater instability, with a half-life measured in terms of hours compared with 60 days for the normal G6PD B. Together, G6PD A- and G6PD Mediterranean account for the most

common variants associated with clinically significant hemolysis.

With this sex-linked pattern of inheritance, the clinical expression of G6PD deficiency is largely limited to hemizygous males and homozygous females. On occasion, heterozygous females may exhibit hemolysis owing to the variable nature of X chromosome inactivation. Three distinct clinical syndromes are recognized: acute hemolytic anemia induced by oxidant stress, chromic hereditary nonspherocytic anemia, and favism. Although favism is an idiosyncratic reaction occurring on exposure to the fava bean in a subset of G6PD-deficient individuals, the severity of the enzyme defect is the essential element in defining the two remaining clinical syndromes.

Oxidant-Induced Acute Hemolytic Anemia. An acute hemolytic episode associated with oxidant stress is the most frequent expression of G6PD deficiency. It is seen with both G6PD A- and G6PD Mediterranean variants. Administration of a variety of drugs has been associated with acute hemolytic episodes in these individuals (Table 41.9). Additionally, infection is a common precipitating factor. Salmonellae, coliforms, ß-hemolytic streptococci, rickettsiae, influenza, and viral hepatitis have each been associated with hemolysis in G6PD A- individuals through poorly elucidated mechanisms.

TABLE 41.9. DRUGS OR CHEMICALS PRODUCING HEMOLYTIC ANEMIA IN INDIVIDUALS DEFICIENT IN ERYTHROCYTE GLUCOSE-6-PHOSPHATE DEHYDROGENASE

Antimalarials
Chloroquine
Primaquine
Quinacrine
Quinine
Antipyretics and analgesics
Acetaminophen
Acetanilide
Acetylsalicylic acid, etc.
Aminopyrine
Antipyrine
Phenacetin acetophenetidin
Sulfonamides
Azulfidine
Sulfacetamide
Sulfadiazine
Sulfamethoxazole
Sulfamethoxypyridazine
Sulfanilamide
Sulfapyridine
Sulfisoxazole
Nitrofurans
Furazolidone
Nitrofurantoin
Nitrofurazone
Sulfones
Dapsone
Sulfoxone
Thiazolesulfone
Other drugs
Acetylphenylhydrazine
Ascorbic acid
p-Aminosalicylic acid
Chloramphenicol
L-Dopa
Methylene blue
Nalidixic acid
Naphthalene
Niridazole
Phenylhydrazine
Probenecid
Quinidine
Toluidine blue
Trinitrotoluene
Vitamin K (water-soluble analogs)

Laboratory and clinical evidence of acute hemolysis is usually apparent within days of onset of infection or drug administration. Signs and symptoms are often more severe with G6PD Mediterranean, as the enzyme is much less stable. The finding of bite cells or the demonstration of Heinz bodies in this clinical setting are strong evidence for G6PD deficiency. If exposure to the inciting agent is brief, the episode of hemolysis is generally self-limited. Symptoms may also abate as compensatory release of reticulocytes into the circulation occurs. These young red cells demonstrate normal G6PD activity.

Confirmation of the diagnosis requires quantitative G6PD assays. Caution must be exercised in interpreting G6PD levels during periods of brisk reticulocytosis, as enzyme levels during these periods may not reflect steady-state conditions.

Hereditary Nonspherocytic Hemolytic Anemia. G6PD variants that lead to severe enzyme deficiency result in chronic hemolysis. On occasion, this subset of individuals may present in the neonatal period with hemolytic disease of the newborn. Moderate, partially compensated hemolytic anemia is the rule. Splenectomy is generally unsuccessful in diminishing the rate of hemolysis in these individuals.

Favism. Acute, potentially fatal intravascular hemolysis may occur as an idiosyncratic reaction on exposure to fava beans in a subset of G6PD-deficient individuals. G6PD Mediterranean is the most commonly afflicted subtype. Signs and symptoms of brisk intravascular hemolysis occur within hours of ingestion. Maintenance of intravascular volume and renal function together with prompt transfusional support are required.

Miscellaneous Enzyme Deficiencies of the HMP Shunt

Deficiencies of the following HMP pathway enzymes have also been described: γ-glutamyl cysteine synthetase, glutathione synthetase, glutathione reductase, and glutathione peroxidase. Clinical manifestations of these enzymes are widely variable, spanning the recognized clinical syndromes of G6PD deficiency.

Enzyme Deficiencies of the EM Pathway

A syndrome of hemolytic anemia has been described for deficiencies of nearly all enzymes in the EM pathway. Hemolysis occurs in these disorders as the late-developing erythroblasts lose their mitochondrial machinery and begin to rely on glycolysis for energy production.

Pyruvate Kinase Deficiency

Pyruvate kinase (PK) deficiency is the most common enzyme deficiency of the EM pathway. Numerous dysfunctional variants of the PK enzyme are known to exist, affecting, for example, enzyme activity, stability, or kinetic rate. A single copy of a defective gene usually causes no significant alteration of enzyme activity. Hemolytic anemia typically occurs in the double heterozygote, an individual inheriting two mutant genes. The clinical variability of PK deficiency likely reflects the wide range of combinational possibilities.

Hemolytic anemia presenting in infancy or childhood is usually severe, requiring maintenance transfusions and splenectomy, whereas disorders presenting in adulthood are largely compensated and result in mild to moderately severe anemia. Paradoxically, adaptation to anemic hypoxia is facilitated by the PK deficiency itself, which results in a buildup of pathway intermediates proximal to the site of PK activity, including 2,3-DPG. Elevated levels of 2,3-DPG also lead to a peculiar laboratory finding in PK-deficient red cells. By inhibiting enzymes of the HMP pathway, red cells exhibit increased Heinz body formation on exposure to oxidant stress.

Red cell morphology is often unremarkable. In more severe cases, echinocytes may be present, which become even more plentiful after splenectomy. Diagnosis of PK deficiency is based on enzymatic assays demonstrating quantitative or qualitative abnormalities. Hemolysates should be free of contaminating white blood cells, which normally contain high levels of PK.

An acquired form of PK deficiency may occur in conjunction with the myelodysplastic syndromes, myeloproliferative disorders, or AML.

Miscellaneous Enzyme Deficiencies of the EM Pathway

Deficiencies of hexokinase, glucosephosphate isomerase, phosphofructokinase, aldolase, triosephosphate isomerase, and phosphoglycerate kinase have all been reported to result in congenital nonspherocytic hemolytic anemia. With the exception of hexokinase, these enzyme defects have also resulted in impairment of other organ systems, most commonly involving the neuromuscular system.

Disorders of Nucleotide Metabolism

Three uncommon disorders of purine and pyrimidine metabolism in red cells are responsible for hemolytic anemia, the most frequent of which is pyrimidine-5′-nucleotidase deficiency. This enzyme is responsible for degradation of ribosomal RNA during reticulocyte maturation. A pyrimidine-5′-nucleotidase deficiency results in aggregation of residual ribosomes, which are manifested in peripheral blood smears by coarse basophilic stippling. It is an acquired pyrimidine-5′-nucleotidase deficiency seen in lead poisoning that leads to this heralded morphologic clue to the diagnosis. Congenital enzyme deficiencies usually result in moderate anemia. Diagnosis is established through an assay of nucleotidase activity.

A deficiency of adenylate kinase has also been described as a rare cause of congenital nonspherocytic hemolytic anemia, as has hyperactivity of the enzyme adenosine deaminase.

Acquired and Extrinsic Red Cell Abnormalities

Immune-Mediated Red Cell Destruction

The basic antigen-antibody interactions that serve to protect through red cell recognition of nonself can mediate red cell destruction. Such immunohemolytic anemias may be caused by antibodies of the IgG or IgM class, the intrinsic properties of which lead to different modes of red cell injury.

IgM antibodies may cause red cell destruction through two mechanisms: physical agglutination and complement activation, each facilitated by the large pentameric structure of the IgM molecule, which allows it to easily bridge multiple antigenic sites. The capacity of an IgM antibody to mediate red cell destruction is directly linked to its thermal amplitude, the temperature range over which it most avidly binds to an antigen. IgM antibodies are frequently referred to as cold-reacting antibodies because they exhibit increasing binding avidity as temperatures approach 4°C. In contrast, complement activity occurs only near body temperature. Consequently, only IgM antibodies, which exhibit avid binding properties near body temperature where complement is active, cause hemolysis.

When complement is activated, immediate red cell lysis may proceed through completion of the complement cascade. Complement activation may, however, be stalled through a series of protective checks and balances after the liberation of C3b from C3. Red cells coated with C3b may be cleared by phagocytic cells bearing C3b receptors; particularly rich in C3b receptors are the hepatic reticuloendothelial cells (Kupffer cells). Alternatively, C3b may be inactivated to C3d by a plasma inactivator. Importantly, C3d-coated red cells serve as a marker of IgM-mediated hemolysis, a property that is clinically utilized in the Coombs' or direct antiglobulin test (DAT).

In contrast to IgM, antibodies of the IgG class exhibit their greatest binding activity at body temperature. The much smaller size of the IgG molecule makes it less capable of bridging the gap between two or more red cell antigens, a prerequisite to physical agglutination or complement activation. Consequently, IgG-induced hemolysis is not mediated through complement-dependent lysis but rather through reticuloendothelial clearance.

When bound to the red cell surface, the Fc portion of the IgG molecule is exposed, becoming a target for the mononuclear cells of the reticuloendothelial system that bear Fc receptors. Particularly rich in Fc receptors are the splenic macrophages, making the spleen a major site of destruction of IgG-coated red cells. Unlike the rapid demise of IgM-coated red cells, red cells coated with IgG are readily found in circulation. The circulating IgG-coated erythrocytes may be detected by the DAT. The longevity of IgG-coated red cells is determined by the number of antigenic sites per red cell, the serum concentration of the IgG antibody, and the functional capacity of the reticuloendothelial system. When bound through Fc receptor attachment, the erythrocytes undergo a shape change, becoming spherical owing to a loss of red cell membrane. The diminished surface area-to-volume ratio makes these spheroidal red cells increasingly susceptible to hypotonic (or osmotic) stress. These morphologic and physical properties are indistinguishable from the findings of hereditary spherocytosis. Destruction occurs through internalization or

piecemeal destruction by macrophages, as they surround portions of the red cell membrane with fingerlike extensions.

Isoantibodies

Immunohemolytic anemias may be caused by isoantibodies (alloantibodies), i.e., antibodies to blood group antigens. For purposes of this discussion, isoantibodies are antibodies to the major and minor blood group antigens. Isoantibodies may be naturally occurring or may arise on exposure to foreign blood group antigens through blood transfusion or during pregnancy as small amounts of fetal blood leak across the placental barrier to gain access to the maternal circulation.

Transfusion of Incompatible Blood

Antibodies to the ABO blood group antigens are of the IgM class and exhibit a broad range of thermoreactivity. Transfusion of ABO-incompatible blood results in immediate and severe symptoms related to brisk complement-mediated hemolysis. Flushing, hyperventilation, tachycardia, urticaria, and shock occur secondary to the release of vasoactive substances encountered with complement activation; severe pain results from vasoocclusion secondary to red cell agglutination. Disseminated intravascular coagulation and renal failure may also ensue.

Non–ABO-related hemolytic transfusion reactions are IgG mediated. The majority are secondary to antibodies within the Rh system, notably anti-D. Antibodies to c, E, and to the Kell, Kidd, and Duffy systems round out the most frequent causes of non–ABO-related transfusion reactions. These IgG-coated erythrocytes are predominantly cleared by the spleen. A chill/fever reaction typically occurs after approximately 1 hour, the severity of which is related to the antibody titer and the number of antigenic sites. Hemoglobinemia reflects a backwash of hemoglobin into the circulation from the splenic macrophages and consequently never reaches the magnitude seen with intravascular hemolysis of ABO transfusion reactions. Hemolysis is nonetheless thorough, as evidenced by the rise in serum bilirubin.

Very low titers of antibody may lead to a delayed transfusion reaction. In such circumstances, reexposure to a foreign antigen through transfusion will cause an anamnestic rise in antibody titer. Clinically, hemolysis becomes evident 3 to 7 days posttransfusion as the patient exhibits jaundice and signs and symptoms of anemia.

Hemolytic Disease of the Newborn

During pregnancy, tiny amounts of fetal erythrocytes enter the maternal circulation. If these red cells bear foreign (paternally restricted) antigens, they will serve as an immune stimulus. Maternal IgG antibodies have the capacity to cross the placenta and coat fetal erythrocytes, leading to splenic sequestration and destruction. The magnitude of this hemolytic disease of the newborn is related to the titer of the maternal antibody; in severe cases, *in utero* exchange transfusion may be necessary to prevent fetal demise. Historically, anti-D has accounted for the majority of severe cases. With the prophylactic administration of Rh immune globulin, severe cases are now more frequently encountered with antibodies to c, E, or other minor blood group antigens.

Although ABO antibodies are largely naturally occurring and incapable of placental transfer because of their pentameric structure, small amounts of IgG type anti-A and anti-B may also be found, typically in type O individuals. These IgG antibodies are capable of placental transfer and, similar to anti-D, may cause hemolytic anemia. ABO hemolytic disease of the newborn, although much more frequent than that owing to anti-D, is typically less severe.

Anemia and evidence of accelerated erythropoiesis, the hallmarks of hemolytic anemias, are present in hemolytic disease of the newborn. Nucleated red blood cells are frequent and may be striking. Hyperbilirubinemia is largely of the unconjugated type owing to hepatic immaturity. Spherocytes are frequent in ABO-related hemolysis but are a minor feature of Rh incompatibility. In contrast, a strong DAT is more frequent in hemolytic disease of the newborn owing to anti-D; antibodies in the ABO system are loosely bound to the erythrocyte surface, resulting in a weakly positive DAT, and may be detected in the circulation with an indirect antiglobulin test.

Autoimmune Hemolytic Anemias

Autoimmune hemolytic anemias are caused by autoantibodies—self-induced antibodies directed at one's own red cells. The autoimmune hemolytic anemias are divided into two broad categories based on the class and the corresponding thermal activity of the responsible antibody: warm (IgG) and cold (IgM) autoimmune hemolytic anemias. Both warm and cold autoimmune hemolytic anemias may be further stratified into idiopathic and secondary forms.

Warm Autoimmune Hemolytic Anemias

The warm autoimmune hemolytic anemias are predominantly mediated through the binding of IgG antibodies to the erythrocyte's surface, resulting in splenic sequestration and destruction. Consequently, the morphologic hallmarks of this group of anemias is the spherocyte. The severity of the anemia and the number of spherocytes, however, are quite variable. In severe cases, the frequency of spherocytes rivals that seen in hereditary spherocytes. In such cases, the true identity of a warm autoimmune hemolytic anemia is revealed with a positive direct antiglobulin test. Red cells are typically coated with IgG occurring alone or in combination with varying amounts of intact C3. Rarely, the red cells exhibit only C3 on their surface. The responsible antibody may be serologically nonspecific; however, in many cases, preferential reactivity with the Rh locus is apparent. Seldom, however, is the antibody directed at specific Rh determinants.

Idiopathic Warm Autoimmune Hemolytic Anemia. In as many as half of the cases, the warm autoimmune hemolytic anemia is not associated with an underlying disease process and is termed idiopathic. The remaining cases are associated with or secondary to an underlying disease.

Secondary Warm Autoimmune Hemolytic Anemias. Secondary warm autoimmune hemolytic anemias most frequently arise in association with an underlying lymphoproliferative disorder or systemic autoimmune disorder. Chronic lymphocytic leukemia is the most frequently associated lymphoproliferative disease; warm autoimmune hemolytic anemia may be seen in

10% to 30% of cases sometime during the course of the disease. Other lymphoproliferative disorders are less frequently complicated by warm autoimmune hemolytic anemias but may also give rise to cold autoimmune hemolytic anemia. The onset of the anemia may precede the diagnosis of lymphoma/leukemia by months or years. Hence, thorough systemic evaluation is necessary before labeling an individual with an idiopathic autoimmune hemolytic anemia. Systemic lupus erythematosus is the prototype of the systemic autoimmune disorders associated with warm autoimmune hemolytic anemias. Rheumatoid arthritis, lupoid hepatitis, and ulcerative colitis may also exhibit red cell autoantibodies. Clinically significant hemolysis is less frequent than the finding of a positive DAT among these disorders. As in other forms of secondary autoimmune hemolytic anemia, treatment is directed at the underlying disease processes.

Cold Autoimmune Anemias: Cold Agglutinin Disease

As previously discussed, cold-reacting IgM antibodies may cause hemolysis through physical agglutination or complement activation. The thermal range of most IgM antibodies is sufficiently low that dissociation occurs before reaching temperatures compatible with complement activation. Consequently, hemolysis is frequently mild unless a broad range of thermal activity is present or ambient temperature sufficiently cool. Approximately 95% of cold-reacting antibodies show anti-I specificity, the remainder demonstrating activity with the i-antigen normally found on fetal erythrocytes.

Primary Cold Agglutinin Disease. Primary cold agglutinin disease is marked by episodic painful acrocyanosis induced by agglutination of red cells in the peripheral circulation on exposure to cold. It is typically a disease seen in the elderly and is managed largely by avoidance of cold environments. Exclusion of an underlying chronic lymphoproliferative disorder is mandatory. The diagnosis is suspected by the clinical history and the finding of red cell agglutination as blood is drawn or from blood smears prepared at room temperature. Cold agglutinin titers are established by serially diluting the patient's serum and incubating it with a suspension of group O erythrocytes. Normal individuals exhibit low concentrations of cold active anti-I, resulting in physiologic titers less than 1:16. Pathologic titers are typically greater than 1:256 and may exceed 1:100,000 in idiopathic cold agglutinin disease. The DAT typically reveals the presence of C3.

Secondary Cold Agglutinin Disease. Secondary cold agglutinin disease is associated with a variety of infections, the prototype being mycoplasma pneumonia and infectious mononucleosis. Cold agglutinin titers are often used as a supportive test in establishing a presumptive diagnosis of mycoplasma pneumonia. Titers rise within 1 to 3 weeks after the onset of symptoms. Titers typically reach 1:640, although on occasion they may reach values exceeding 1:1,000 and may be associated with symptomatic anemia. In contrast to the anti-I antibodies of mycoplasmal pneumonia, cold agglutinins arising in association with infectious mononucleosis are directed against the i-antigen and require fetal cells for detection. Cold agglutinin disease may occur secondarily in association with a variety of lymphoproliferative disorders. In these cases, the monoclonal IgM antibody expressed by the malignant B-cell clone is the responsible hemolytic agent. As previously stressed, discovery of either a warm or cold autoimmune hemolytic anemia necessitates a thorough investigation to rule out an underlying occult lymphoproliferative disorder.

Paroxysmal Cold Hemoglobinuria. Paroxysmal cold hemoglobinuria (PCH) is a rare acquired hemolytic anemia that may be seen as a complication of syphilis or a recent viral illness, particularly measles, mumps, or infectious mononucleosis. It is the result of an anti-P autoantibody occurring in individuals expressing the nearly universal P_1 and P_2 red blood cell antigens. This 7S IgG antibody, termed the Donath-Landsteiner antibody, binds to the erythrocyte at temperatures below 20°C. This IgG antibody is unique not only in its binding avidity at cold temperatures but also in its capacity to cause complement binding. Subsequent activation of complement occurs with warming; at the same time, the Donath-Landsteiner antibody dissociates from the red cell membrane. This biphasic property is utilized clinically to establish the diagnosis. Hemolysis occurs only after the blood has been chilled and then warmed to 37°C. Bound IgG can be detected in the DAT only when performed at temperatures less than 24°C; at body temperature, only C3 is detected. Clinical hemolysis is precipitated by exposure to cold and results in symptoms similar to those seen in acute intravascular hemolysis associated with ABO incompatibility. The disorder is usually a self-limited, postviral complication. Treatment for syphilis is required for luetic PCH.

Drug-Induced Autoimmune Hemolytic Anemia

The administration of a number of pharmaceutical agents has been complicated by the development of antibody-mediated hemolysis. Although, strictly speaking, many cases are not examples of true autoantibody-mediated hemolysis, these disorders are best considered as a group. Four basic hemolytic mechanisms have been described.

Hapten Mechanism. When given intramuscularly in large doses (greater than 10 million units), penicillin is absorbed to the red cell membrane. The combination of drugs and hapten (red blood cell membrane) serves as an antigen to which formed antibody binds. The antibody is usually of the IgG type and results in hemolysis through reticuloendothelial destruction. Discontinuation of the drug results in abrupt cessation of hemolysis, although the positive DAT usually persists for several months. Penicillin analogs such as cephalothin are also known to cause similar hemolytic complications.

Immune Complex or Innocent Bystander Mechanism. Several drugs, including quinidine, quinine, p-aminosalicylate, phenacetin, ethacrynic acid, and the nonsteroidal antiinflammatory agents, may result in hemolysis through the formation of immune complexes that nonspecifically coat red cells. These drugs bind to plasma proteins, which act as haptens. Antibodies then bind to the circulating drug-protein complex, and the resulting immune complex is deposited on a variety of innocent bystanders such as erythrocytes, platelets, and glomerular endothelium. The antibodies are often IgM, resulting in complement

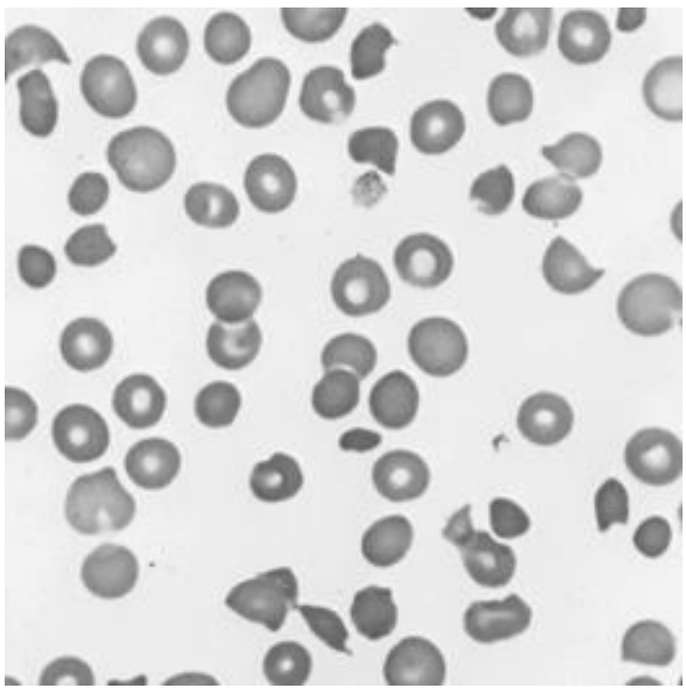


FIGURE 41.22. Variably sized, irregularly shaped erythrocyte fragments characterize the macroangiopathic and microangiopathic hemolytic anemias. Spherocytes may also be found. When seen in conjunction with red cell fragments, the presence of spherocytes should not raise confusion with hereditary spherocytosis or warm autoimmune hemolytic anemia.

deposition. The picture is usually one of acute intravascular hemolysis and does not require large amounts of drug. Renal failure becomes a problem in 50% of patients.

True Autoimmune Hemolytic Anemia: The Aldomet Model. The prototype for this type of drug-related hemolysis is Aldomet (methyldopa). The frequency of antibody formation with Aldomet therapy is directly linked to the dose of drug administered. Overall, approximately 15% of patients treated with Aldomet will develop a positive DAT after 3 months of treatment. The antibody is similar to the true autoantibodies encountered in idiopathic warm autoimmune hemolytic anemia, an IgG antibody exhibiting specificity toward the Rh locus. The exact mechanism of antibody formation is unknown, although it is postulated that a state of unregulated B-cell activity is created by the inhibitory effects of the drug on T-suppressor cells. A positive DAT does not equate to hemolysis, as clinical hemolysis develops in only approximately 1% of patients. Interestingly, it is a more frequent complication at lower dosages. To prevent hemolytic complications, the drug is discontinued after discovery of a positive DAT. Hemolysis seldom occurs after cessation of the drug, although a positive DAT may persist for months to years. L-Dopa, melanemic acid, flufenamic acid, chlordiazepoxide hydrochloride, cimetidine, and cefazolin have also been associated with similar induction of autoantibodies.

Nonimmunologic Absorption of Proteins. Although proven hemolysis is rare, a weakly positive DAT may develop in association with the administration of cephalothin or cephaloridine. These agents cause nonspecific absorption of a wide variety of proteins, including IgG, IgM, and complement to the red cell surface.

Paroxysmal Nocturnal Hemoglobinuria

Paroxysmal nocturnal hemoglobinuria (PNH) is an uncommon cause of hemolysis with a pathophysiology uniquely distinct from the hemolytic disorders previously discussed. PNH is an acquired clonal hematopoietic stem-cell disorder characterized by intravascular hemolysis and blood cells with characteristic defects. Multiple membrane proteins have been found to be lacking in the abnormal cells, including erythrocytes, granulocytes, monocytes, platelets, and, sometimes, lymphocytes. These proteins are all linked to the cellular membranes by a glycosylphosphatidylinositol (GPI) anchor. Recent work showed that the first step in the biosynthesis of the GPI anchor, which involves the synthesis of *N*-acetylglucosaminyl-phosphatidylinositol (GlcNAc-PI), is defective in patients with PNH. Although several genes may be involved in this process, *PIG-A*, mapped to the X chromosome, appears to play the primary role and defects within this gene have been found in most studied patients with PNH.

As a consequence of the absence of the GPI anchor protein, the affected proteins are degraded within the cell. The GPI-linked proteins belong to several different groups such as enzymes, receptors, complement regulators, adhesion molecules and other proteins without well-defined functions. These deficiencies can be complete or partial, which influences the severity of the symptoms. Two of the most important proteins lacking on the membrane are decay-accelerating factor (CD55) and membrane inhibitor of reactive lysis (CD59), which ordinarily protect cells from complement-mediated attack. Deficiency of these two proteins, particularly membrane inhibitor of reactive lysis, results in increased red blood cell susceptibility to complement lysis.

Traditional diagnosis of PNH has relied on the demonstration of complement-mediated hemolysis. In the acidified serum lysis test (Ham test), PNH red cells will lyse in acidified serum, which activates complement. The Ham test may not detect a small population of abnormal cells. The sucrose hemolysis test relies on the fact that a low ionic strength medium, such as sucrose, activates complement; PNH red cells will also lyse in serum mixed with isotonic sucrose. Although sensitive, the sucrose hemolysis test is less specific.

However, recent studies utilizing antibodies directed against GPI-anchored proteins such as CD55, CD59, CD16, and CD66b are shown to be much more practical and reliable for diagnosis of PNH. Flow cytometric analysis of granulocytes or red cells incubated with monoclonal antibodies appears to be sensitive and specific and can give information regarding the proportion of abnormal cells.

Although the name implies that PNH is characterized by episodes of hemolysis occurring at night, this classic presentation is not very common. In fact, the clinical manifestations of PNH are highly variable, but all patients have at least some degree of intravascular hemolysis. The extent of hemolysis and the degree of anemia vary greatly from patient to patient. Peripheral blood values and peripheral blood smear features are nonspecific. PNH may present with pancytopenia with marrow hypoplasia or aplasia, and the differential diagnosis of pancytopenia should always include PNH. However, the degree of bone marrow cellularity also varies from hyperplastic owing to erythroid hyperplasia to severe aplasia. Unexplained iron deficiency resulting from persistent hemoglobinuria and hemosiderinuria should also raise suspicion for PNH. Thrombosis, especially venous, is a grave complication of PNH and accounts for approximately one third to one half of deaths in patients with PNH. Patients with PNH/aplastic anemia may have a lower risk of thrombosis (5%). Patients with aplastic anemia appear to have an increased risk of developing PNH. Recent studies showed that GPI anchor–deficient hematopoietic cell populations frequently arise in the setting of long-term survival after aplastic anemia.

The median survival in PNH patients is between 10 and 15 years from the time of diagnosis. The natural course of the disease is often prolonged but marred by episodic bouts of hemolysis and a number of often fatal complications. Myeloid aplasia is accompanied by all the complications of neutropenia and thrombocytopenia. Anemia may be exacerbated by the development of iron deficiency with prolonged intravascular hemolysis or folate deficiency that may accompany brisk hemolysis and reticulocytosis. A small number of patients may develop AML. However, the incidence of AML in PNH appears to be similar to the risk of aplastic anemia patients developing AML. Therefore, it is possible that the PNH clone by itself may not be responsible for the increased risk of developing AML but rather that the aplastic anemia predisposes to clonal hematopoietic disorders such as AML. Some patients are reported to survive for prolonged periods (more than 25 years) and as many as 15% show

spontaneous recovery from PNH with no sequelae attributable to their disease.

Traumatic Red Cell Injury

Excessive physical trauma to the erythrocyte on its circulatory journey may lead to intravascular lysis or cell injury predisposing to premature red cell destruction (Fig. 41.22). Red cell fragments, or schistocytes, serve as the morphologic hallmark of this type of hemolysis. These beaten and battered red cells are of various sizes and shapes. Having withstood the initial injury, these resealed erythrocyte remnants eventually succumb to splenic destruction. Traditionally, the red cell fragmentation syndromes have been separated based on injury sustained in the heart or great vessels (macrovascular) versus that sustained in the prearteriolar or capillary bed (microvascular).

Macrovascular Hemolysis

Although capable of tolerating tremendous shear forces, erythrocytes become increasingly susceptible to injury with alterations in blood flow. Such changes in the macrovasculature occur with outflow disease such as aortic stenosis. A calcified valve is a double-edged sword, as turbulent jet flow is created past a roughened valvular surface. Similarly, prosthetic cardiac devices employed in the surgical repair of valvular or parenchymal defects may present bare or nonendothelialized surfaces injurious to red cells. The mechanical design of early prosthetic valves, which predispose to physical trauma and red cell fragmentation, represent another cause of macrovascular injury; this complication has now largely been resolved.

Hemolysis secondary to macrovascular injury is usually mild and well compensated. Prolonged periods of red cell injury may lead to depletion of iron stores through urinary iron loss accompanying intravascular hemolysis. Exacerbations of hemolysis may be encountered with sudden valvular incompetence and may be a sign of a failing prosthetic implant.

Microangiopathic Hemolytic Anemia

Microangiopathic hemolytic anemia is the result of red cell fragmentation in the prearteriolar or capillary bed. Erythrocyte damage occurs as the force of the blood flow carries red cells through the narrowed or damaged microvasculature. Lysis or sublethal assault occurs through two mechanisms: deposition of a fibrin meshwork that acts as a sieve through which red cells are pushed (Fig. 41.23) or diffuse endothelial damage, which exposes flowing red cells to a rough and hostile surface. The interplay of these two mechanisms can be seen with endothelial damage because the denuded endothelial surfaces cause fibrin deposition.

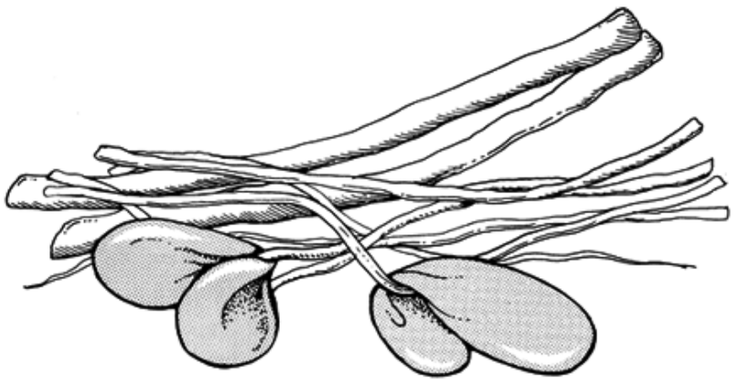


FIGURE 41.23. Fibrin strands in the microvasculature form a sharp mesh, causing fragmentation of circulating erythrocytes.

The causes of microangiopathic hemolytic anemia are numerous (Table 41.10). Disseminated intravascular coagulation, occurring as a complication of infections, disseminated carcinoma, or obstetric catastrophes; thrombotic thrombocytopenic purpura and the related hemolytic uremic syndrome; and malignant hypertension are the most frequent underlying disorders leading to red cell fragmentation and anemia.

TABLE 41.10. CAUSES OF MICROANGIOPATHIC HEMOLYTIC ANEMIA

Thrombocytic thrombocytopenic purpura
Hemolytic uremic syndrome
Disseminated intravascular coagulation
Infections: gram-negative septicemia
Snake bites
Hemolytic transfusion reactions
Obstetric complications: abruptio placentae, amniotic fluid embolism, retained dead fetus, eclampsia
Malignant hypertension
Immunologic disorders: vasculitis
Acute glomerulonephritis
Polyarteritis nodosa
Wegener's granulomatosis
Systemic lupus erythematosus
Scleroderma
Renal and hepatic allograft rejection
Disseminated carcinoma: gastrointestinal, breast, and pulmonary adenocarcinoma
Congenital vascular malformations
Cavernous hemangioma
Hepatic hemangioendothelioma

The peripheral blood findings are similar to those seen with macrovascular red cell injury. Additionally, thrombocytopenia is a frequent finding; this reflects the utilization of platelets and

formation of platelet thrombi in thrombotic thrombocytopenic purpura and hemolytic uremic syndrome or the consumption of platelets occurring with activation of the clotting cascade in disseminated intravascular coagulation.

Multisystem support may be necessary in treatment of the complications of anemia and coagulopathy. Nonetheless, therapy of the primary underlying disease must not be neglected.

March Hemoglobinuria

Repetitive forceful contact of body parts with hard surfaces can lead to erythrocyte lysis within the local microvasculature. This phenomenon was originally described in soldiers subjected to prolonged periods of marching. Indeed, long-distance walking or running is the most frequent cause. Because not all individuals are plagued by this complication, individual factors must come into play that govern the destructive forces affecting the erythrocyte.

Clinically, patients experience self-limited bouts of hemoglobinuria after such episodes of forceful contact. Red cell morphology is generally unremarkable, suggesting immediate lysis of erythrocytes within the local microvasculature. Differentiation from myoglobinuria can be established chemically, whereas other causes of hemoglobinuria must be distinguished based on clinical and laboratory features.

Drugs, Chemicals, and Venoms

Drugs

Administration of a number of drugs may lead to red cell damage and lysis through oxidative injury to hemoglobin, red cell membranes, or intracellular enzymes. Particularly vulnerable are individuals deficient in G6PD. In rare circumstances, the oxidant stress associated with administration of some drugs is sufficient to cause hemolysis in apparently normal erythrocytes.

Chemicals

Copper, arsenic, mercury, and lead are the most common heavy metals known to result in intravascular hemolysis after ingestion or inhalation. These chemicals are highly reactive with membrane thiols and interfere with red cell volume control.

Devastating intravascular hemolysis occurs within 2 to 12 hours of exposure to arsine fumes, produced by the action of water on metallic arsenide. The deficiency of ceruloplasmin found in Wilson's disease leads to the accumulation of copper in tissue, which may result in complications of hemolytic anemia. The introduction of significant amounts of water into the circulation has also been associated with hemolysis. This phenomenon has been reported as a complication of tissue irrigation accompanying prostate surgery or as a complication of near drowning in fresh water.

Venoms

The action of phospholipases released by *Clostridium perfringens* or contained within the venoms of some snakes or spiders can lead to hemolysis through dissolution of the red cell membrane. The most devastating example is the frequently fatal hemolysis of clostridial sepsis. Rapid hemolysis in this disorder is attributable to the phospholipase activity of the toxin and is grossly visible as deep cherry-red plasma. Microscopically, a kaleidoscope of red cell changes is visible, mimicking HPP.

The bite of the brown recluse spider results in necrotic arachnidism, an intense local area of tissue necrosis and ulceration, which may be complicated within days by hemolytic anemia resulting from the sphingomyelinase activity of the spider venom. Similar enzyme activity may be found in the venom of various snakes, particularly pit vipers. The intensity of hemolysis depends on the amount of venom absorbed.

Hypersplenism

Anemia may arise from primary splenic sequestration or destruction through a variety of mechanisms: hyperactivity of the splenic reticuloendothelial function; impediment of transit through the splenic cords; or increased transit time owing to splenic pooling. Prominent splenomegaly is common and results from a variety of conditions, including portal hypertension, extramedullary hematopoiesis, metabolic storage disorders, lymphoreticular neoplasia, and infection.

Anemia may be accompanied by thrombocytopenia and/or neutropenia mediated through the same mechanisms. The term hypersplenism has been used to describe such otherwise unexplained cytopenias. To properly invoke hypersplenism, one must demonstrate paradoxical peripheral cytopenia and marrow hyperplasia, both of which normalize after removal of the spleen or effective therapy of the underlying cause for splenic enlargement.

Polycythemia

A discussion of erythrocyte disorders would not be complete without consideration of polycythemia. Polycythemia, or erythrocytosis, is defined as an increased red cell volume. As with the functional definition of anemia, red cell volume is approached in the more practical quantitative terms of hemoglobin or hematocrit. Erythrocytosis therefore equates to an increased hemoglobin level or hematocrit.

Relative erythrocytosis, or stress erythrocytosis, is secondary to diminished plasma volume and must be distinguished from an absolute erythrocytosis, a true increase in red cell volume. Absolute erythrocytosis may occur as a primary autonomous expansion of red cell mass: polycythemia vera (PV), which is also discussed in Chapter 45.

PV must be differentiated from the many causes of erythrocytosis that occur secondary to the effects of excess erythropoietin. Elevated erythropoietin levels may be physiologically appropriate in response to a state of tissue hypoxia, which may arise through habitation at high altitudes, pulmonary disease, congenital heart disease with right-to-left shunts, hypoventilatory states, and high oxygen–affinity hemoglobinopathies. If the tissue hypoxia is reversed in these conditions, erythropoietin levels fall to normal, accompanied by a return of blood hemoglobin levels to normal.

Inappropriately elevated levels of erythropoietin may be encountered in the absence of tissue hypoxia. Several neoplastic and nonneoplastic disorders are associated with the production or release of erythropoietin or erythropoietin-like substances (Table 41.11). A few endocrine disorders have also been associated

with inappropriate erythrocytosis secondary to the effects of androgens on erythropoiesis. Androgens act by stimulating erythropoietin production and by directly promoting erythroid stem cells.

TABLE 41.11. CLASSIFICATION OF ERYTHROCYTOSIS

Relative: Diminished plasma volume
Absolute
Primary: Polycythemia vera
Secondary:
Physiologically appropriate
High altitude
Pulmonary disease
Pickwickian syndrome
Cardiovascular disease with right-to-left shunt
High oxygen-affinity hemoglobinopathies
Congenitally decreased erythrocyte 2,3-diphosphoglycerate
Physiologically inappropriate
Neoplasms
Renal cell carcinoma
Cerebellar hemangioblastoma
Hepatocellular carcinoma
Uterine leiomyoma
Adrenal adenoma
Ovarian carcinoma
Nonneoplastic renal disease
Renal cysts
Hydronephrosis

The approach to the patient depends on identifying the primary or secondary nature of the erythrocytosis. The peripheral blood and bone marrow findings are largely limited to the elevated red cell mass except in the case of PV, in which characteristic increases in all three cell lines are encountered. Splenomegaly, basophilia, elevated leukocyte alkaline phosphatase levels, and increased levels of vitamin B_{12} or B_{12} binding proteins are additional associated findings of PV. The underlying cause of the decreased plasma volume leading to relative erythrocytosis is usually evident. Similarly, symptoms of hypoxemia are often readily recognizable. Geographic altitude and the extreme obesity of the Pickwickian syndrome are difficult to disguise. Hemoglobin electrophoresis, radiographic evaluation, and quantitative erythropoietin levels are useful after exhaustion of the more common causes of erythrocytosis.

Ideally, one would like to maintain the hematocrit below 52% through periodic phlebotomy. As the hematocrit rises above this level, so does blood viscosity, placing the patient at risk for thrombosis, particularly in areas of low flow rates. Additionally, as viscosity increases, the oxygen-carrying capacity of blood decreases; in cases of hypoxia-related secondary erythrocytosis, this will serve to further stimulate erythropoietin production.

SUGGESTED READINGS

Andrews NC. Disorders of iron metabolism. *N Engl J Med* 1999;341:1986–1995.

Bessler M, Mason PJ, Hillmen P, et al. Paroxysmal nocturnal haemoglobinuria (PNH) is caused by somatic mutations in the PIG-A gene. *EMBO J* 1994;13:110–117.

Cook JD. The measurement of serum transferrin receptor. *Am J Med Sci* 1999;31:269–276.

Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 1996;13:399–408.

Feder JN, Tsuchihashi Z, Irrinki A, et al. The hemochromatosis founder mutation in HLA-H disrupts beta2-microglobulin interaction and cell surface expression. *J Biol Chem* 1997;272:14025-14028.

Feder JN, Penny DM, Irrinki A, et al. The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. *Proc Natl Acad Sci U S A* 1998;95:1472–1477.

Fukuda MN. HEMPAS. Hereditary erythroblastic multinuclearity with positive acidified serum lysis test. *Biochim Biophys Acta* 1999;1455:231–239.

Gasparini P, Miraglia del Giudice E, Delaunay J, et al. Localization of the congenital dyserythropoietic anemia II locus to chromosome 20q11.2 by genomewide search. *Am J Hum Genet* 1997;61:1112–1116.

Hoffman R, Benz EJ, Shattil SJ, et al., eds. *Hematology: Basic principles and practice*, 3rd ed. Philadelphia: Churchill-Livingstone, 2000.

Jandl JH. *Blood: Textbook of hematology*, 2nd ed. Boston: Little, Brown, 1996.

Lee GR, Roerster J, Lukens J, et al., eds. *Wintrobe's clinical hematology*, 10th ed. Baltimore: Williams & Wilkins, 1999.

Lebron JA, Bennett MJ, Vaughn DE, et al. Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. *Cell* 1998;93:111–123.

Lind L, Sandstrom H, Wahlin A, et al. Localization of the gene for congenital dyserythropoietic anemia type III, CDAN3, to chromosome 15q21-q25. *Hum Mol Genet* 1995;4:109–112.

Miyata T, Yamada N, Iida Y, et al. Abnormalities of PIG-A transcripts in granulocytes from patients with paroxysmal nocturnal hemoglobinuria. *N Engl J Med* 1994;27:330:249–255

Sandstrom H, Wahlin A. Congenital dyserythropoietic anemia type III. *Haematologica* 2000;85:753–757.

Tamary H, Shalmon L, Shalev H, et al. Localization of the gene for congenital dyserythropoietic anemia type I to a <1-cM interval on chromosome 15q15.1-15.3. *Am J Hum Genet* 1998;62:1062–1069.

Tse WT, Lux SE. Red blood cell membrane disorders. *Br J Haematol* 1999;104:2–13.

Wickramasinghe SN. Congenital dyserythropoietic anemias. *Curr Opin Hematol* 2000;7:71–78.

Wickramasinghe SN. The wide spectrum and unresolved issues of megaloblastic anemia. *Semin Hematol* 1999;36:3–18.

42

The Thalassemia and Hemoglobinopathy Syndromes

James D. Hoyer

Donald R. Hoffman

Much of our knowledge concerning the structure and function of proteins and the genes that control their production has its origin in studies of the hemoglobin molecule.

The history of these studies is peppered with landmark discoveries. The first observation that more than one form of hemoglobin existed, unrelated to its state of oxygenation, was that of Körber who, in 1866, reported his finding of two distinct hemoglobins in the newborn based on their resistance to denaturation by acid and alkali (1). In 1910, Herrick coined the term *sickle cell* to describe the shape of red cells seen in the blood of a student from the Caribbean who had chronic anemia (2). In 1925, Cooley and Lee (3) suggested that the chronic anemia, splenomegaly, and bone changes seen in a series of five children represented a distinct type of anemia. The link between Cooley's anemia and the hemolytic anemia described in the Italian literature, known as the Rietti-Greppi-Micheli type (4,5 and 6) and later as Mediterranean anemia, had to await the work of Wintrobe and colleagues (7). The term *thalassemia*, a euphonious attempt to convert various previous titles into one word, was first suggested by Whipple and Bradford (8). In fact, the word suggests sea (*thalassa*) in the blood (*-emia*) and shows a degree of Western chauvinism because there probably are more people with thalassemia in Asia than there are people in the Mediterranean basin.

Pauling and colleagues (9) are usually given credit for first demonstrating a chemical abnormality in an inherited hemoglobin disorder when in 1949 they observed an electrophoretic difference between adult hemoglobin (Hb A) and sickle hemoglobin (Hb S). However, 1 year earlier, Hörlein and Weber (10) suggested that abnormal hemoglobin might explain their findings in a patient with hereditary methemoglobinemia [later shown to be Hb-M Saskatoon (11)]. The search for the specific defect in abnormal hemoglobin culminated in 1956 when Ingram (12) pinpointed the amino acid substitution in Hb S. A torrent of discoveries followed this observation. To date, more than 750 structurally abnormal hemoglobins (13) have been defined. Establishing the structural and functional relationships of the various molecular lesions became possible largely through the detailed radiographic crystallographic analysis of the hemoglobin molecule conducted by Perutz and colleagues (14,15) that began in the 1960s and continues.

The genetic basis for many of these disorders became evident to several investigators, and the relationship of the more common forms to malaria was suggested by Haldane (16) in 1949. The exact details of the number of globin genes and the precise definition of the lesions that underlie the various disorders had to await the explosion of knowledge resulting from the development of DNA and RNA technologies. The abilities of the cytogeneticist to look at the details of the chromosome have now converged with those of the “gene prober” to produce a complete picture of the globin genes.

The purpose of this chapter is to relate this knowledge to a discussion of the hemoglobinopathies, emphasizing the laboratory aspects of this varied group of disorders.

STRUCTURE AND FUNCTION OF NORMAL HEMOGLOBIN MOLECULES

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

Normal human hemoglobin molecules consist of two pairs of polypeptide (globin) chains, each of which carries a heme portion. The tetramer is roughly spherical in shape (64 × 55 × 40 Å) with a molecular weight (MW) of 64,400 d; the globin portion makes up approximately 96% of the molecule. The hemoglobins that are produced at various stages of embryonic, fetal, and adult life are listed in Table 42.1. Each of these hemoglobins consists of a pair of α- or α-like (ζ) globin chains and a pair of ß- or ß-like globin chains (δ,γ,ε). There is considerable homology within each of these two groups. The structure of the various globin chains is known precisely (17) and can be considered under four headings. The primary structure consists of 141 amino acids in the α-globin chain and 146 amino acids in the ß- and ß-like globin chains. The primary sequence of the ß and δ chains is very similar and differs in only 10 of the 146 residues. The secondary structure involves the α-helical arrangement of amino acids present in 75% of each globin chain. There are eight α-helical segments in the ß chains (letters A through H) and seven in the α chains. The tertiary structure is represented by the larger folds superimposed on the α-helical arrangement that occur in the nonhelical portions of the molecule. Finally, the quaternary structure involves the interrelationship between the four globin chains.

TABLE 42.1. THE NORMAL HUMAN HEMOGLOBINS

Hemoglobin	Globin Chains	Major Source	Embryo	Neonate	Adult
			%	%	%
Gower 1	ζ_2/ε_2	Yolk sac	50	0	0
Gower 2	α_2/ε_2	Yolk sac	25	0	0
Portland	ζ/γ_2	Yolk sac	25	0	0
Hb F	$\alpha_2/{}^{G}\gamma_2$[a]	Liver, spleen	0	75	<1
	$\alpha_2/{}^{A}\gamma_2$	Liver, spleen	0	75	<1
Hb A	α_2/β_2	Bone marrow	0	25	97
Hb A^2	α_2/δ_2	Bone marrow	0	<1	3

[a] Two structurally and genetically distinct γ-globin chains are normally produced, one with glycine (^{G}) and the other with alanine (^{A}) as the 136th amino acid residue.

The heme portion (MW 614 d) of the hemoglobin molecule consists of a porphyrin ring with an iron atom in its center. This porphyrin ring is held in place by two histidine residues in each

globin chain. The iron atom is covalently bonded to the proximal histidine at F8 (eighth amino acid in the F helix) and is stabilized by a large number of other interatomic links within the heme pocket. There are four heme groups in each hemoglobin molecule, allowing it to bind as many as four molecules of oxygen.

The hemoglobin molecule is beautifully suited to its main function—the delivery of oxygen. It has the ability to pick up oxygen at the relatively high oxygen tension in the lungs and deliver it at the lower oxygen tension in the tissues, remaining soluble throughout, undamaged by the oxygen it carries, and utilizing none of it. The hemoglobin molecule exhibits a cooperative binding effect; that is, as each heme group becomes oxygenated, it is easier for the other heme groups to become oxygenated as well. This is attributable to a complex set of steric changes within the molecule that occur at the interface between the four globin subunits. The differences between oxyhemoglobin and deoxyhemoglobin involve numerous amino acids at or near the points of contact between the globin chains. Therefore, amino acid substitutions in these positions can significantly alter heme function and result in high or low oxygen affinity hemoglobin variants.

Several factors facilitate the delivery of oxygen. Unrelated to the hemoglobin molecule itself are the large surface area and pliability of red cells, the blood flow that is systemically and locally controlled, and the rate of diffusion of oxygen into and through the cells adjacent to capillaries. In addition, the oxygen affinity of some cells, particularly those containing myoglobin, attracts oxygen into those cells. Factors directly related to the hemoglobin molecule, and therefore of importance when considering the hemoglobinopathies, include the amount of hemoglobin circulating and its ability to deliver oxygen. The latter is represented by the difference between the oxygen content of arterial and venous blood and is dependent on the oxygen dissociation curve, which indicates the amount of oxygen delivered or released for a unit decrease in oxygen tension (PO_2).

The oxygen dissociation curve (Fig. 42.1) can be depicted as the relationship between the amount of the hemoglobin that is in the form of oxyhemoglobin (percentage of saturation) and the partial pressure of oxygen (oxygen tension) in the surrounding medium. The sigmoidal shape of this curve is owing to the cooperative binding effect as previously described. The oxygen affinity of hemoglobin is often reported as the P_{50}, which is that partial pressure of oxygen at which the hemoglobin is half-oxygenated; P_{50} is inversely related to oxygen affinity. Oxygen dissociation curves of clinical import are those obtained from whole blood under standard physiologic conditions: temperature of 37°C, a PCO_2 of 40 mm Hg, and a pH of 7.4. The P_{50} for normal men is close to 27 mm Hg; it is slightly higher in women. Under normal circumstances, hemoglobin is 97% saturated in the lungs and approximately 75% saturated at the normal mixed venous oxygen tension of 40 mm Hg.

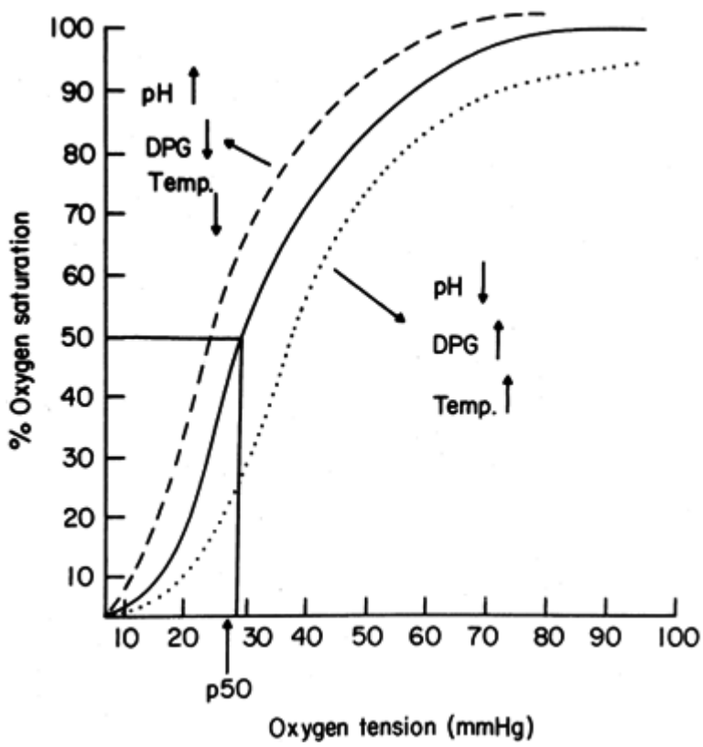


FIGURE 42.1. The oxygen dissociation curve of human adult hemoglobin (Hb A). Alteration in the position of the curve resulting from changes (↑, increase; ↓, decrease) in pH, 2,3-diphosphoglycerate (DPG), and temperature (temp) are indicated by the dashed and dotted lines. (From Weatherall DJ, Clegg JB. *The thalassemia syndromes*, 3rd ed. Oxford: Blackwell Scientific Publications, 1981, with permission.)

There are three main factors, other than the structure of the hemoglobin molecule itself, that affect the shape and position of the oxygen dissociation curve (Fig. 42.1): pH, temperature, and 2,3-diphosphoglycerate (2,3-DPG). Bohr and colleagues (18,19) were the first to show that oxygen affinity was reduced by carbon dioxide. The reciprocal relationship between oxygen affinity and carbon dioxide affinity (an obvious advantage in a system in which oxygen is adsorbed in the lungs and delivered to the tissue, while the reverse is occurring for carbon dioxide) was later shown to be effected by a change in pH. As carbon dioxide is expelled through the lungs, pH increases and oxygen affinity

increases; there is a shift to the left in the oxygen dissociation curve. The lowering of pH that occurs in the tissues has the opposite effect, decreasing the oxygen affinity. Thus, at either end of the system, the Bohr effect of a change in pH is beneficial to the uptake or delivery of oxygen.

Temperature change also alters oxygen affinity (20) in an apparently appropriate fashion. Increase in temperature, with its increased metabolic demands, reduces oxygen affinity and thus increases delivery of oxygen. Conversely, reduction in temperature increases oxygen affinity and reduces delivery of oxygen.

2,3-DPG, the most abundant organic phosphate in the red cell at approximately 5 mmol/L of packed red cells, has a marked effect on the oxygen affinity of hemoglobin (21). 2,3-DPG lowers oxygen affinity in two ways: by binding preferentially to the ß chain of deoxyhemoglobin, thus tending to stabilize it, and by reducing intracellular pH. The concentration of 2,3-DPG within the red cell does not remain constant; it varies inversely with the hemoglobin concentration both in anemic and in normal individuals. Thus, as the hemoglobin falls, the amount of 2,3-DPG increases, oxygen affinity decreases, and more oxygen is unloaded from the hemoglobin.

In summary, hemoglobin, aided by several mechanisms active in the red cell, is a very efficient respiratory pigment. Hemoglobin within the red cell meets the four requirements suggested by Barcroft in 1928 for such a pigment, namely, the ability to (a) carry large quantities of oxygen, (b) take up and release the oxygen at appropriate pressures, (c) remain soluble, and (d) act as a buffer (22).

GENETIC CONTROL

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

Each globin chain is under separate genetic control. The α-gene cluster lies at the extreme end of the short arm of chromosome 16 (16p13.3 pter) (23) and consists of four genes and three pseudo-genes[a] (Fig. 42.2) (24). There are two functional α-genes on each chromosome 16-α1 and α2; they have minor structural differences but transcribe into mRNAs that encode identical α-globin products in unequal quantities; α1 is responsible for approximately one third of the α chains and α2 for approximately two thirds (25).

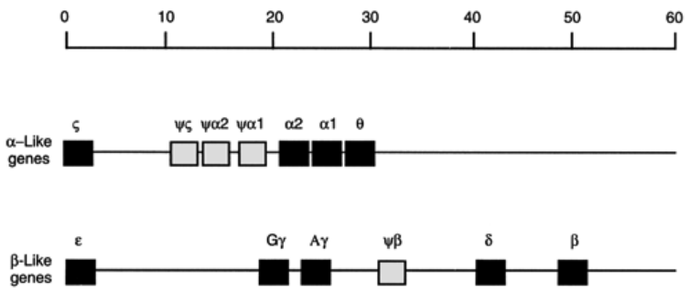


FIGURE 42.2. Relative positions of genes and pseudogenes of the α-like (chromosome 16) and ß-like (chromosome 11) series. The scale represents kilobases; the sequences from left to right represent 5′ (N-terminal) to 3′ (C-terminal) positions on the chromosome.

The ß-globin gene cluster lies on the short arm of chromosome 11 distal to band p14 and between the genes for parathyroid hormone and insulin (26). It consists of five genes (ε, ^{G}γ, ^{A}γ,δ, ß) and one pseudogene (Fig. 42.2). The two γ-genes encode for γ chains that differ slightly structurally, ^{A}γ having alanine and ^{G}γ having glycine as the 136th amino acid residue.

Each globin gene has a similar organization, consisting of three exons and two introns or intervening sequences. All five regions are transcribed into nuclear or pre-mRNA, but the RNA from the introns is spliced out in the formation of cytoplasmic mRNA, which therefore contains only information transcribed from the exons. Regions close to the gene (usually upstream at the N-terminal or 5′-end) are involved in initiating and controlling gene function. Hence, mutations occurring near a gene as well as those within a gene may significantly affect the amount or the structure of its ultimate product.

Under normal circumstances, the product of an α- or α-like gene combines with the product of a ß- or ß-like gene to form a hemoglobin tetramer (Table 42.1). As maturation occurs *in utero* and postpartum, the activity of the genes changes (Fig. 42.3). This change has a downstream direction with the embryonic genes, active in yolk sac erythropoiesis, ceasing to function at approximately the 10th week to be superseded by α- and γ-(hepatic and splenic erythropoiesis); γ-chain production is gradually replaced by ß-chain production starting at approximately the sixth week of fetal life and proceeding until the sixth to 12th month of extrauterine life. By that time, γ-chain production is less than 1% that of the ß chain. Production of the δ chain is approximately 0.5% of total globin-chain production at birth and rises to adult levels by 1 year. (In the vocabulary of the hemoglobin chains, adulthood is reached by approximately 1 year.)

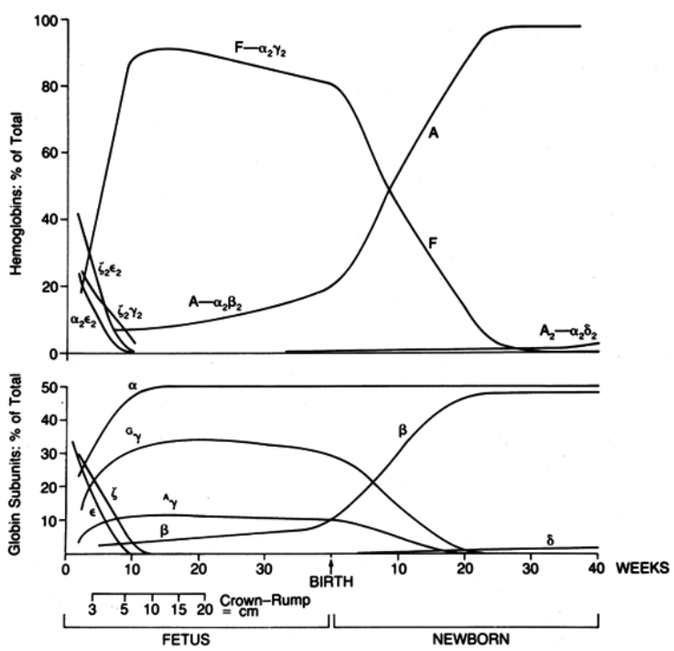


FIGURE 42.3. Changes in globin-chain and hemoglobin production during human development. (From Bunn HF, Forget BG. *Hemoglobin: molecular, genetic and clinical aspects*. Philadelphia: W.B. Saunders, 1986, with permission.)

Clinically, the most significant of these changes is the γ to ß switch, which explains why at birth a normal neonate possesses approximately 75% to 90% Hb F ($\alpha_2\gamma_2$) and 10% to 25% Hb

A ($\alpha_2\beta_2$), whereas by approximately 12 months, the infant possesses less than 1% Hb F, approximately 96% Hb A, and approximately 3% Hb A_2 ($\alpha_2\delta_2$).

CLASSIFICATION AND NOMENCLATURE

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

Most hemoglobinopathies are inherited abnormalities. A few result from *de novo* mutation (27), and there is a small but fascinating group of acquired defects mimicking α thalassemia and persistence of fetal hemoglobin (28,29).

The inherited group of disorders can be divided into three broad categories (Table 42.2):

- *Structural, or qualitative abnormalities*, in which the amino acid sequence of one or more of the globin chains is altered (by substitution, addition, deletion, or fusion) as a result of an incorrect DNA code. Hb S is a common example.
- *Quantitative abnormalities*, in which the production of one or more of the globin chains is reduced or absent (owing to inadequate amounts of mRNA, to unstable mRNA, or to mRNA that contains an untranslatable "nonsense" message) resulting in an imbalance in globin-chain production. These form the various thalassemia syndromes that are classified according to the globin chain whose production is deficient.
- *Hereditary persistence of fetal hemoglobin* (HPFH), in which there is complete or partial failure of the γ to ß switch. As noted later, there is an overlap between HPFH and the ß thalassemias, resulting in a spectrum with HPFH at one end, ß thalassemia at the other, and δß thalassemia as an intermediate.

TABLE 42.2. A CLASSIFICATION OF THE MORE COMMON HEMOGLOBINOPATHIES

Structural defects
Involving Hb S
Without *in vivo* sickling
Hb AS (sickle cell trait)
Hb AS-G-Philadelphia
Hb AS α-thalassemia
Hb S HPFH
With *in vivo* sickling
Hb SS (sickle cell anemia)
Hb SC
Hb SD-Los Angeles
Hb SO-Arab
Hb Sß-thalassemia
HB C disorders
Hb D disorders
Hb E disorders
Hb G-Philadelphia disorders
Unstable hemoglobins
Hemoglobins with high or low O_2 affinity
Hb M (methemoglobinemia)
Globin chain imbalance (thalassemia)
α-Thalassemia
Silent (one-gene defect)
Thalassemia minor (two-gene defect)
Hb H disease (three-gene defect)
Hb Barts hydrops fetalis (four-gene defect)
Result of a structural defect
Hb constant spring
ß-thalassemia
$ß_0$-Thalassemia
$ß_+$-Thalassemia
Result of a structural defect
Hb Lepore
Hb E
Unstable hemoglobins (some)
δß-Thalassemia
HPFH
Black type
Swiss type
Greek type
Associated with structural defect
Hb Kenya

Hb, hemoglobin; HPFH, hereditary persistence of fetal hemoglobin.

Because many of these disorders are common, combinations occur as a result of compound heterozygosity for abnormal genes (Fig. 42.4). Furthermore, structural defects, such as Hb Lepore, or Hb E may cause an imbalance in globin-chain production and hence a thalassemia syndrome. In general, the various genetic abnormalities are inherited in a codominant fashion.

Nomenclature commonly used for the hemoglobinopathies is somewhat confusing because as knowledge has been accumulated, it has not always been translated into more accurate designations. The structural hemoglobin variants can be specifically named according to their structural abnormality. Thus, sickle hemoglobin is $\alpha_2\beta_2^{6glu\rightarrow val}$ or $\alpha_2\beta_2^{6(A3)glu\rightarrow val}$, indicating that the sixth amino acid residue [or the third amino acid in the first (a) α helical segment] of the ß-globin chain, which is normally glutamic acid (GAG), is replaced by valine. However, this is difficult to verbalize, if not to write; hence, it is generally referred to as Hb S.

Normal adult hemoglobin is designated Hb A, and fetal hemoglobin as Hb F. As hemoglobin variants were discovered, they were given letters of the alphabet, beginning with Hb C (Table 42.3). However, it was soon realized that the alphabet would not be large enough to accommodate all hemoglobin variants, and a somewhat disorganized system of common nomenclature evolved. Because several hemoglobin variants can show identical mobilities (particularly on alkaline electrophoresis), they may be given the same letter designation followed by the place of origin of the individual carrying the hemoglobin variant. For example, there are several variants with the Hb D designation that all migrate in the S position or alkaline electrophoresis. The most common of these is Hb-D Los Angeles (D-Punjab).

However, other hemoglobin variants have been named in a variety of ways, such as after countries (Hb-Nigeria, Hb-Mexico), states (Hb-Ohio, Hb-Michigan) cities or towns (Hb-Seattle, Hb-Mulmö), urban areas (Hb-Queens), rivers (Hb-Volga), hospitals (Hb-Barts), and even historic figures (Hb-Abraham Lincoln, Hb-Osler). Another area of confusion occurs when hemoglobin variants were given different designations but found subsequently to be structurally identical. For example, Hb G-Philadelphia, $\alpha_2^{68[E17]Asn\rightarrow lys}\beta_2$ has nine different names attached to it.

In the laboratory, the specific designation of the structural abnormality would seem to be the most logical nomenclature to use. However, at least for the common varieties, the exact structure is seldom verified, so the eponym is usually applied.

When referring to heterozygotes, shorthand is used, listing the hemoglobin present in greatest quantity first. Hb AS, for example, refers to a sickle cell trait in which there is approximately 55% Hb A and 45% Hb S. The same system can be applied to double heterozygotes; for instance, Hb SC indicates an individual who inherited a β^s gene from one parent and a β^c gene from the other.

Hb S β thalassemia indicates an individual who inherited a β^s gene from one parent and a β-thalassemia gene from the other (Fig. 42.4).

The nomenclature commonly used for the thalassemia syndromes is relatively simple but somewhat less accurate; the Greek letter of the globin chain that is reduced in quantity is used. Thus, β thalassemia indicates reduction of β-chain production, α thalassemia reduction of α-chain production, and so on. However, there are almost 200 different lesions in or near the α gene that can cause β thalassemia. Attempts to name the various subsets of the thalassemias are discussed later.

The term *trait* is used for heterozygotes (sickle cell trait, Hb C trait) and disease for the homozygote (sickle cell disease, Hb C disease); however, not all authors would agree that sickle cell anemia is synonymous with sickle cell disease. This system becomes unwieldy when four genes are involved, as in α-chain abnormalities such as Hb G-Philadelphia. The adjectives minor and major, used in describing the thalassemias, were originally used in both a clinical and a genetic (heterozygote and homozygote)

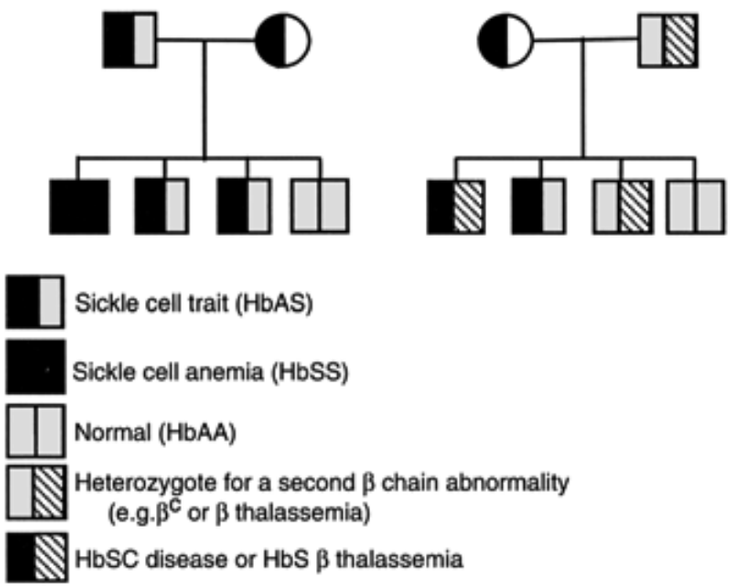


FIGURE 42.4. Inheritance pattern of hemoglobinopathies. β-Globin chain defects are used in these examples. α-Globin chain defects have a more complicated inheritance pattern because two α-globin genes are inherited from each parent. (Reprinted with permission from Hoffman GC. The sickling disorders. *Lab Med* 1990;21:797–807.)

TABLE 42.3. ORDER OF DISCOVERY OF HEMOGLOBIN VARIANTS

Hemoglobin Variant	Year
Hb S	1949
Hb C	1950
Hb D-Los Angeles	1951
Hb E	1954
Hb G (now Hb Korle-bu)	1955
Hb H	1955
Hb I	1955
Hb J	1956
Hb K	1957

Hb, hemoglobin.

sense. As the large number of different thalassemia genes with varying clinical expressions became known, the terms have reverted to a strictly clinical interpretation (Table 42.4).

TABLE 42.4. CLASSIFICATION OF THE THALASSEMIA SYNDROMES ACCORDING TO CLINICAL SEVERITY

Thalassemia major	Severe anemia; red cell transfusion required to maintain life.
Thalassemia intermedia	Moderate anemia; red cell transfusion occasionally required.
Thalassemia minor	Slight if any anemia, microcytic red cells, often with erythrocytosis.
Thalassemia minima	Silent thalassemia; no clinical or hematologic abnormality. Demonstrable by family studies or gene analysis.

INCIDENCE, DISTRIBUTION, AND THE MALARIA HYPOTHESIS

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

Most of the more than 700 hemoglobinopathies are rare and known to occur in a handful of families at the most; the affected individuals are almost exclusively heterozygotes, with homozygotes occurring mainly in consanguineous families. The common abnormalities, particularly those that have reached polymorphic levels in some populations, such as Hb S, Hb C, Hb E, and the α and ß thalassemias achieved their popularity through natural selection.

Inherited abnormalities in the three major components of the red cell (membrane, hemoglobin, and a rudimentary glycolytic pathway) represent some of the best examples of Darwinism in action among human populations. All are related to malaria (30). The remarkably high incidence of stomatocytic elliptocytosis in Melanesia is an example of a membrane abnormality that is beneficial in areas of endemic *Plasmodium falciparum* malaria (31). The parasite has difficulty entering these abnormal red cells. The Duffy blood group provides another striking example (32); *Plasmodium vivax* is unable to enter Duffy-negative red cells. In Africa, Duffy-negative individuals, Fy(a-b-), are found mainly in western regions where *P. vivax* malaria is not found. In this situation, it appears that the blood group distribution preceded and prevented the spread of *P. vivax* from the east (33).

This Duffy relationship is completely different from that with Hb S, whose incidence is greatest where malaria (in this case *P. falciparum*) is most prevalent. There are areas in central Africa in which the incidence of sickle cell trait reaches 40%. This high incidence is maintained (despite the loss of sickle genes from the gene pool with the neonatal death of most Hb SS individuals) because falciparum malaria is much milder and seldom lethal in Hb AS individuals than in normal (Hb AA) individuals. This is an example of balanced polymorphism in which the heterozygote has an advantage over both the normal and abnormal homozygote (34). The high incidence of Hb E in southeast Asia can also be explained by the malaria hypothesis (35).

The distribution of α and ß thalassemias also matches the distribution of malaria. The classic studies of Siniscalco and colleagues (36) in Sardinia and Flint et al. (37) in Melanesia are two among many population studies demonstrating this relationship. In Sardinia, the incidence of ß thalassemia minor [and glucose-6-phosphate dehydrogenase (G6PD) deficiency] is directly related to the incidence of malaria, which in turn is related to the altitude at which the populations live (mosquitoes being limited to valleys by temperature and water requirements). In Melanesia, malaria and α thalassemia were shown to be directly related, the diagnosis of α thalassemia being proved by gene analysis. One exception to the relationship between malaria and α thalassemia is the relatively low prevalence of the latter in central and western Africa, where Hb S is the major hemoglobinopathy. A likely explanation is that two mutations, both beneficial, are unlikely to coexist if both homozygotes are lethal. Hb S seems to have won, perhaps as the more beneficial trait with respect to *P. falciparum* infection.

G6PD deficiency is another guard against malaria, and, like sickle cell trait, has a high incidence in Africa (38).

The incidence of Hb S (present in approximately 8% of African Americans), G6PD deficiency (with a similar incidence), α thalassemia minor (approximately 30% of African Americans have single-gene deletion type), and ß thalassemia minor (3% to 5% among African Americans and individuals of Mediterranean extraction) can be explained by their origin from areas of high malarial endemicity.

LABORATORY APPROACH

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

The diagnosis of all hemoglobinopathies eventually lies in the laboratory. The number and degree of sophistication of the tests needed depend on the clinical needs of the patient population and the interests of the investigator. At one extreme, a simple solubility test to rule out the presence of Hb S may suffice, whereas at the other, precise amino acid sequencing or gene analysis may be required.

There are three main reasons that a laboratory may be asked to investigate the possible existence of a hemoglobinopathy. There may be a hint in the patient's history or physical examination; the blood count may indicate the possibility; or the investigation may be required as part of a screening program for either a population such as neonates or individuals as a preoperative or other type of requirement. Hemoglobinopathies of real or potential clinical significance will usually present with some abnormality in the blood count. However, many of the rarer examples detected in screening programs are noticed because of an abnormal electrophoretic pattern; many of these will be of little or no clinical significance but nevertheless need to be investigated. Whatever the entry point, hemoglobin electrophoresis and a blood count, if not already performed, will almost certainly be required.

In this section, the application and value of the more commonly used tests are discussed. Some more esoteric tests, including those involving molecular and gene studies, are mentioned in later sections dealing with specific diseases.

Blood Count and Red Cell Morphology

Three types of abnormal blood counts are associated with clinically significant hemoglobinopathies:

1. Microcytic hypochromic red cells often without significant anemia (as in thalassemia minor);
2. A hemolytic anemia (as in sickle cell anemia and Hb C disease);
3. A combination of the two (as in thalassemia intermedia and major).

In addition to these common findings, a physiologically appropriate erythrocytosis is associated with high oxygen affinity hemoglobins, and the appropriate anemia is associated with low oxygen affinity hemoglobins.

Red cell morphology may provide a general or specific indication of a hemoglobinopathy. The latter include sickle cells, Hb C crystals, and the bizarre-shaped red cells seen in Hb SC disease.

Less specific and often requiring additional staining techniques are Heinz bodies associated with many unstable hemoglobins and the Hb H inclusions of Hb H disease and some forms of α thalassemia minor. Target cells are particularly common when Hb C or Hb D is present.

Heinz bodies, which consist of aggregates of denatured hemoglobin within the red cell, are seen in two main situations: oxidative hemolysis and instability of the hemoglobin molecule. They are not visible in red cells stained with Wright's or other trichrome stains; vital or supravital stains such as those used to demonstrate reticulocytes are required. Crystal violet or methyl violet has an advantage over stains such as new methylene blue and brilliant cresyl blue because the latter redox stains are themselves able to cause Heinz body formation. However, this latter property may be advantageous in situations in which few Heinz bodies are present. This concept can be carried further by the use of phenylhydrazine, a drug that in correct concentration *in vitro* will cause the formation of many more Heinz bodies in red cells containing an unstable hemoglobin (or with other abnormalities such as G6PD deficiency that predispose to Heinz body formation) than in normal red cells (39,40). Whichever staining method is use, normal controls must be incorporated in the procedure.

Hb H (β_4) is an unstable hemoglobin composed of four ß chains and forms a distinct type of Heinz body when exposed to a redox dye such as brilliant cresyl blue or new methylene blue (41). A suspension of red cells containing Hb H when incubated at 37°C for 1 hour with a few drops of 1% citrate/saline solution of brilliant cresyl blue will contain many round, small-blue-stained inclusions of precipitated Hb H, giving the cell a golf-ball appearance. These inclusions can be distinguished from the precipitated RNA of reticulocytes and the larger Heinz bodies associated with other unstable hemoglobins under similar test conditions.

Solubility Test

The relative insolubility of deoxygenated Hb S compared with other hemoglobins is the basis of a simple test for its presence (42).

Several acceptable commercial kits are available that consist of a mixture of a lysing agent and a reducing agent in a high-phosphate buffer solution. When red cells containing Hb S are present, the hemoglobin S will precipitate forming an opaque solution that is readily distinguished from the clear pink solution obtained with other hemoglobins.

False-positive results are rare and may be attributable to incomplete lysis of red cells because of erythrocytosis or to the presence of large amounts of protein in the patient's plasma, lending an opacity to the solution. The presence of many nucleated red blood cells in the peripheral blood can also cause a false-positive result. False-negative results, conversely, are not uncommon and depend mainly on the amount of the Hb S present. When less than 20% to 25% Hb S is present [as may occur in a transfused or exchange-transfused patient with sickle cell anemia, in an individual with both sickle cell trait (Hb AS) and two-deletion α thalassemia (α—/α—), in a severely anemic patient, and most important, in a neonate with sickle cell anemia or trait], the result is often negative. Doubling the quantity of blood tested alleviates this problem.

A positive solubility test indicates only the presence of Hb S or any hemoglobin containing the Hb S mutation (such as Hb C-Harlem, a doubly substituted hemoglobin variant) but does not provide a quantitative measure and does not therefore distinguish sickle cell trait, sickle cell anemia, or any combination of Hb S and another hemoglobinopathy.

Electrophoresis

Electrophoresis on various media is usually the first step in demonstrating and specifying a hemoglobin variant. Four commonly used techniques are outlined here: electrophoresis at alkaline pH (43,44), acid electrophoresis (45,46,47 and 48), isoelectric focusing (IEF) (49,50 and 51), and electrophoresis of separated globin chains (52,53).

Alkaline electrophoresis is a popular method used in the evaluation of hemoglobinopathies. It is also often called cellulose acetate electrophoresis, as this is the support medium used by many laboratories. In alkaline solution (tris-ethylenediaminetetraacetic acid (EDTA)-borate buffer at pH 8.6), all hemoglobin molecules have a net negative change and thus will migrate toward the anode when placed in an electric field. The amino acid substitution in many hemoglobin variants alters their net charge and thus their electrophoretic mobility (Fig. 42.5). Because substitution of either different amino acids or the same amino acid at different points in a globin chain may result in identical alterations in net change, electrophoresis at alkaline pH seldom provides a specific diagnosis. For example, although the substitution occurs at different positions, Hb C, Hb E, and Hb O-Arab all have a glutamic acid to lysine substitution, and thus they migrate together on alkaline electrophoresis. Similarly, many of the Hb D or Hb G variants co-migrate with Hb S or alkaline electrophoresis.

However, many types of hemoglobin that migrate together on alkaline electrophoresis can be separated utilizing electrophoresis at an acid pH (pH 6.0). The most commonly used method utilizes agar as the support medium and a citrate-citric acid buffer. The separation of variants depends not only on change differences but also on the ability of the variant to combine with a compound within the agar, termed agaropectin. This molecule combines with Hb variants at the surface of the molecule, most notably Hb C and Hb S. Most Hb variants with a substitution in the internal portion of the globin subunit do

not show an altered electrophoretic mobility from Hb A on this medium. Acid electrophoresis is most useful for differentiating Hb C from Hb E and Hb O-Arab and to confirm homozygous Hb S (i.e., that a single band on the S position or alkaline electrophoresis is composed only of Hb S and not Hb S and another Hb variant). The migration pattern of many common Hb variants on both alkaline and acid electrophoresis is shown in Fig. 42.6).

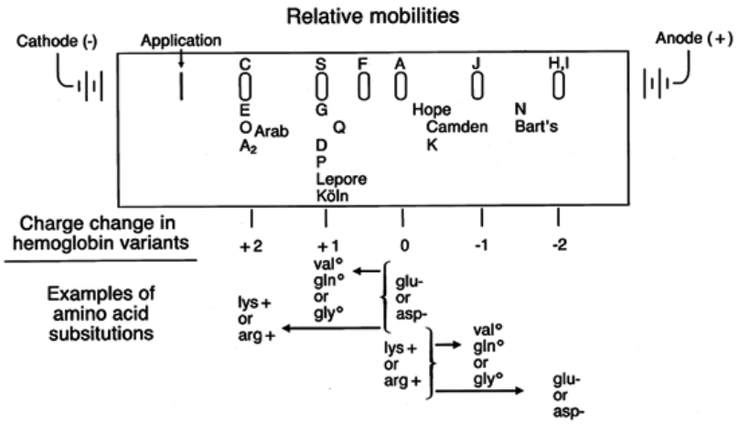


FIGURE 42.5. Hemoglobin electrophoresis at pH 8.6 (alkaline electrophoresis). At this pH, the hemoglobin molecule has an overall negative charge and migrates toward the anode. An amino acid substitution, which alters the overall charge of the molecule, will alter the electrophoretic mobility relative to Hb A.

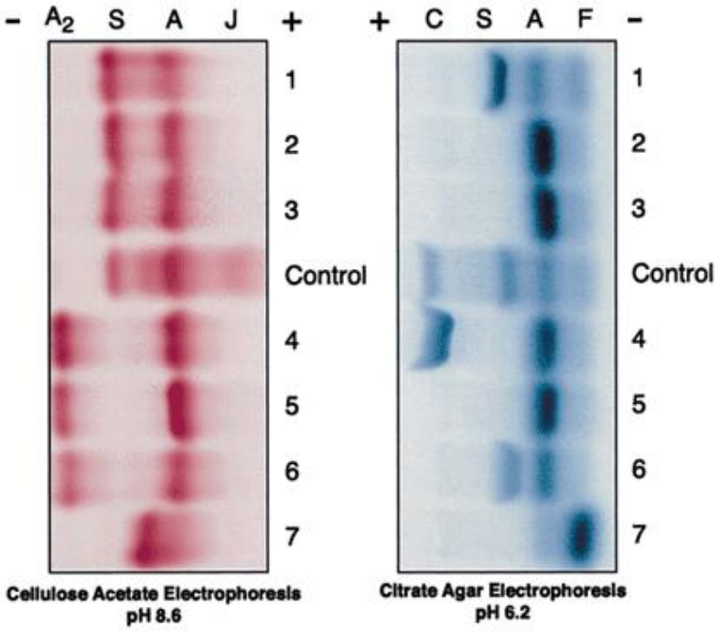


FIGURE 42.6. Examples of common hemoglobin variants on both cellulose acetate and citrate agar electrophoresis. 1, Hb S trait; 2, Hb G-Philadelphia trait; 3, Hb D-Punjab trait; 4, Hb C trait; 5, Hb E trait; 6, Hb O-Arab trait; 7, increased Hb F in a neonate.

IEF is another popular method used by laboratories that have a large number of specimens or very small sample volumes, such as laboratories that perform newborn screening. This electrophoretic method utilizes carrier ampholytes, small proteins that are able to carry both current and pH (Zwitterions). These compounds have MWs of 300 to 1,000 and are used in mixtures of 50 to 100 individual compounds. The ampholytes are incorporated into the support medium (either agarose or polyacrylamide). When a current is applied to the support medium, these ampholytes will gradually establish a pH gradient throughout the gel (for example, a pH range of 6 to 8 for hemoglobin analysis). High voltages must be used because the carrier ampholytes are present in high concentrations. Samples (RBC lysates) then placed on the gel will travel to their isoelectric point (the point at which they carry a net zero charge), where migration stops. Unlike alkaline electrophoresis, there is no danger of running the gel too long and running fast variants off the end of the gel. Because of minor differences in isoelectric points of various hemoglobin variants, IEF gives better separation of hemoglobin variants that show similar mobilities on alkaline electrophoresis. The bands present are much sharper than those seen on alkaline electrophoresis. Some hemoglobin variants, such as Hb-Malmö, show separation from Hb A, which is not seen on alkaline electrophoresis. Additionally, minor bands (such as Hb H, Hb-Barts, and δ-chain variants) are easily seen. However, minor bands (owing to glycosylated hemoglobins) and aging bands (methemoglobin,

glycerated hemoglobin) are also seen and may cause confusion in interpretation. Some typical patterns obtained with IEF are shown in Fig. 42.7 and Fig. 42.8.

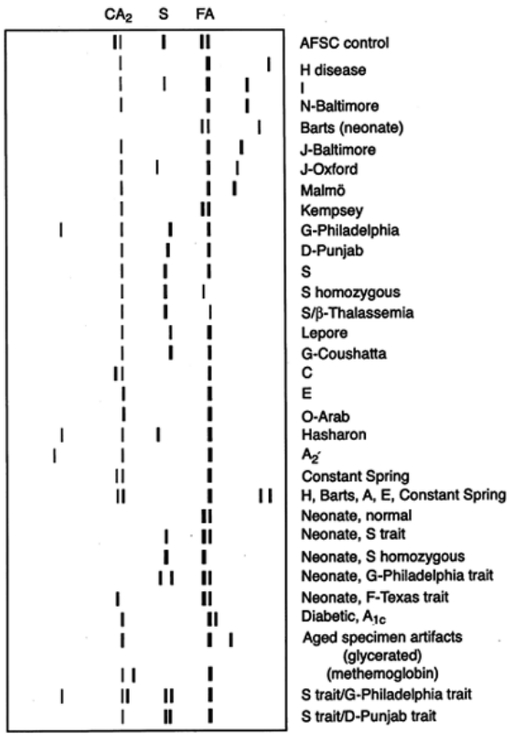


FIGURE 42.7. Isoelectric focusing. The electrophoretic mobilities of many hemoglobin variants are shown. (From Fairbanks VF, Klee G. Biochemical aspects of hematology. In: Burtis CA, Ashwood ER. *Tietz' textbook of clinical chemistry*, 2nd ed. Philadelphia: Saunders, 1994:2037, with permission.)

Globin-chain electrophoresis is an ancillary procedure in which a hemoglobin lysate with mercaptoethanol (to remove heme) and 8 mol/L urea (to dissociate the globin chains). When the resultant preparation is run on cellulose acetate, the α- and ß-globin chains are clearly separated. Furthermore, any hemoglobin variants present will often show altered mobility from the normal globin chains. Globin-chain electrophoresis is usually run at both an alkaline and acid pH, as some hemoglobin variants show slight differences in mobility. This method often gives additional information on hemoglobin variants that have similar mobilities by other methods. Examples of different variants are shown in Fig. 42.9 and Fig. 42.10. A summary of the electrophoretic mobilities of many hemoglobin variants is given in Fig. 42.11.

Although the references cited here and the instructions provided by manufacturers enable one to produce clear electrophoretic patterns, there are a few practical points that require emphasis. The clarity of separation depends to a significant degree on the specimen used and its handling. Blood collected in Na-EDTA as an anticoagulant is adequate for most purposes and may be stored for at least 24 hours at room temperature. This permits overnight mailing without special handling. Unstable hemoglobins are an exception because, if very unstable and therefore present in small amounts, they may disappear during such storage.

Given an adequate blood sample, there are several approaches to preparing it for electrophoresis. Because most plasma proteins such as albumin that are present in a significant quantity have an electrophoretic mobility much greater (faster) than any hemoglobin, the washing of red cells before preparing a lysate need not be tremendously thorough. In fact, if only a small quantity of blood is available, as in a capillary tube sample, rinsing can be omitted for screening purposes. Nevertheless, washing the red cells before preparing a lysate provides the best specimen.

Generally, application is best made with one of the several commercially available applicator wells, thus ensuring uniform and reproducible amounts. The quantity of lysate applied can be doubled when the amount of an abnormal hemoglobin under investigation is known to be small in cord blood specimens, for example, where the amount of Hb A and its variants are low compared with Hb F, or when attempting to demonstrate an abnormal Hb A_2 when an α-chain defect is suspected. The concentration and amount of lysate required for IEF are less than those for the other techniques.

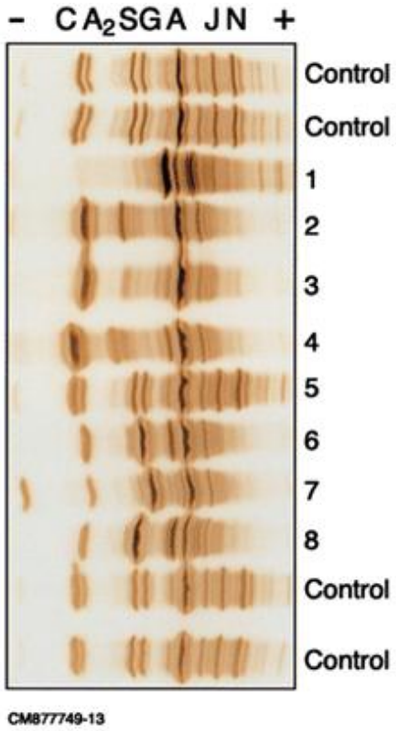


FIGURE 42.8. Isoelectric focusing gel with examples of common hemoglobin variants. 1, increased Hb F in a neonate. There is also Hb-Barts present, indicating alpha thalassemia; 2, Hb O-Arab trait; 3, Hb E trait; 4, Hb C trait; 5, control consisting of Hbs C-S-G, J, and N; 6, Hb D-Punjab trait; 7, Hb G-Philadelphia trait; 8, Hb S trait. Most specimens also show the aging bands owing to glycerated Hb and methemoglobin.

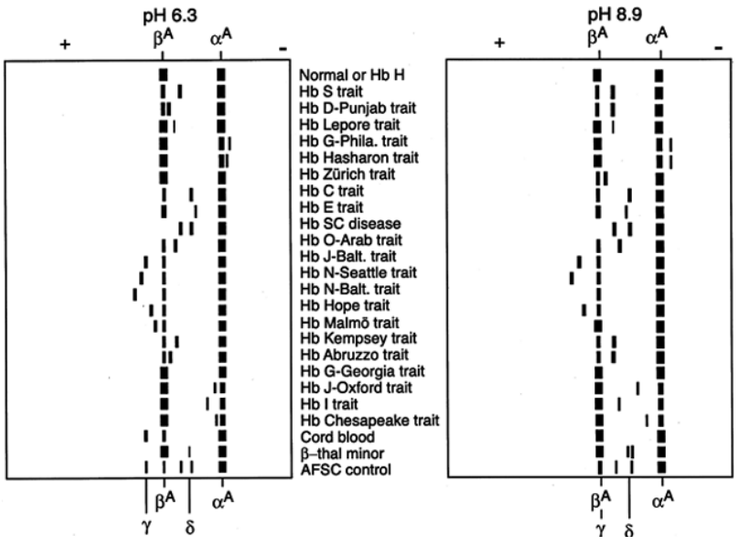


FIGURE 42.9. Globin chain electrophoresis. Examples of both α- and ß-chain variants are shown. (From Fairbanks VF, Klee G. Biochemical aspects of hematology. In: Burtis CA, Ashwood ER. *Tietz' textbook of clinical chemistry*, 2nd ed. Philadelphia: Saunders, 1994:2036, with permission.)

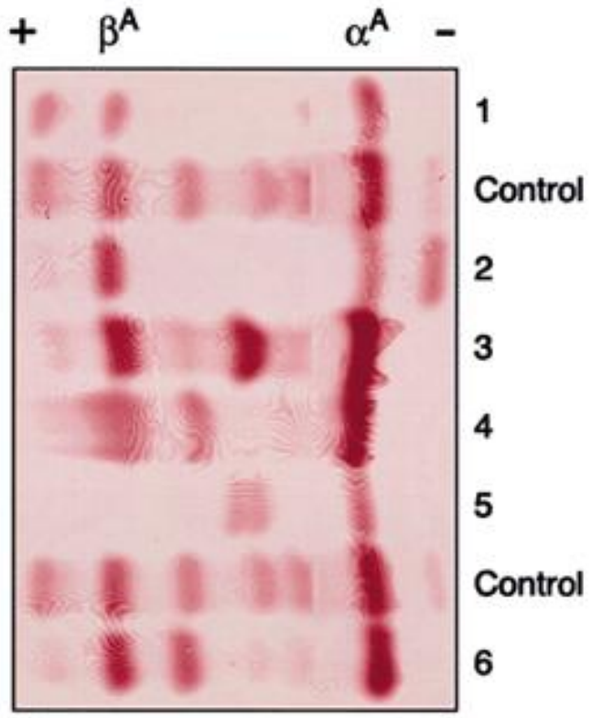


FIGURE 42.10. Globin-chain electrophoresis, pH 6.3. 1, Hb J-Baltimore trait; 2, Hb G-Philadelphia trait; 3, Hb E trait; 4, Hb S trait; 5, Hb C with Hb O-Arab (there are no normal β chains present); 6, Hb D-Punjab trait.

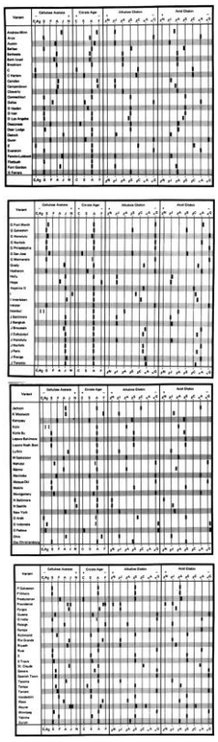


FIGURE 42.11. Hemoglobin electrophoresis, the relative migration of some hemoglobin variants on cellulose acetate at alkaline pH and in citrate agar at acid pH, and the relative migration of separated globin chains in alkaline and in acid buffers. (From Bunn HF, Forget BG. *Hemoglobin: molecular, genetic and clinical aspects*. Philadelphia: Saunders, 1986, with permission.)

It is helpful to keep the concentration constant because quantitative comparisons of the electrophoretic bands are often of value. In this context, the density of the carbonic anhydrase band is often useful because its concentration in the mature red cell is constant and therefore is usually proportional to the amount of lysate applied.

Quantitation

The quantitation of normal and variant hemoglobins is often of diagnostic importance. With a well-prepared electrophoretic pattern, “eyeballing” the plate is often as useful as more specific quantitation. Estimating the relative proportions of Hb S and Hb A can often be done in this manner. In those circumstances in which a more precise estimate is required, densitometry can be used on cellulose acetate plates. However, this method has disadvantages, particularly when comparing large with small quantities, as when measuring Hb A_2 as a percentage of a relatively large amount of Hb A; the smaller quantity is usually overestimated. Elution of the hemoglobin bands from the electrophoretic plates and direct measurement of their hemoglobin content provide even greater accuracy, but the method is time-consuming and seldom justified.

For some hemoglobins, specific methods of quantitation are readily available and simple to perform. Hb F can be quantitated base on its resistance to alkali (54); toluene may be used in place of carbon tetrachloride called for in this method. Hb A_2 can be measured by elution from an anion-exchange column (55,56), Hb-Barts and Hb H by elution from a cation-exchange column (57).

High-Performance Liquid Chromatography

High-performance liquid chromatography (HPLC) is a method that has been available for many years but not routinely used in the analysis of hemoglobinopathies. However, in the past few years, HPLC instruments have become available that are compact, user friendly, and dedicated to the detection of hemoglobins and their variants (58,59). Run lengths have been shortened from more than 20 minutes to 6 to 7 minutes. These instruments are approved by the U.S. Food and Drug Administration for the measurements of Hb S, A_2, and F but also give useful information for other hemoglobin variants that may be present (Fig 42.12). These instruments generally utilize a weak cation exchange column. Gradually increasing the ionic strength of the eluting solution causes the hemoglobin protein to come off the column at a particular retention time. Amino acid substitutions that are present in the hemoglobin variant will alter the

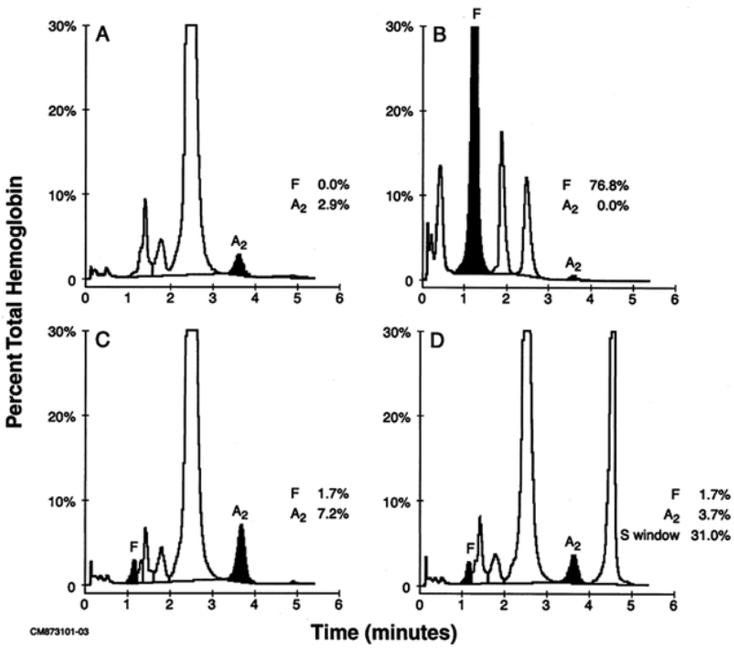


FIGURE 42.12. High-performance liquid chromatography. **A**, Normal adult; **B**, neonatal specimen with high Hb F (a hemoglobin variant is also present); **C**, ß-thalassemia trait with elevated Hb A_2; **D**, Hb S trait.

retention time relative to Hb A. There is some analogy between the retention times obtained by HPLC and the patterns seen on alkaline electrophoresis. Amino acid substitutions that give the molecule an overall more negative change at alkaline pH will run faster than Hb A. Similarly, these same substitutions usually result in a shorter retention time than Hb A on the HPLC column. Conversely, those amino acids substitutions that would result in a more positively changed molecule on alkaline electrophoresis will usually result in a longer retention time than Hb A on HPLC. This method also has the advantage that Hb C does not co-elute with Hb A_2, and so Hb A_2 can be measured in the presence of Hb A_2. Hb E and O-Arab, however, still co-elute with Hb A_2 by this method.

Globin-Chain Synthesis

All the methods described so far provide a static view of the hemoglobin content of red cells. Under some circumstances, particularly in the thalassemia syndromes, the relative rates of production of the various globin chains provide useful information. These globin chain-synthesis studies (60) require the separation and quantification of globin chains that have been generated in a reticulocyte-enriched preparation in the presence of an adequate supply of all necessary amino acids, one of which (usually leucine) is radiolabeled. The latter is then used as an indicator of the quantity of various globin chains produced. The individual globin chains are separated by high-pressure liquid chromatography. In many situations, such as in the diagnosis of the α-thalassemia syndromes, this methodology has been replaced by specific gene identification.

Molecular Methods

Molecular methods are available for the diagnosis of hemoglobinopathies and thalassemias but are generally restricted to use at major medical centers. Most α thalassemia mutations are attributable to deletions of one or both α chains; either Southern blot or polymerase chain reaction (PCR) methods can be utilized to detect most of these mutations. Molecular detection of ß-thalassemia mutations is more difficult because most mutations are nondeletional and there are a far greater number. In many geographic regions, however, relatively few mutations account for the majority of cases. Specific molecular assays can then be set up to look for only those specific mutations (25,61,62 and 63).

In very unusual cases, DNA sequencing of the three coding regions (exons) of the α- and ß-globin genes can be performed to definitely identify the mutation present in hemoglobin variants. This is particularly useful in confirming previously undiscovered hemoglobin variants or those hemoglobin variants that do not separate from Hb A on electrophoretic methods by HPLC (owing to neutral change substitutions). This process is accomplished by PCR-based assays using primers that flank the three exons and has been made much simpler with the advent of automated DNA sequencing machines. DNA sequencing can also be used to identify α-thalassemic mutations as well. However, these also require primers that will amplify the noncoding portions (introns) of ß-globin genes.

THALASSEMIA SYNDROMES

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

The thalassemia syndromes result from an imbalance in globin-chain production that is almost always owing to underproduction of one or two types of globin chains; rare examples in which overproduction of a globin chain causes the imbalance have been reported (64).

The thalassemia syndromes are classified according to the deficient globin chain or chains and according to their clinical severity (Table 42.3). The two largest groups are the α and ß thalassemias; δ and γ thalassemia is much less common and of little clinical significance. Therefore, they are discussed only briefly here.

δ Thalassemia

δ Thalassemia has been reported in the heterozygous and homozygous states (65). Hb A_2 ($\alpha_2\delta_2$) is reduced in the heterozygote and is usually absent in the homozygote because the genetic lesion results in failure to produce any δ-globin chains (δ^0); δ^+ thalassemia, in which there is a reduced production of δ chains, has also been reported (66). Neither the heterozygote nor the homozygote exhibits any clinical or hematologic abnormality. Co-inheritance of δ and ß thalassemia results in a form of ß thalassemia with a normal level of Hb A_2 (65). Reduction or absence of Hb A_2 is also seen in δß thalassemia, some forms of HPFH, and in association with Hb Lepore.

γ Thalassemia

Because a normal individual inherits two pairs of γ-globin genes, one pair coding for ^{G}γ-globin chain and the other for ^{A}γ-globin chains, which differ in possessing glycine or alanine as the 136th amino acid residue, several different forms and combinations of γ thalassemia might be expected. ^{G}γ Thalassemia appears to be the most common but even in the homozygote results only in slight reduction in the total quantity of Hb F in the fetus and newborn and in no clinical or hematologic abnormality in the adult (67). Homozygous loss of the activity of both ^{G}γ- and ^{A}γ-globin genes has not been reported and probably would be incompatible with life.

α Thalassemia

α Thalassemia results from a deficiency of α-globin chains and, except in very rare circumstances, is an inherited disorder (25,68). Most α-thalassemia syndromes are caused by a deletion of all or part of one or both α-globin genes on chromosome 16.

There are eight known deletions involving one of the α-globin genes. Two of these, which involve deletions of -3.7 kb and -4.2 kb, are very common worldwide and are found in multiple ethnic groups (69). These are probably the most common single gene disorders in the world. Single-gene deletion α thalassemia has an incidence of approximately 30% among African Americans (70), and the incidence is as high as 80% in some areas of Papua New Guinea (37). There are 21 known mutations that involve the deletion of both α-globin genes. The three most common are (a) the $--^{SEA}$ mutation found in southeast Asians, (b) the $--^{MED-1}$ deletion found in the Mediterranean population, (c) a 20.5-kb deletion that involves

all the α^2 and the 5′ end of the α^1 gene, also found in the Mediterranean population (69).

There are at least 32 known types of nondeletional α-thalassemia mutations; most are extremely rare. Fourteen mutations result in the synthesis of hyper-unstable hemoglobins, and five are owing to terminating codon mutations, such as Hb-Constant Spring. The remainder of the mutations affects some aspect of mRNA translation (69).

Because the normal complement of α-globin genes is four, there are five possible combinations of deletion (Table 42.5). Deletion of both genes from the same chromosome is called α thalassemia 1, whereas deletion of only one gene is known as α thalassemia 2; this apparently illogical nomenclature arose because the double deletion was described first. α Thalassemia 1 is particularly prevalent in southeast Asia and in the Mediterranean basin but is very rare among African Americans.

The severity of the clinical and hematologic effects is directly related to the number of genes deleted or affected (Table 42.5). As a consequence of the deficiency of α-globin chains, there is a relative excess of γ-globin chains in the fetus and infant and of ß chains in the adult; these excess chains can themselves form tetramers: γ_4 (Hb-Barts) and $ß_4$ (Hb H). These two hemoglobins are of little physiologic value because of their exceedingly high oxygen affinity. They are both unstable, a characteristic that underlies the hemolytic component of Hb H disease. On hemoglobin electrophoresis, they are fast migrators (Fig. 42.14) and can be quantitated by separation on cation-exchange columns (57). Various reticulocyte stains such as brilliant cresyl MA blue and new methylene blue cause Hb H to precipitate within the red cells, giving them a typical golf-ball appearance (41).

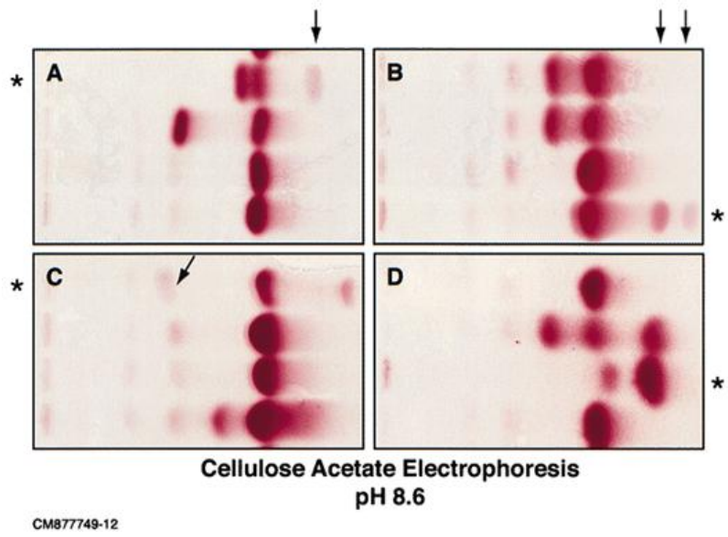


FIGURE 42.14. Cellulose acetate electrophoresis. **A:** Specimen from a neonate. Hb-Barts is present (*arrow*) indicating α thalassemia. **B:** Specimen from a 3-month-old. In addition to Hb-Barts, there is also a small amount of Hb H (*double arrow*). **C:** Hb H-Hb-Constant Spring. There is a fullness in the Hb A_2 area indicating the presence of Hb-Constant Spring. **D:** Hb-Barts hydrops fetalis. There is no Hb A present. The majority of the hemoglobin is Hb-Barts, with a small amount of Hb-Portland.

Heterozygous a Thalassemia 2

This form of α thalassemia is very common, but because it is associated with minimal imbalance between α- and δ- or ß-chain production, there is no clinical effect and the blood count, including the mean cell volume (MCV), is normal. In the adult, the diagnosis may be inferred from family studies or proven by gene studies; in the neonate, there is a slight increase (1% to 3%) in Hb Barts (71).

Heterozygous a Thalassemia 1 and Homozygous a Thalassemia 2

These genetic disorders result in α thalassemia minor (71) with the blood count typical of all forms of thalassemia minor (low MCV, minimal if any anemia). Because α-globin chains are present in all three normal hemoglobins, they are equally affected by a slight deficiency and there is no change in the relative percentages of Hb A, Hb A_2, and Hb F. It might be expected that the imbalance between α- and ß-globin chain production would result in the presence of a small amount of Hb H ($ß_4$) in the adult. This is usually not the case in the α thalassemia 2 homozygote. However, in the α thalassemia 1 heterozygote, Hb H can be precipitated in red cells by staining with brilliant cresyl blue, but it is not present in sufficient quantity to be detected by routine electrophoretic methods (72). Because α thalassemia 1 is exceedingly rare in African Americans but is quite common in southeast Asians, the cost-effectiveness of performing Hb H stains in individuals with thalassemia minor and normal levels of Hb A_2 and Hb F depends on the ethnic origin of the population under study.

In the neonate with either form of α thalassemia minor, an increase in Hb Barts (γ_4) in the range of 4% to 10% is usually found (71). This finding is a useful side benefit of neonatal screening programs for hemoglobinopathies because it provides positive evidence for a diagnosis using simple techniques and can be helpful information when an African-American child presents later in life with a microcytic hypochromic blood picture associated with normal levels of Hb A_2 and Hb F.

Hb H Disease

Hb H disease (clinically classified as a form of thalassemia intermedia) is a moderate to severe hemolytic disease (68,73). The

Hb H tetramer (β_4) is unstable and forms Heinz bodies (denatured hemoglobin) within the red cell. These inclusions are recognized by the reticuloendothelial system, and a portion of the red cell or the complete red cell may be removed from the circulation. The morphology of the red cells is often bizarre.

The clinical severity of Hb H disease is related to the amount of Hb H in the red cells, which in turn depends on the underlying genetic lesion. The most common type is the double heterozygote for α thalassemia 1 and α thalassemia 2 in which three α-globin genes are deleted (because α thalassemia 1 is rare among African Americans so also is Hb H disease). In southeast Asians, another frequent cause of Hb H disease is the combination of two deleted α-globin genes, a gene for the elongated α-globin chain of Hb-Constant Spring and one normal α-globin gene (--/$\alpha^{CS}\alpha$) (74,75). The mutant α-globin chain of Hb-Constant Spring is produced in small quantities (1% to 2%), and the α-globin chain deficiency therefore is nearly equivalent to that resulting from a deleted α gene (76).

The Hb-Constant Spring mutation changes the normal termination signal at codon 142 (TAA) to a codon that codes for glutamine (CAA). Thus, an additional 31 amino acids are added to the α-globin chain until the next termination codon is reached (75) (Fig. 42.13). This abnormal α globin is produced in very low amounts (1% to 2%) and so acts as an α-thalassemia mutation. The heterozygote for Hb-Constant Spring is clinically and hematologically normal. The homozygote for Hb-Constant Spring has mild normocytic hemolytic anemia with splenomegaly, not the typical thalassemia minor picture that might be expected (77,78). Hb-Constant Spring migrates between carbonic anhydrase and Hb A_2 on cellulose acetate at alkaline pH. Because the quantity is small, it may be overlooked unless specifically looked for (Fig. 42.14).

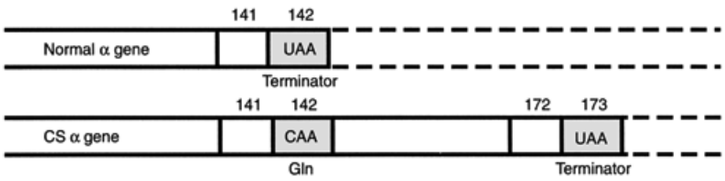


FIGURE 42.13. Substitution in mRNA of the abnormal α-globin chain of Hb-Constant Spring (CS).

In addition to the Hb-Constant Spring, there are four other Hb variants with elongated α chains. These differ only in the amino acid which is present at codon 142: Hb-Seal Rock (79) (GAG), Hb-Icaria (lysine) (80), Hb-Koya Dora (serine) (81),

TABLE 42.5. THE α-THALASSEMIAS RESULTING FROM GENE DELETION: NOMENCLATURE, CLINICAL EFFECT, AND PERCENTAGES OF Hb BARTS (γ_4) AND Hb H(β_4).

α-Genes	Subtype	Disorder	Hb Barts (Neonatel) (%)	Hb H (Adult) (%)
□-□[b]	Normal	Normal	<1	0
□-□				
▪-□	α-Thal 2	Silent α-thalassemia	1–3	0
□-□	heterozygote			
▪-□	α-Thal 2	Thalassemia minor	4–10	0
▪-□	homozygote			
▪-▪	α-Thal 1	Thalassemia minor	4–10	Trace
□-□	heterozygote			
▪-▪	α-Thal 1/α-Thal 2	Hb H disease	10–25	10–25
▪-□	heterozygote			
▪-▪	α-Thal 1	Hb Barts hydrops fetalis	75	—
▪-▪	heterozygote			

[a]The 25% of hemoglobin (Hb) unaccounted for in the α-thalassemia 1 homozygote consists mainly of the embryonic Hb Portland.

[b] ▪, Deleted α-globin gene; □, Normal α-globin gene.

Thal, thalassemia.

and Hb-Paske (tyrosine) (82). Similar to Hb H-Constant Spring, cases of Hb H disease in combination with Hb-Icaria (80,83) and Hb-Paske (82) have been described.

Homozygous a Thalassemia 1 (Hb-Barts Hydrops Fetalis)

Deletion of all four α-globin genes results in absence of α-globin chain production and is incompatible with life. Affected fetuses are hydropic, die *in utero* or soon after premature birth, and possess Hb-Barts (γ_4), 10% to 30% Hb-Portland [$(\zeta)_2\gamma_2$], and a small amount of Hb H (84,85) (Fig. 42.14). Hb-Portland is probably the main deliverer of oxygen because it has an oxygen dissociation curve not far from the physiologic one, compared with Hb-Barts, which has very high oxygen affinity (68). The blood smear shows large hypochromic red cells with many reticulocytes, nucleated red cells, target cells, and red cell fragments. Hb-Barts hydrops fetalis is rarely seen in populations outside southeast Asia and southern China, where it is the most common cause of hydrops fetalis.

ß-Thalassemia Syndromes

Any imbalance in globin-chain production involving a reduction or absence of ß-globin chains results in ß thalassemia. The simple concept of heterozygotes (ß thalassemia minor) and homozygotes (ß thalassemia major), although broadly applicable to these syndromes, is complicated by many possible genetic lesions. In contrast with α thalassemia, the mutations resulting in ß thalassemia are almost exclusively point mutations (86). Of the almost 200 known ß-thalassemia mutations, 179 are nondeletional mutations. In contrast, only 17 deletional forms of ß thalassemia have been described (87). These nondeletional mutations occur in or near the ß gene and may interfere with the initiation, transcription, termination, or splicing of mRNA; mutations within the gene may produce nonsense codes that stop transcription or splicing of mRNA or produce sense codes that result in a ß-globin chain so unstable that it is found only in red cell precursors (e.g., ß-Terre Haute) (88). The large number and variety of genetic causes are reflected in a spectrum of clinical syndromes in heterozygotes, homozygotes, and compound heterozygotes that range from silent ß thalassemia through thalassemia minor and intermedia to thalassemia major. For most practical purposes, it is possible to divide the genetic lesions into those resulting in the absence of ß-globin chain production ($ß^0$) and those resulting in reduction of ß-chain production ($ß^+$). Many of the ß-thalassemia alleles are clustered in particular populations so that, for instance, six alleles account for 92% of genes for ß thalassemia found in the Mediterranean area (86). This makes specific prenatal screening with gene probes less complicated than might be expected from the large total number of known alleles. These genetic lesions, affecting only the production of ß-globin chains, underlie the pure ß thalassemias.

Another group of lesions, mostly relatively large deletions, in addition to the ß gene, involves other regions of the ß-like gene cluster, including the γ and δ genes. These result in the δß and γδß thalassemias and HPFH.

ß Thalassemia Minor

ß Thalassemia minor is seldom symptomatic and requires no therapy (86). The uniformly microcytic red cells, typical of all types of thalassemia minor, are usually the first indication of its presence. Anemia, if present, is mild, and slight erythrocytosis is common. It is not possible to distinguish between $ß^0$ thalassemia minor and $ß^+$ thalassemia minor (or other forms of thalassemia minor) on a hematologic basis. However, as a group, individuals, usually black, with $ß^+$ (mild) thalassemia minor have less anemia and a higher MCV than those with $ß^0$ or $ß^+$ (severe) thalassemia minor (89). The three types are most easily distinguished when a sickle gene also segregates within the family; the S/$ß^0$-thalassemia compound heterozygote will have no Hb A, the S/$ß^+$-thalassemia compound heterozygote will have Hb A levels falling between 5% and 30%.

An approximate doubling of the level of Hb A_2 ($\alpha_2\delta_2$) is the hallmark of $ß^0$ and $ß^+$ thalassemia minor. Hb F may be normal and, when increased, is seldom greater than 5%.

Normal levels of Hb A_2 and Hb F are found occasionally in phenotypic examples of ß thalassemia minor. Several families have been reported in which one parent of an offspring with typical ß thalassemia major had phenotypic ß thalassemia minor with a normal Hb A_2 level. The presence of a separate gene for ß thalassemia, either in *cis* or in *trans*, is an explanation in at least some of these families (65). Hb A_2 may be reduced in iron deficiency states (90); hence, Hb A_2 may fall within the reference range when iron deficiency and ß thalassemia minor coexist.

ß Thalassemia Major

Almost all $ß^0$-thalassemia homozygotes, many $ß^+$ homozygotes, and compound heterozygotes present as thalassemia major. The deficiency of ß-globin chains is masked at birth by the production of γ-globin chains (Hb F), but as the switch from γ-globin chain production to ß-globin chain production proceeds, the severity of the clinical and hematologic expression increases. Diagnosis of the majority of cases that have not been anticipated by prenatal study of the parents or antenatal study of the fetus (86) is usually made during the second month of life. Where appropriate medical facilities are available, transfusion therapy aimed at maintaining a hemoglobin level above 10.0 or 12.0 g/dL should be instituted at the time of diagnosis, to be followed by chelation therapy to reduce iron overload (91).

If no therapy or inadequate therapy is given, a series of clinical and hematologic events ensue mainly as a result of hypoxia owing to the severe anemia (91). Anemia develops for two main reasons: dyserythropoiesis and hemolysis. The excess α-globin chains in the erythron do not form a soluble tetramer, α_4, comparable with the Hb-Barts (γ_4) and Hb H ($ß_4$) found in the α thalassemias. However, the excess α-globin chains damage the red cell precursors and mature red cells so that, despite the vast erythroid hyperplasia, red cell production is reduced and the red cells formed have a shortened life span. In untreated patients, the erythroid hyperplasia causes expansion of the marrow cavity with deformities of the facial, cranial, and long bones; extramedullary erythropoiesis causes splenomegaly and hepatomegaly; the persistent hypoxia retards mental and physical growth. In many

countries, this picture has been relegated to textbooks by optimal transfusion therapy.

In an untreated patient, the small amount of hemoglobin produced consists almost entirely of Hb F (90%) with little ($ß^+$ thalassemia major) or no Hb A ($ß^0$ thalassemia major). The distribution of Hb F among the red cells is not uniform, and those with the most Hb F have the longest life span, thus accentuating the percentage of Hb F in the circulation.

The bone marrow in untreated ß thalassemia major shows vast erythroid hyperplasia, megaloblastic, dysplastic, and bizarre precursors with many degenerate forms that reflect the intramedullary destruction of the α-globin chain–laden precursors. Although the formation of a relatively soluble tetramer $α_4$ comparable with Hb-Barts ($γ_4$) and Hb H ($ß_4$) found in α thalassemia does not occur, aggregates of ß-globin chains can be demonstrated in red cell precursors and in mature red cells (particularly after splenectomy) by supravital staining (92). The morphology may suggest erythroleukemia (FAB M6) even to the extent of exhibiting periodic acid–Schiff-positive inclusions. Iron is present in large amounts, and ringed sideroblasts may be present in moderate numbers. Anemia is severe, MCV and mean corpuscular hemoglobin (MCH) are low, but the mean corpuscular hemoglobin concentration (MCHC) may be only slightly reduced despite the apparent pallor of the red cells in a peripheral blood smear. There is marked variation in red cell size and shape. Reticulocytosis and polychromasia are present, but because of the ineffective erythropoiesis, not in the quantity expected for the degree of anemia; erythroid precursors at various stages of maturation are also found in the peripheral blood.

There is biochemical evidence of hemolysis, including increased unconjugated bilirubin levels. Serum iron and transferrin are increased even in the untransfused individual. The urine may appear brown owing to the presence of dipyrroles.

The pattern of hemoglobin electrophoresis depends on the type of ß thalassemia major. Hb A will be absent in untransfused homozygous $ß^0$ thalassemia; a variable amount will be seen in homozygous $ß^+$ thalassemia. Homozygous $ß^+$ thalassemia with relatively large amounts of Hb A, clinically classified as thalassemia intermedia, is a particularly common type among Africans and African Americans. In $ß^0$-thalassemia major Hb F represents approximately 90% of the hemoglobin present; however, this is 90% of very little. Hb A_2 may be low, normal, or, less often, increased in percentage but is always proportionally high compared with Hb A. Unexpectedly low levels of Hb A_2 may be accounted for by the selective survival of red cells with the greatest amount of Hb F, which also contain the lowest amounts of Hb A_2 (93).

δß-Thalassemia

The absent or reduced production of δ- and ß-globin chains that occurs in the δß type of thalassemia is usually the result of deletion of all or part of both the δ- and ß-globin genes (94,95). Ten different δß-thalassemia mutations have been described (96). Because two chains are underproduced, it might be assumed that the resulting thalassemias would be more severe than their ß-thalassemia counterparts. This is not the case because the deletion is associated with a significant persistent production of the γ chains of Hb F, thus diminishing the α:non-α-chain imbalance. The δß thalassemias form part of the spectrum that lies between ß thalassemia and HPFH.

Heterozygous δß thalassemia is often hematologically indistinguishable from other types of thalassemia minor, although the red cell indices may be closer to normal. In a few instances, the indices may be normal. These types can be classified as a silent form of thalassemia. δß Thalassemia minor is readily distinguished from other forms of thalassemia minor by the presence of 5% to 20% Hb F, which is heterogeneously distributed among the red cells, and a normal or reduced level of Hb A_2.

Homozygous δß thalassemia is a mild disorder, clinically classifiable as a thalassemia intermedia or even minor (97). Anemia is mild (8.0 to 12.0 g/dL) with microcytic hypochromic red cells and evidence of reticulocytosis. Hemoglobin electrophoresis shows 100% Hb F with no Hb A or Hb A_2.

HEREDITARY PERSISTENCE OF FETAL HEMOGLOBIN

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

HPFH comprises a group of disorders, clinically benign and usually hematologically normal, in which Hb F production persists throughout life. Several distinct genetic lesions underlie the various types of HPFH (95,98,99). In seven types, there is a large deletion involving both the δ and ß genes, which in at least one instance may bring an enhancer at the 3+ end of the deletion into close apposition with the α genes, probably accounting for their continued expression. A very similar deletion results in δß thalassemia; however, the DNA enhancer sequence is included in the deletion, perhaps accounting for the lower level of γ-globin chain (and Hb F) production and the imbalance in α:non-α-globin chain production typical of the disorder. The other types of HPFH are nondeletional and associated with single-base substitutions that occur in or near promoter sequences of the $^Aγ$- and $^Gγ$-globin genes (99). There is even evidence that some types may be attributable to genetic lesions unconnected with the γδß-gene complex (98).

The amount of Hb F varies from one type of HPFH to another (Table 42.6) as does its distribution among the red cells. HPFH may be subdivided into those in which all red cells contain Hb F (pancellular types) and those in which only a portion of the red cells contain Hb F (heterocellular types). The distinction is made based on the acid-elution technique of Kleihauer et al. (100,101) (Hb F is acid resistant as well as alkali resistant) or by the use of anti-Hb F antibodies (102,103). The latter is more sensitive; however, it may be difficult with either technique to be sure of the distribution of Hb F, particularly when at a low level. The discovery of several types of HPFH with a heterocellular distribution of Hb F reduces the value of demonstrating the type of distribution of Hb F among red cells because a heterocellular distribution is also found in all other forms of increased Hb F in adults (29).

TABLE 42.6. PERCENTAGE OF Hb F, RATIO OF γ-GLOBIN CHAINS, AND DISTRIBUTION AMONG RED CELLS IN FOUR TYPES OF HEREDITARY PERSISTENCE OF FETAL HEMOGLOBIN.

	Hb F %			
Type	**Heterozygote**	**Homozygote**	**$^Gγ^Aγ$**	**Distribution Among Red Cells**
Black	17–36	100	0.4	Pancellular
Greek	10–30	—	0.1	Pancellular
British	4–12	19–21	0.08	Heterocellular
Swiss	1–5	—	Variable	Heterocellular

Hb, hemoglobin.

Hemoglobin Lepore

Hb Lepore in the heterozygote is associated with a typical thalassemia minor blood count and in the homozygote with thalassemia

intermedia or even major (104). This unusual hemoglobin is composed of a pair of normal α-globin chains and a pair of non-α-globin chains that consist of the beginning (N-terminal end) of a δ-globin chain and the end (C-terminal) of a ß-globin chain. This globin chain is the product of a fusion gene that has arisen from unequal chromosomal crossover (Fig. 42.15). The exact point of crossover and therefore the proportion of δ and ß gene present in the Lepore gene varies. Three different hemoglobins have been reported: Hb $Lepore_{Hollandia}$, Hb $Lepore_{Baltimore}$, and Hb $Lepore_{Boston}$, the last being the most common (104). As a result of the crossover, the δ and ß genes are represented only by the fusion gene that in the adult heterozygote would be expected to produce approximately 50% of the non-α-globin chains; however, Hb Lepore makes up only approximately 10% of the circulating hemoglobin. This deficient production results in an imbalance between α- and non-α-globin chain production, causing thalassemia minor in the heterozygote with its typical hematologic picture. Homozygous Hb Lepore (105) presents as thalassemia major or thalassemia intermedia. The circulating hemoglobin consists of approximately 75% Hb F and 25% Hb Lepore with no Hb A or Hb A_2.

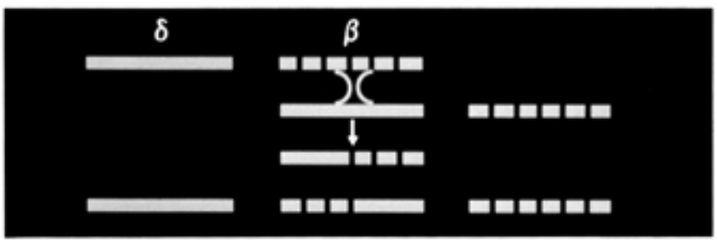


FIGURE 42.15. Unequal crossover in chromosome 11 between the δ and ß genes, resulting in the production of the δß fusion gene of the non-α-globin chain of Hb Lepore with no normal δ- and ß-globin genes in *cis* and the ßdelta; fusion gene of Hb anti-Lepore with normal δ- and ß-globin genes in *cis*.

Hb Lepore migrates with Hb S on cellulose acetate in alkaline buffer and with Hb A in acid citrate agar. By HPLC, Hb Lepore migrates with Hb A_2. The finding of approximately 10% of a hemoglobin migrating like Hb S on cellulose acetate and in the A_2 position on HPLC in an untransfused patient is virtually diagnostic of Hb Lepore.

The converse of Hb Lepore is Hb anti-Lepore (Fig. 42.15) (105). In this situation, δ and v genes in *cis* are intact with the addition of an anti-Lepore gene that is responsible for the production of a small amount of a non-α chain made up of the beginning of a ß- and the end of a δ-globin chain. The amount of Hb Anti-Lepore produced is small, and the imbalance of globin-chain production is minimal. The blood count and all red cell indices are normal. Three different anti-Lepore hemoglobins have been reported; as in the case of Hb Lepore, they differ in the point at which chromosomal crossover has occurred.

Differential Diagnosis of Thalassemia Minor

Thalassemia minor is a benign syndrome with few if any symptoms but with an abnormal hematologic picture that is often described as a microcytic hypochromic anemia (Table 42.7). The anemia is slight or nonexistent in Africans but usually present in Asians, and the hypochromasia, as measured by the MCHC, is seldom more than slight. The striking features of the blood count are the low MCV, the high red cell count, and the lack of anisocytosis as measured by the red cell distribution widths (RDW). Iron-deficiency anemia and iron-deficient erythrocytosis are the main differential diagnoses. Morphologically, the red cells of thalassemia minor differ little from those of mild iron deficiency; the presence of target cells is not a useful discriminator; however, basophilic stippling occurs in thalassemia minor but is seldom seen in iron deficiency. Various regression formulas have been proposed to distinguish the blood counts of iron deficiency and thalassemia minor, indicating whether the patient's iron status or hemoglobin should be studied first. England and Fraser (106) suggested that thalassemia is the probable diagnosis if the result of the formula MCV - RBC - (5 × Hb) × 3.4 is negative. Green and King (107) suggested the formula (MCV^2 × RDW) ÷ (Hb × 100), with a result less than 73 indicating thalassemia minor.

TABLE 42.7. BLOOD COUNTS OF 10 PATIENTS WITH ß-THALASSEMIA MINOR[a]

Patient	Hb (g/L) 120–160[b]	RBC (10^{12}/L) 4.5–5.5[b]	MCV (fl) 80–100[b]	RDW <15	A'_2 (%) 1.5–3.5[b]
1	124	5.7	64	13.3	6.2
2	132	7.2	65	12.5	5.2
3	127	6.2	65	12.7	6.5
4	125	6.0	64	14.4	6.3
5	140	7.2	58	14.2	6.8
6	113	6.1	58	13.6	7.3
7	120	6.0	63	14.7	6.2
8	140	6.0	61	12.8	6.5
9	126	6.0	69	15.2	6.3
10	125	5.5	71	14.4	6.6

[a] The counts, with minimal or no anemia, increased red cell count (RBC), low mean cell volume (MCV), and normal red cell distribution width (RDW) might be found in any type of thalassemia minor. The approximate doubling of the Hb A'_2 percentage confirms the specific diagnosis of ß-thalassemia minor.

[b] Reference range.

Hb, hemoglobin.

Whichever pragmatic approach is adopted, three factors need to be kept in mind. First, because iron deficiency and thalassemia minor are common disorders, coincidence of the two is not uncommon. Second, the Hb A_2 level is reduced in iron deficiency, and occasionally the Hb A_2 level in an individual with ß thalassemia minor and iron deficiency may fall within the normal range (90). Third, because thalassemia minor may be caused by a structurally abnormal hemoglobin (e.g., Hb Lepore) or may coexist with an abnormal hemoglobin (e.g., Hb S α thalassemia), the complete study of a patient with the blood picture of thalassemia minor should include hemoglobin electrophoresis. In addition to indicating the presence of a hemoglobin variant, hemoglobin electrophoresis may provide other useful diagnostic information. It can be argued that with increased Hb A_2, indicated by the presence of an Hb A_2 band significantly denser than

the carbonic anhydrase band, with or without an increased Hb F band in a patient with a thalassemic blood picture, is strong enough evidence for the diagnosis of β thalassemia minor without specific quantification of Hb A_2. Similarly, the combination of an electrophoretic band suggesting 10% to 15% Hb F and a thalassemic blood count strongly suggests δβ thalassemia minor. α Thalassemia minor is often a diagnosis of exclusion; a thalassemic blood count with normal levels of Hb A_2 and Hb F and normal iron studies is consistent with such a diagnosis. Specific evidence for the presence of one or more defective (thalassemic) α-globin genes may come from the presence of increased quantities of Hb-Barts at birth (Table 42.5), the presence of rare Hb H inclusions in the αα/– – type (heterozygous α thalassemia 1) but not in the α—/α— (homozygous α thalassemia 2), and by the use of α-gene probes.

STRUCTURAL HEMOGLOBIN VARIANTS

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

More than 750 structural hemoglobin variants have been reported, most of which are rare and of no clinical or hematologic significance. An updated compendium of all reported examples is published regularly (13). A few generalizations can be made. Except in the case of two of the most common variants, Hb S and Hb C, structural alterations affecting the surface of the hemoglobin molecule seldom have any deleterious effect. Conversely, almost all variants in which the structural defect lies near the heme pocket, lies at points of contact between globin chains, or affects their helical structure are associated with alterations in the stability or function of the molecule.

The various types of genetic lesions that underlie the structurally abnormal hemoglobins include:

- Single-base substitutions within a DNA triplet codon that result in an amino acid substitution (e.g., Hb S, $\alpha_2\beta_2^{6\ Glu\rightarrow Val\ A\rightarrow T}$, GAG to GTG);
- Two separate base substitutions probably arising from chromosomal crossover involving two abnormal genes (e.g., Hb C-Harlem, $\alpha_2\beta_2^{6\ Glu\rightarrow Val}$, 73 Asp→Asn, a combination of the substitutions found singly in Hb S and Hb-Korle Bu) (108);
- Deletion of one or more amino acid residues (e.g., Hb-Leiden, $\alpha_2\beta_2^{6\ or\ 7\ Glu\rightarrow 0}$) (109,110);
- Shortening of globin chain by premature termination (e.g., Hb-McKees Rocks) (111);
- Elongation of globin chain by insertion (e.g., Hb-Grady, $\alpha_2^{116+118\ duplicated}\beta_2$ (112);
- Elongation by mutation in chain terminator (e.g., Hb-Constant Spring $\alpha_2^{171}\beta_2$ (75); Hb Tak (113);
- Elongation by frame-shift mutation (e.g., Hb-Wayne) (114);
- Fusion of globin chains (e.g., Hb Lepore $\alpha_2\delta\beta_2$ (104).

The Sickling Disorders

Hb S ($\alpha_2\beta_2^{6\ glu\rightarrow val}$) results from a point mutation in the sixth codon of the β chain, so that GAG is replaced by valine (GTG). Geographically, Hb S is found in highest concentration in Africa. In west central Africa, it is found in highest concentration the modern countries of Nigeria and Ghana as well as in Gabon and Zaire. On the eastern coast, it is found in highest concentration in Kenya with a decreasing percentage proceeding southward. Hb S is also found in the northeastern corner of Saudi Arabia, a localized area of east central India, and in low concentration (less than 1%) in areas of Greece, Sicily, and Turkey. Many, but not all, of these areas correspond to endemic areas of falciparum malaria. In American blacks, the prevalence of Hb S trait is 8.0% and that of sickle cell anemia is 0.14% (115). There is evidence from studies of DNA segments (haplotype analysis) near the β-globin gene locus (using restriction length polymorphisms) that the mutation probably occurred in several different population groups and then reached polymorphic levels through its beneficial effect in areas of endemic malaria (34,116). There are slight differences between these groups in the severity of sickle cell anemia.

The pathophysiologic basis of the various sickling disorders is the tendency for molecules of Hb S to precipitate within the red cell in the form of liquid crystals or tactoids under conditions of reduced oxygen tension (Fig. 42.16). The characteristic sickle cell is the result. This tendency to precipitate depends on the concentration of Hb S within the red cell, the nature of any other

TABLE 42.8. SOME COMBINATIONS OF Hb S AND α-THALASSEMIA, β-THALASSEMIA, OR GLOBIN CHAIN STRUCTURAL ABNORMALITIES

	Percentage in Adults						
Hemoglobinopathy	**Hb A**	**Hb S**	**Hb X**	**Hb A_2**	**Hb F**	**Sickle Cells**	**Clinical Severity**
Sickle cell trait (Hb AS)	55–60	40–45	0	N	N	0	1
Sickle cell anemia (Hb SS)	0	90–95	0	N	5–10	+	3
Hb S α-thalassemia 1	75	25	0	N	N	0	1
Hb S β^0 thalassemia	0	90–95	0	↑	5–10	+	3
Hb S β^+ thalassemia	5–30	60–90	0	↑	5–10	+	2
Hb S HPFH	0	70–80	0	N	20–30	0	1
Hb SC	0	50	50	N	N	±	2
Hb SD-Los Angeles	0	50	50	N	N	±	2
Hb SN	0	50	50	N	N	0	1
Hb SO	0	50	50	N	N	+	3
Hb AS-G-Philadelphia	25	25	50[a]	N	N	0	1

[a] Approximately 25% Hb G (migrating with Hb Sl and 25% of the hybrid $\alpha^G_2\beta^S_2$ (migrating as Hb C).

N, normal; Hb X, abnormal hemoglobin other than Hb S; 0, sickle cells absent; +, sickle cells present; ±, sickle cells possible; 1, benign; 2, moderately severe; 3, similar to sickle cell anemia.

hemoglobin present, as well as the oxygen tension. Thus, at the normal PO_2 of postcapillary venules, sickle cells will form in a homozygote (Hb SS, sickle cell anemia) whose red cells contain approximately 90% Hb S, whereas they will not form in a heterozygote (Hb AS, sickle cell trait) whose red cells contain approximately 55% Hb A and only 40% Hb S. The severity of a sickling disorder is directly related to the tendency of red cells to sickle (Table 42.8).

FIGURE 42.16. Electron micrograph of a portion of a sickled red cell in a renal biopsy. The tactoids of Hb S molecules are seen arranged in tubular fashion cut horizontally, vertically, and tangentially. (From Hoffman GC. The sickling disorders. *Lab Med* 1990;21:797–807, with permission.)

In Hb SS, a continuous cycle (Fig. 42.17) of peripheral deoxygenation and sickling and central oxygenation and desickling occurs as red cells pass from periphery through the lungs. Two events may alter this cycle. Repeated deformation of the red cell membranes eventually leads to the formation of irreversibly sickled red cells. Sickled red cells, unsuited to pass through small blood vessels, may block some of those vessels, causing local stasis, further deoxygenation of the red cells, and further sickling; the temporary or permanent loss of blood flow to a distal tissue segment causes the pain typical of the vasoocclusive sickle cell crisis (117).

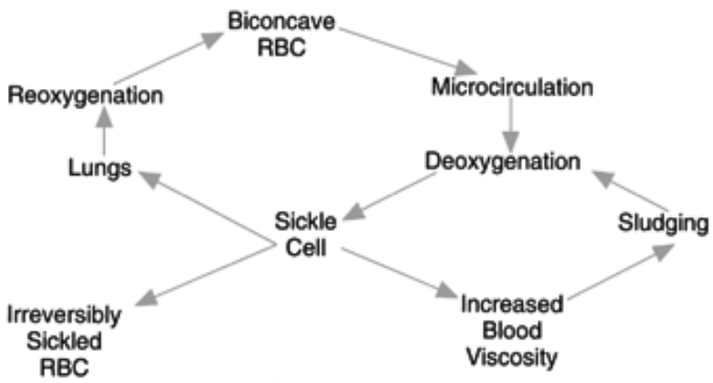


FIGURE 42.17. The sickling cycle that occurs in patients with sickle cell anemia. Vasoocclusive (painful) crises may develop if "sludging" leads to blockage of blood vessels. (From Fishleder AJ, Hoffman GC. A practical approach to the detection of hemoglobinopathies: part II. The sickle cell disorders. *Lab Med* 1987;18:441–443, with permission.)

Sickle Cell Trait

Sickle cell trait (Hb AS) is an essentially benign disorder because, under normal physiologic conditions, sickle cells are not formed (115). There is, however, some evidence that in rare hypoxic situations, sickle cells may form *in vivo*; sudden death among individuals involved in strenuous exercise at high altitudes is more common among African Americans than among whites and may be associated with *in vivo* sickling (118). Localized hypoxia in renal papillae is the probable explanation for the instances of painless hematuria, usually from the left kidney, which may occur rarely in sickle cell trait but more commonly in sickle cell anemia and Hb SC disease. Isosthenuria may also occur.

The blood count is normal; hemoglobin electrophoresis shows approximately 40% Hb S, 55% Hb A, with a normal or slightly increased amount of Hb A_2 (Fig. 42.18), and the solubility test gives a positive result.

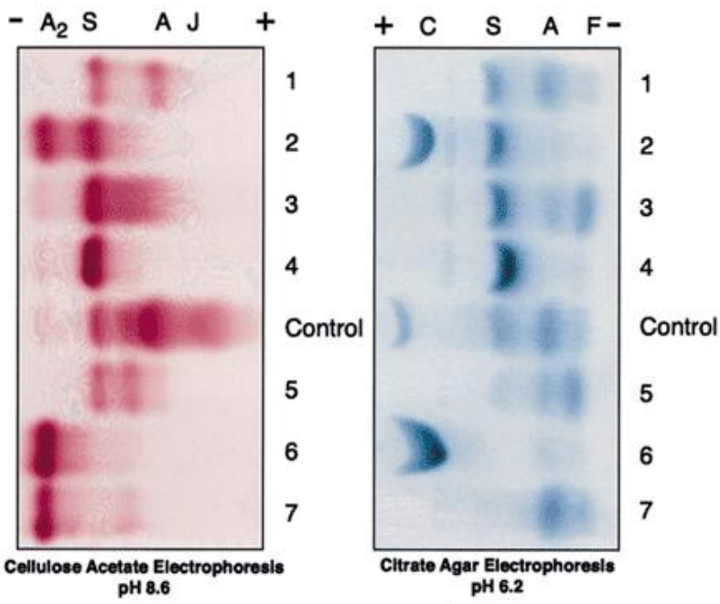


FIGURE 42.18. Examples of many conditions discussed in the text on both cellulose acetate and citrate agar electrophoresis. 1, Hb S trait; 2, Hb S/C disease; 3, Hb S/β+thalassemia. The small amount of Hb A is better seen on citrate agar electrophoresis. Hb F is also elevated. 4, Homozygous Hb S; 5, Hb S/Hb D Punjab in an infant. Hb F is elevated. 6, Homozygous Hb C; 7, homozygous Hb E.

Sickle Cell Anemia

The varied clinical manifestations of sickle cell anemia are the result of the altered hematologic properties of the abnormal sickle erythrocyte. There is a chronic hemolytic anemia owing to shortened red cell survival. This hemolytic process is often associated with gallstone formation. There are various acute vasoocclusive crises that are caused by obstruction of the microcirculation by the sickled cells. These include painful crises (musculoskeletal, abdominal), the acute chest syndrome, splenic sequestration crises, neurologic crises (seizures, strokes), and the hand-foot syndrome (in children). Hematologic crises include aplastic crises (often secondary to parvovirus infection) and megaloblastic crises (secondary to folate deficiency). There is chronic end organ damage that is the sequela of both multiple vasoocclusive events and chronic anemia. These include cardiac (myocardial hypertrophy), pulmonary (multiple infarcts), genitourinary (concentrating defects, nephrotic syndrome, renal failure, priapism), skeletal (bone infarctions, aseptic necrosis of the femoral/humeral head), neurologic (deficits secondary to strokes), ocular (proliferative retinopathy, glaucoma) and dermatologic (lower leg ulcers). One complication of particular note is the phenomenon of autosplenectomy. Owing to repeated infarcts, the spleen by adulthood is not more than a fibrotic nub. Because of this, these patients have increased susceptibility to infections, particularly by encapsulated organisms such as *Pneumococcus*. There are also general systemic manifestations such as impairment of growth and sexual development (119,120).

The severity of the disease, as reflected by the degree of anemia and frequency of the painful crises, varies widely from individual to individual and between ethnic groups. The disease is usually mild among Saudi Arabians; there is little variation among the genetically distinct types of sickle cell anemia found in Africa (121). These variations depend in part on the level of Hb F, which tends to inhibit gelation, and on the number of α

genes; the coexistence of one or two α-gene deletions reduces the MCHC and the tendency for Hb S to gel. A third major factor in the morbidity and mortality of sickle cell anemia is infection, usually bacterial, that is at least in part a reflection of the functional asplenia of infancy and childhood and splenic infarction in the adult. During fetal and early neonatal life, the presence of Hb F reduces the sickling tendency. Accordingly, signs or symptoms of the disease do not develop until the third or fourth month of life. The first vasoocclusive events usually do not appear until 6 to 12 months. In milder forms of the disease, they may not appear for several years. Rare asymptomatic cases occur.

Early childhood is a particularly dangerous time for patients with sickle cell anemia, partly because of their inability to describe symptoms but mainly because of potentially lethal complications that may occur in addition to vasoocclusive episodes. Overwhelming infection, particularly pneumococcal, may prove lethal in a matter of hours, and acute splenic sequestration of red cells that may be equivalent to loss of one third of the blood volume is a medical emergency. The importance of prenatal or neonatal diagnosis cannot be overstated; it allows the parent and pediatrician to take prompt action when indicated and to provide prophylaxis against infection. Vasoocclusive episodes or painful crises may occur at any age. They occur sporadically and vary widely in frequency from patient to patient (122). A precipitating event is often not obvious; however, infection, bacterial or viral, precedes approximately one third of crises, probably as result of pyrexia (123). Any organ may be affected. The spleen, enlarged during the first years of life, is seldom palpable after the age of 7 years owing to repeated infarctions.

The degree of anemia varies widely (average hematocrit, 24.6 ± 3.8) (124); the hemoglobin level may fall precipitously during the temporary “aplastic” crises that are associated with parvovirus infection (125) as well as during sequestration crises; hyperhemolytic episodes are seldom if ever the cause. Several long term approaches to therapy of sickle cell anemia, in addition to supportive measures during crises, have been suggested. The roles of red cell transfusion, exchange transfusion, and hypertransfusion protocols such as those used in the treatment of β thalassemia major remain a matter of controversy, even when considered during times of particular risk such as major surgery and pregnancy (126,127,128 and 129). Hydroxyurea can increase the production of Hb F and thus reduce the tendency for the Hb S to form tactoids (130). Replacing the red cell factory either *in toto* by bone marrow transplantation (131) or in part by specific gene replacement are possible future alternatives.

The main laboratory features of sickle cell anemia include variable but usually moderate to severe normocytic anemia, primarily hemolytic, in types associated with sickle cells and reticulocytosis; the presence of Howell-Jolly bodies and target cells, indicating hyposplenism or splenic infarction; 90% to 95% Hb S with 5% to 10% Hb F and no Hb A on hemoglobin electrophoresis (Fig. 42.18); and a positive solubility test.

Sickling Disorders Involving a Second Hemoglobinopathy

The more common combinations of Hb S with other hemoglobinopathies are listed in Table 42.8. Inheritance of a β^S gene from one parent and a second β-gene variant from the other results in a compound heterozygote such as Hb SC disease or Hb S β^0 thalassemia. An α-gene variant can be inherited by an individual who has either sickle cell trait or sickle cell anemia, for instance, Hb AS/G-Philadelphia or Hb SS/G-Philadelphia.

The combination of hemoglobin S and C (Hb SC disease) produces a sickling syndrome with clinical manifestations that are milder than that of sickle cell anemia. The prevalence is 0.13% in African Americans. The hemolytic anemia is mild to moderate and complications from the hemolytic anemia are less frequent. Growth, development, sexual maturation, and body habitus are usually normal; life expectancy is only slightly shortened. Painful vasoocclusive crises (abdominal, musculoskeletal, neurologic) are less frequent and infarctive damage (e.g., bone, lung) are usually less disabling. Approximately half of patients have splenomegaly. Although asplenia occurs in only 25% of patients, painful splenic infarcts are common, and episodes of acute splenic sequestration may also occur. Because the function of the spleen is usually intact, the frequency of pneumococcal infections (meningitis sepsis) is not increased. The frequency of cerebrovascular accidents is approximately equal to that of sickle cell anemia patients. Interestingly, renal papillary necrosis and proliferative retinopathy are more frequent in Hb SC disease. Aseptic necrosis of the femoral head is approximately half as frequent as in sickle cell anemia; however, aseptic necrosis of the humeral head is more common. Although there are fewer complications at time of delivery, pregnant women with Hb SC disease have a higher rate of spontaneous early abortions (132,133 and 134).

In the peripheral blood smear classic sickle cells are rare; more often one sees irregular, boat-shaped cells. Characteristic SC poikilocytes are normally seen. These are dense, very misshapen cells that often contain Hb C crystal. Target cells and spherocytes can also be seen (135). Hemoglobin electrophoresis at alkaline pH shows approximately equal amounts of Hb S and Hb C (Hb A_2 is contained within the latter). Acid electrophoresis is necessary for confirmation to distinguish Hb SC from Hb S/O Arab and Hb S/E (Fig. 42.18).

Hb D-Los Angeles and Hb O-Arab also copolymerize with Hb S and therefore produce sickling disorders (136,137). Acid electrophoresis is needed to distinguish Hb S/D-Los Angeles from homozygous Hb S as these two variants show identical mobilities on alkaline electrophoresis but are different on acid electrophoresis. Acid electrophoresis will differentiate Hb S/O-Arab from Hb S/C (Fig. 42.18).

In contrast, Hb S in combination with HPFH (Hb S/HPFH) is a benign disorder. Patients do not usually have anemia, and there are no vasoocclusive events. Hemoglobin electrophoresis in the compound heterozygote usually shows approximately 60% Hb S and 40% Hb F with no Hb A (138,139).

Sickling Disorders Involving a Thalassemic Gene

Hb S β^0 thalassemia may be indistinguishable clinically from sickle cell anemia, although splenomegaly usually persists in the adult; hematologically, the low MCV and MCH of Hb S β^0 thalassemia help to distinguish the two (140). Hb S β^+ thalassemia is usually a less severe sickling disorder than either sickle cell anemia

or Hb S β^0 thalassemia and may present as thalassemia intermedia (141). The β^+-thalassemia gene is responsible for the production of some Hb A, ranging from 5% to 30% of the circulating hemoglobin. The amount of Hb A is inversely related to the severity of the disease. Hemoglobin electrophoresis shows more Hb S than Hb A, a finding that in itself is virtually diagnostic of Hb S β^+ thalassemia in an untransfused individual; in addition, there is increased Hb A_2 and a small amount of Hb F (Fig. 42.18).

α Thalassemia of either the one- or two-gene deletion variety may be inherited by individuals also having sickle cell trait or sickle cell anemia. In sickle cell trait, the percentage of Hb S is determined by the number of α genes—approximately 35% to 40% when the normal four α genes are functioning, 30% to 35% when three α genes are functioning (this combination of Hb AS with single-gene deletion α thalassemia is found in approximately 30% of individuals with sickle cell trait), and 25% when only two α genes are functioning (found in approximately 3% of individuals with sickle cell trait) (142). Iron deficiency also lowers the percentage of Hb S in sickle cell trait (90). Inheritance of α thalassemia in individuals with Hb C trait (Hb AC) has a similar effect on the proportion of Hb C. This effect occurs because the β^S- and β^C-globin chains have a positive charge relative to β^A and therefore are less likely to combine with the limited number of σ-globin chains, which are also positively charged. The reverse occurs when an abnormal β-globin chain (e.g., many examples of Hb J) has a relatively negative charge, and the percentage of the abnormal hemoglobin increases when combined with α thalassemia (143).

Hb C Disorders

Hb C ($\alpha_2\beta_2^{6\ Glu\rightarrow Lys}$) occurs in 2.4% of African Americans; 0.02% are homozygotes (Hb CC) (144). The mutation for Hb C apparently occurred in the Ghana and Upper Volta region of western Africa, where as many as 25% of the population in some areas are heterozygotes. Hb C is found almost exclusively in blacks (145).

Hb C trait (Hb AC) is an entirely benign disorder. The blood count is normal, but the blood smear shows many normochromic target cells. Hemoglobin electrophoresis shows approximately 45% Hb C and 55% Hb A (Fig. 42.6). Iron deficiency or the deletion of one or two α-globin genes results in a lower percentage of Hb C. Hb A_2 migrates with Hb C on cellulose acetate, and they are eluted together from ion-exchange columns; therefore, it is necessary to use chromatographic methods (such as HPLC) for the quantitation of Hb A_2 in the presence of Hb C.

Hb C disease (Hb CC) is associated with a mild to moderate chronic hemolytic anemia and splenomegaly (146). Apart from the aplastic crises and cholelithiasis, patients homozygous for Hb C have few complications and do not experience vasoocclusive crises, sequestration crises, or infections prevalent among patients with sickle cell anemia. The red cells show the unusual combination of reduced MCV but increased MCHC (147). The blood smear shows even more target cells than Hb C trait; it also contains occasional spherocytes, a reflection of the increased MCHC. Intracellular or extracellular Hb C crystals, blocklike structures 1 to 3 μm in length, may be seen in the blood smear and are pathognomonic of Hb C. In Hb C disease, there is at least 90% Hb C on cellulose acetate electrophoresis at pH 8.4, with a slight increase in Hb F and no Hb A (Fig. 42.18). In acid agar, pH 6.0, the normal or slightly increased amount of Hb A_2 may be visible, migrating as Hb A. This apparent combination of Hb C with a small amount of Hb A must not be confused with Hb C/β^+ thalassemia; the absence of Hb A on cellulose acetate is helpful in this distinction. Hb E and Hb O migrate with Hb C on cellulose acetate but can be distinguished by acid agar electrophoresis (Fig. 42.6).

Compound heterozygotes for Hb S and Hb C are discussed in the section on the sickling disorders. Hb C/β^0 thalassemia and Hb C/β^+-thalassemia are moderate and mild hemolytic anemias, respectively. The electrophoretic pattern of Hb C/β^0 thalassemia is indistinguishable from that of Hb C disease (Hb CC), but the MCV is somewhat lower (50 to 70 fl) than in Hb CC (mean, 72 fl) (148). In Hb C/β^+ thalassemia, the quantity of Hb C exceeds that of Hb A, and in Hb C/α thalassemia, the reverse is found in a manner analogous to the proportions of Hb S and Hb A in the Hb S β^+ and Hb S α thalassemia.

Hemoglobin E

Hb E ($\alpha_2\beta_2^{26\ Glu\rightarrow Lys}$) is the second most common hemoglobin variant worldwide after Hb S. It occurs mainly in southeast Asian populations and is common in Thailand, Laos, and Cambodia; it is rare in ethnic Vietnamese or Chinese. It has also been reported rarely in blacks and whites (149).

Hb E trait (Hb AE) is a benign disorder associated with microcytic red cells (average MCV 72 fl) but no anemia, not unlike thalassemia minor (150). The proportion of Hb E is 30% to 35% (a useful differential point from Hb C trait, in which the abnormal hemoglobin usually makes up 40% to 50%) (Fig. 42.6). This proportion may be even lower when there is concomitant iron deficiency or α thalassemia.

Homozygous Hb E (Hb EE) is also a benign disorder, but there may be mild anemia, and the red cells are more microcytic than in Hb E trait, mimicking β thalassemia minor (151). Red cells contain more than 90% Hb E and a slightly increased quantity of Hb F, but no Hb A (Fig. 35.18).

In contrast to the mild nature of Hb E disease, the combination of Hb E/β^0 thalassemia may present as thalassemia intermedia or even thalassemia major (152). There is a greater degree of anemia and microcytosis than in homozygous Hb E. The electrophoretic pattern may be similar to that of a homozygote for Hb E but usually shows more Hb F.

Hb E migrates with Hb C, Hb O, and Hb A_2 on cellulose acetate at pH 8.4 but can be distinguished by electrophoresis in acid citrate agar, where it migrates with Hb A (Fig. 42.6). Furthermore, Hb E is mildly unstable and may be precipitated by heat (50° to 60°C) or 17% isopropanol; however, there is no evidence of hemolysis even in the homozygote. The thalassemic nature of this variant can be explained by the amino acid substitution at β^{26}, which creates a new splicing sequence and interferes with mRNA processing. If the new splicing site is used, a defective mRNA is likely produced; if the original splicing site is used, Hb E is produced. Thus, this mutation acts similar to other β+ thalassemia mutations that affect mRNA splicing (153).

OTHER HEMOGLOBIN VARIANTS

Hb G-Philadelphia (α68(E17)Asn→lys) is the most common α-chain variant seen in the United States and is found primarily in African Americans (144). Because of this, it can be seen in combination with ß-chain variants also seen in African Americans, such as Hb S or Hb C. These combinations produce electrophoretic patterns with multiple bands but do not cause any clinical symptoms. Hb G-Philadelphia trait is clinically benign but is often associated with the α-thalassemia 2 trait (154,155).

Hb-Hasharon (α47[CE5] asp→his) is found in two distinct ethnic groups either in descendants of Ashkenazi Jews of Central Europe or Italians from the Ferrara district of Italy. Although this hemoglobin variant is unstable *in vitro*, it is not associated with clinical or hematologic effects (156,157).

Hb J-Baltimore (ß16(A13)gly-asp) (158) Hb N-Baltimore (ß95(F62)lys-glu) (159), and Hb I-(α16(A14)lys-gly) (160) are "fast" hemoglobin variants that are uncommonly seen. These variants are clinically benign; however, they will give spuriously high Hb A_1c values using some ion-exchange chromatography methods (161).

HEMOGLOBIN VARIANTS AFFECTING STABILITY OR FUNCTION

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

Amino acid substitutions, deletions, or additions that affect the heme pocket, the points of contact or interaction between globin chains, their secondary helical structure, or the hydrophobic interior of the subunit almost invariably alter the stability or function of the hemoglobin molecule. These variants may be divided into the unstable hemoglobins, hemoglobins with high or low oxygen affinity, and the hereditary methemoglobins (Hb Ms). This separation is not absolute; for instance, an unstable hemoglobin may also have an altered oxygen affinity and result in slight methemoglobinemia.

Unstable Hemoglobins

Hemolysis associated with the precipitation of denatured hemoglobin (Heinz bodies) within the red cell results from two main aberrations: oxidative hemolysis affecting either normal red cells or red cells deficient in one of the enzymes in the glycolytic pathway and the presence of an unstable hemoglobin. More than 150 unstable hemoglobin variants have been identified (13,162). Of these, approximately 70 result in clinical and laboratory evidence of hemolytic anemia. The remainder are unstable in vitro but do not produce clinical syndromes. Most are ß-globin chain defects, which on the whole are more severe than α-globin chain defects because the latter affect only one of four α genes as opposed to one of two ß genes. Table 42.9 contains examples chosen to demonstrate the differences in severity (which are unrelated to the quantity of unstable hemoglobin in the circulation), the proportion of abnormal hemoglobin, and associated functional defects.

TABLE 42.9. UNSTABLE HEMOGLOBINS—APPROXIMATELY 150 VARIANTS[a]

Hb	Globin Chain	Affected Abnormal Hb	Percentage of O_2 Affinity	Clinical Severity
Köln	ß	10	Increased	Mild
Gun Hill	ß	30	Increased	Mild/moderate
Zurich	ß	25	Increased	Mild
Hammersmith	ß	30	Decreased	Severe
Cranston	ß	30	Increased	Mild
Hasharon	α	15	Normal	Normal
Kansas	ß	45	Decreased	Normal
Terre Haute	ß	0	—	Moderate/severe

[a] These hemoglobins (Hb) are all heterozygotes and indicate the globin chain affected, the percentage of abnormal hemoglobin found in the circulation, their oxygen affinity, and clinical severity. In cases (e.g., Hb Indianapolis and Hb Hasharon) in which the percentage of abnormal hemoglobin is significantly less than 50% (ß-globin abnormalities) or 25% (α-globin abnormalities), the imbalance also results in a thalassemia syndrome.

The denatured hemoglobin in the form of single or multiple Heinz bodies becomes linked to the inner side of the red cell membrane, decreasing the pliability of the red cell and its ability to traverse the microcirculation, particularly in the spleen. The Heinz bodies may be removed from the red cell together with a portion of the red cell membrane, leaving a "bite" cell, or the whole red cell may be removed from the circulation. Not surprisingly, splenectomy results in an increased number of circulating red cells containing Heinz bodies, while at the same time lengthening the red cell life span.

Hb Köln (ß98(FG5)val→met) is the prototype of unstable Hb variants and is the most common unstable Hb encountered. The substitution occurs at the ß98 position, which is the fifth amino acid at the "FG corner." This substitution is just inside the heme pocket (which accounts for its instability) but is also a contact point between the α_1, and $ß_2$ subunits. Thus, Hb Köln has been considered a high oxygen affinity Hb variant. It is found in many different ethnic groups and has probably arisen several times as a *de novo* mutation (163).

In addition to the anemia, reticulocytosis, and bilirubinemia common to most hemolytic anemias, there are laboratory tests that may indicate the presence of an unstable hemoglobin. Two simple screening tests for an unstable hemoglobin involve applying a physical or ionic stress to the hemoglobin molecule. Heating a buffered lysate at 60°C for half an hour or 50°C for 1 hour will cause flocculation of any unstable hemoglobin present (164). The addition of 17% isopropanol to an equal volume of buffered lysate (the Carrell test) will also cause flocculation, within 10 minutes, of unstable hemoglobin (165). In either test,

the flocculus may be small if the quantity of unstable hemoglobin is small. The flocculus is white if heme has been lost but retains the red color of hemoglobin if heme remains attached to the precipitated hemoglobin. Hemoglobin electrophoresis may demonstrate an abnormal fraction, but several unstable hemoglobins migrate with Hb A.

It is always worthwhile to perform all available tests and even repeat them on a fresh sample of blood because the amount of unstable hemoglobin present may be small and its instability not marked.

Hemoglobins with Altered Oxygen Affinity

The presence of hemoglobins with increased or decreased oxygen affinity results in a resetting of the homeostatic mechanism that controls the amount of circulating hemoglobin. Thus, a hemoglobin with increased oxygen affinity will accept at least normal amounts of oxygen in the lungs but will unload less oxygen to the tissues. This relative hypoxia will turn on the production of erythropoietin and increase the number of circulating red cells (and hemoglobin) so that more oxygen can be delivered to the tissues. A physiologic equilibrium will be reached at a hemoglobin level above or at the upper limit of a reference range based on individuals with functionally normal hemoglobin. The reverse occurs when a hemoglobin of low oxygen affinity is present, and the physiologic level of hemoglobin may be below the reference range. In rare instances, the oxygen affinity of the hemoglobin (e.g., Hb-Kansas, $\alpha_2\beta_2^{102asn\rightarrow Thr}$) may be so low that the uptake of oxygen in the lungs is impaired. Despite the ease with which oxygen is delivered to the tissue, the level of hemoglobin lies within the normal reference range (166).

More than 70 hemoglobin variants with an oxygen affinity high enough to cause erythrocytosis have been identified (13). Affected individuals are asymptomatic except for those rare examples in which the red cell mass is large enough to cause hyperviscosity. Therapy is seldom indicated, and pregnancy involves no risk despite the high maternal oxygen affinity relative to that of the fetus.

The presence of a high oxygen affinity hemoglobin may be first suspected from the blood count. Other laboratory findings include an abnormal electrophoretic pattern in approximately 70% of examples. Measurement of the P_{50} (the oxygen tension at which 50% of the hemoglobin is saturated) will confirm the diagnosis. Oxygen affinity is inversely proportional to the P_{50}, and in general the hemoglobin level or degree of erythrocytosis is directly related to the P_{50}.

Many unstable hemoglobins also have an increased oxygen affinity that may explain the high hemoglobin level found in individuals whose hemoglobin has this combination of defects. More than 25 examples of unstable hemoglobins with reduced oxygen affinity have been reported (167).

The M Hemoglobins

Seven structurally abnormal hemoglobins exist only as methemoglobin because the heme iron is stabilized in the oxidized Fe^{3+} or ferriheme form (13,168,169 and 170). In four, the amino acid substitution involves one of the histidines between which the heme is suspended: Hb M-Boston, $\alpha_2^{58His\rightarrow Tyr}\beta_2$; Hb M-Iwate, $\alpha_2^{87His\rightarrow Tyr}$ Hb M-Saskatoon, $\alpha_2\beta_2^{63His\rightarrow Tyr}$; Hb M-Hyde Park, $\alpha_2\beta_2^{92His\rightarrow Tyr}2$ (also known as Hb M-Milwaukee-2). In Hb M-Milwaukee-1, $\alpha_2\beta_2^{67Va\rightarrow Glu}$, the substituted glutamate in the heme pocket forms a similar bond, fixing the ferriheme. There are two examples of fetal M hemoglobins (Hb F-M-Osaka, $\alpha_2\gamma_2^{63His\rightarrow Tyr}$ and Hb F-M-Fort Ripley, $\alpha_2\gamma_2^{92His\rightarrow Tyr}$), which cause cyanosis at birth and disappear as γ-globin chain production switches to ß-globin chain production. The substitutions in the γ chains are analogous to those in the ß chains of Hb M-Saskatoon and Hb M-Hyde Park.

The cyanosis seen in individuals with Hb M is only in part owing to hypoxia; methemoglobin imparts a brownish color to whole blood, which appears blue when seen through the skin. Most patients are asymptomatic; however, Hb M-Saskatoon and Hb M-Hyde Park have been associated with compensated hemolysis and oxidant drug-induced hemolysis.

Diagnosis is often suggested by the appearance of the patient and a family history of cyanosis. In the laboratory, the absorption spectra of the hemoglobin from affected individuals when it has been fully oxidized with ferricyanide differ from methemoglobin A, but only slightly one from another, with absorption peaks in the 580- to 600-nm range (170,171). Most of the M hemoglobins do not separate from Hb A on alkaline electrophoresis; however, IEF, particularly if the specimen is oxidized with ferricyanide, allows good separation of Hb Ms from Hb A.

HEMOGLOBINOPATHIES IN THE FETUS AND NEWBORN

Part of "42 - The Thalassemia and Hemoglobinopathy Syndromes"

The earlier the diagnosis of the severe hemoglobinopathies such as sickle cell anemia and ß thalassemia major can be made, the sooner appropriate therapy and prophylaxis can be instituted. Prenatal diagnosis (172,173) is seldom available outside major hemoglobinopathy centers because it usually involves the study of DNA from chorionic villus cells, amniocytes, or a fetal blood sample. Obtaining these samples exposes the fetus to some danger; therefore, only in fetuses proven to be at risk (by study of both parents) and in situations in which the parents would consider termination are such studies usually undertaken. The earliest diagnosis can be made from DNA of chorionic villus cells. DNA has the added advantage that it can survive transport across vast distances, allowing remote areas to take advantage of the gene analysis. Amplification of segments of the DNA by PCR techniques increases the material available for study.

The gene for Hb S can be specifically detected (Fig. 42.19), but the situation is different for ß thalassemia major because of the many different lesions, inside and outside the ß gene that may cause it. Such techniques as restriction endonuclease mapping and linkage analysis of restriction fragment length polymorphisms and oligonucleotide probes may be required. The vast amount of genetic data (61) has shown that within any one population, there are usually only a few lesions that need to be considered.

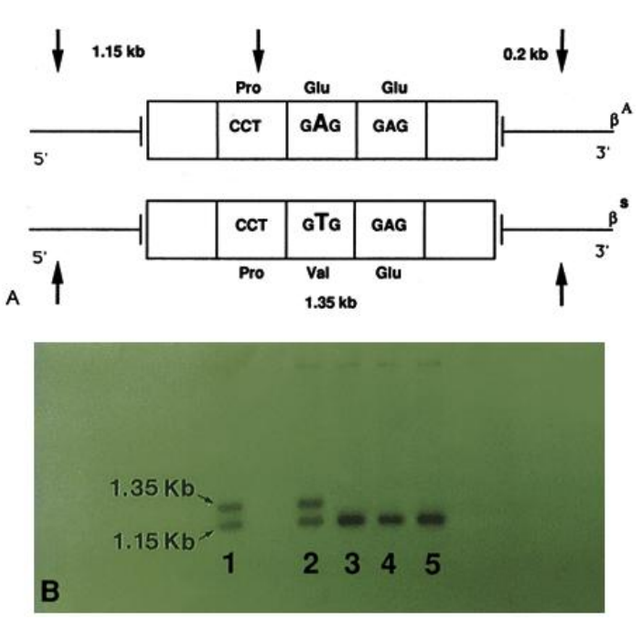


FIGURE 42.19. A: ß-Globin gene with MST II restriction sites (*arrows*). The nucleotide triplets coding for the 5th, 6th, and 7th amino acids are enlarged to illustrate the mutation causing loss of a restriction site in the β^S-globin gene. (Adapted from Law DJ, Frossard PM, Rucknagle DL. High sensitive and rapid gene mapping using miniaturized blot hybridization: application to prenatal diagnosis. *Gene* 1984;28:153–158.) **B:** Autoradiograph of MST II digestions using ß-globin gene probe: β^A, 1.15 kb; β^S, 1.35 kb. Lanes 1 and 2, sickle cell trait; lanes 3–5, normal adult.

The ample availability of cord blood at delivery allows application of many of the methods used in studying adults (174,175 and 176).

The ready availability also permits screening of all populations at risk, and in many states, it is mandated that all newborns be tested for hemoglobinopathies. However, there are significant limitations and pitfalls.

In the case of sickle cell anemia, hemoglobin electrophoresis on cellulose acetate at alkaline pH shows approximately 75% Hb F and 25% Hb S with no Hb A. It is the latter that is important, and it may be difficult to confirm that no Hb A is present. Because Hb A and its variants, particularly in heterozygotes, are present in small quantities, it is advisable to double the amount of hemolysate applied. Electrophoresis in citrate agar will help to confirm the identity of Hb S and make the presence or absence of Hb A more clearcut (Fig. 42.20). It must be remembered that Hb F and Hb S without Hb A is a pattern also present in Hb S $ß^0$ thalassemia and Hb S HPFH; study of the parents will usually clarify the diagnosis.

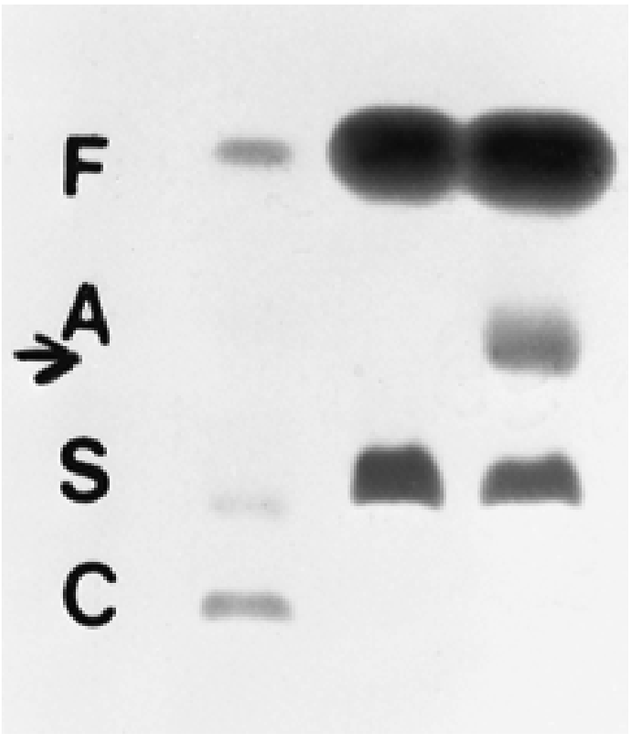


FIGURE 42.20. Hemoglobin electrophoresis in agar using a citric acid–citrate buffer at pH 6.2. **Left lane:** A control mixture of Hb C, Hb S, and Hb F. **Middle lane:** Cord blood from a neonate with sickle cell anemia shows Hb S and Hb F, but not Hb A (a similar pattern would be seen in Hb S-$ß^0$ thalassemia and Hb S HBFH). **Right lane:** Cord blood from a neonate with sickle cell trait shows Hb S, Hb F, and Hb A.

The presence of maternal Hb A may result in a diagnosis of sickle cell trait when in fact the infant has sickle cell anemia. This trap may be avoided by noticing the presence of a Hb A_2 band, which should not be present in cord blood, and therefore must be of maternal origin.

In ß thalassemia major, cord blood electrophoresis will show 100% Hb F. $ß^+$ Thalassemia, conversely, will show a small amount of Hb A, and the pattern will be difficult to distinguish from normal.

REFERENCES

1. Körber E. Inaugural dissertation: Uber differenzen blutfarbstoffes. Dorpat 1866, cited by Bischoff H. *Z Gesamte Exp Med* 1926;48:472–489.

2. Savitt TL, Goldberg MF. Herrick's 1910 case report of sickle cell anemia. *JAMA* 1989;261:266–271.

3. Cooley TB, Lee P. A series of cases of splenomegaly in children with anemia and peculiar bone changes. *Trans Am Pediatr Soc* 1925;37:29–30.

4. Rietti F. Ittero emolitica primitivo. *Atti Acad Sci Med Nat Ferrara* 1925;2:14–22.

5. Greppi E. Ittero emolitica familiare con aumento della resistanza deli globuli. *Minerva Med* 1928;8:1–11.

6. Micheli F, Penati F, Momigliano LG. Ulteriori richerche sulla anemia ipocromica splenomegalica con poichilocitosi. *Atti Soc Ital Ematol Hematol* 1933;16(Suppl):10–20.

7. Wintrobe MM, Matthews E, Pollack R, et al. Familial hematopoietic disorder in Italian adolescents and adults resembling Mediterranean disease (thalassemia). *JAMA* 1940;14:1530–1538.

8. Whipple GH, Bradford WL. Mediterranean disease thalassemia (erythroblastic anemia of Cooley); associated pigment abnormalities simulating hemochromatosis. *J Pediatr* 1936;9:279–311.

9. Pauling L, Itano HA, Singer SJ, et al. Sickle-cell anemia, molecular disease. *Science* 1949;110:543–548.

10. Hörlein H, Weber G. Uber chronische familiare Methamaglobinamie und eine neue Modifikation des Methamaglobins. *Dtsch Med Wochenschr* 1948;73:746–748.

11. Staven P, Strome J, Lorkin PA, et al. Haemoglobin M Saskatoon with slight contrast haemolysis, markedly increased by sulphonamides. *Scand J Haematol* 1972;9:566–571.

12. Ingram VM. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* 1956;178:792–794.

13. Huisman THJ, Carver MFH, Efremov GD. *A syllabus of human hemoglobin variants*, 2nd ed. Augusta, GA: Sickle Cell Anemia Foundation, 1998.

14. Perutz MF. Molecular anatomy, physiology and pathology of hemoglobin. In: Stamatoyannopoulos G, Neinhuis AW, Leder P, et al., eds. *Molecular basis of blood diseases*. Philadelphia: Saunders, 1987:127–178.

15. Fermi G, Perutz MF, Shannon B, et al. The crystal structure of human deoxyhaemoglobin at 1.7 ö resolution. *J Mol Biol* 1984;175:159–174.

16. Haldane JBS. The rate of mutation of human genes. Proceedings of the VIII International Congress on Genetics and Heredity. *Hereditas* 1949(Suppl):267–273.

17. Bunn HG, Forget BG. *Hemoglobin: molecular, genetic and clinical aspects*, 2nd ed. Philadelphia: Saunders, 1986:13–35.

18. Bohr C, Hasselbach K, Krogh A. Ueber einen in biologischer Beziehung Wichtigen Einfluss den die Kohlensaurespann ung des Blutes aut dessen Sauerstoff bindung ubt. *Scand Arch Physiol* 1904;16:402–412.

19. Perutz MF, Muirhead H, Mazzarella L, et al. Identification of residues responsible for the alkaline Bohr effect in haemoglobin. *Nature* 1969;222:1240–1243.

20. Reeves RB. The effect of temperature on the oxygen equilibrium curve of human blood. *Respir Physiol* 1980;42:317–328.

21. Oski FA, Gottlieb AJ. The interrelationships between red blood cell metabolites, hemoglobin, and the oxygen dissociation curve. *Prog Hematol* 1971;7:33–67.

22. Barcroft J. *The respiratory function of the blood. Part II. Haemoglobin*. London: Cambridge University Press, 1928.

23. Buckle VJ, Higgs DR, Wilkie AOM, et al. Localisation of human α globin to 16p13.3 pter. *J Med Genet* 1988;25:847–859.

24. Little PFR. Globin pseudogenes. *Cell* 1982;28:683–684.

25. Higgs DR, Vickers MA, Wilkie AOM, et al. A review of the molecular genetics of the human α-globin gene cluster. *Blood* 1989;73:1081–1104.

26. Sanders-Haigh L, Anderson WF, Francke U. The ß-globin gene is on the short arm of chromosome 11. *Nature* 1980;283:683–686.

27. Nute PE, Stamatoyannopoulos G. Cases of abnormal human hemoglobin produced by de novo mutation. *Hemoglobin* 1988;12:429–439.

28. Higgs DR, Wood WG, Barton C, et al. Clinical features and molecular analysis of acquired Hb H disease. *Am J Med* 1983;75:181–191.

29. Weatherall DJ, Pembrey ME, Pritchard J. Fetal haemoglobin. *Clin Haematol* 1974;3:467–508.

30. Weatherall DJ. Common genetic disorders of the red cell and the "malaria hypothesis." *Ann Trop Med Parasitol* 1987;81:539–548.

31. Anonymous. Ovalocytosis and malaria [Editorial]. *Lancet* 1988;2:608–610.

32. Mitchell GH, Bannister LH. Malaria parasite invasion: interactions with the red cell membrane. *Crit Rev Oncol Hematol* 1988;8:255–310.

33. Livingstone FB. The Duffy blood groups, vivax malaria and malaria selection in human populations: a review. *Hum Biol* 1984;56:413–425.

34. Fleming AF, Storey J, Molineaux L, et al. Abnormal haemoglobins in the Sudan savanna of Nigeria. I. Prevalence of haemoglobins and relationships between sickle cell trait, malaria and survival. *Ann Trop Med Parasitol* 1979;73:161–172.

35. Nagel RL, Roth EF Jr. Malaria and red cell genetics. *Blood* 1989;74:1213–1221.

36. Siniscalco M, Bernini L, Filippi I, et al. Population genetics of human haemoglobin variants, thalassemia and glucose-6-phosphate dehydrogenase deficiency, with particular reference to the malaria hypothesis. *Bull WHO* 1966;34:379–393.

37. Flint J, Hill AV, Bowden DK, et al. High frequencies of α-thalassemia are the result of natural selection by malaria. *Nature* 1986;321:744–750.

38. Bienzle U, Guggenmoos-Holzman I, Luzzatto L. Malaria and erythrocyte glucose-6-phosphate dehydrogenase variants in West Africa. *Am J Trop Med Hyg* 1979;28:619–621.

39. Beutler E, Dern RJ, Alving AS. The hemolytic effect of primaquine: V1. An in vitro test for sensitivity of erythrocytes to primaquine. *J Lab Clin Med* 1955;45:40–50.

40. Beutler E. Heinz body staining. In: Beutler E, Lichtman MA, Coller BS, et al, eds. *Williams hematology*, 5th ed. New York: McGraw-Hill, 1995:L26.

41. Kim HC, Schwartz E. Unstable hemoglobins. In: Beutler E, Lichtman MA, Coller BS, et al, eds. *Williams hematology*, 5th ed. New York: McGraw-Hill, 1995:L33.

42. Schmidt RM, Wilson SM. Standardization in detection of abnormal hemoglobins. Solubility tests for hemoglobin S. *JAMA* 1973;225:1225–1230.

43. Barrett RC. Rapid cellulose acetate electrophoresis, II. Qualitative and quantitative hemoglobin fractionation. *Clin Chem* 1993;9:325–329.

44. Briere RO, Golias T, Batsakis JG. Rapid qualitative and quantitative hemoglobin fractionation: cellulose acetate electrophoresis. *Am J Clin Pathol* 1965;44:695–701.

45. Robinson AR, Robson M, Harrison AP, et al. A new technique for differentiation of hemoglobin. *J Clin Lab Med* 1957;50:745–752.

46. Milner PF, Gooden HM, General RT. Rapid citrate agar electrophoresis in routine screening for hemoglobinopathies using a simple hemolysate. *Am J Clin Pathol* 1975;64:58–64.

47. Winter WP, Yodh J. Interaction of human hemoglobin and its variants with agar. *Science* 1983;221:175–178.

48. Winter WP, Seale WR, Yodh J. Interaction of human hemoglobin with anionic polysaccharides. *Am J Pediatr Heme Oncol* 1984;6:77–81.

49. Koepke JA, Thoma JF, Schmidt RM. Identification of human hemoglobins by use of isoelectric focusing in gel. *Clin Chem* 1975;21:1953–1955.

50. Beuzard Y, Galacteros F, Braconnier F, et al. Isoelectric focusing of human hemoglobins. *Prog Clin Biol Res* 1981;60:177–195.

51. Black J. Isoelectric focusing in agarose gel for detection and identification of hemoglobin variants. *Hemoglobin* 1984;8:117–127.

52. Schneider RG. Differentiation of electrophoretically similar hemoglobins such as S, D, G, and P, or A_2, C, E, and O by electrophoresis of globin chains. *Clin Chem* 1974;20:1111–1115.

53. Schneider RG, Barwick RC. Measuring relative electrophoretic mobilities of mutant hemoglobins and globin chains. *Hemoglobin* 1978;2:417–435.

54. Pembrey ME, McWade P, Weatherall DJ. Reliable routine estimation of small amounts of foetal haemoglobin by alkali denaturation. *J Clin Pathol* 1972;25:738–740.

55. Efremov CD, Huisman THJ, Bowman K, et al. Microchromatography of hemoglobins. II. A rapid method for the determination of hemoglobin A_2. *J Lab Clin Med* 1974;83:657–664.

56. Brosius EM, Wright JM, Baine RM, et al. Microchromatographic methods for Hb A_2 compared. *Clin Chem* 1978;24:2196–2199.

57. Henson J, Carver J, Wilson JB, et al. Carboxymethyl-cellulose microchromatography for the quantitation of hemoglobin Bart's (δ_4) and its use in the detection of α-thalassemia conditions. *J Chromatogr* 1980;198:443–448.

58. Ou CN, Rognerud CL. Rapid analysis of hemoglobin variants by cation-exchange HPLC. *Clin Chem* 1993;39:820–824.

59. Huisman THJ. Separation of hemoglobins and hemoglobin chains by high performance liquid chromatography. *J Chromatogr* 1987;418:277–304.

60. Weatherall DJ, Clegg JB, Naughton MA. Globin synthesis in thalassemia: an in vitro study. *Nature* 1965;208:1061–1065.

61. Kazazian HH Jr. The thalassemia syndromes: molecular basis and prenatal diagnosis in 1990. *Semin Hematol* 1990;27:209–228.

62. Bowie LJ, Reddy PL, Beck KR. Alpha thalassemia and its impact on other clinical conditions. *Clin Lab Med* 1977;17:97–108.

63. Wetherall DJ. The thalassemias. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW, et al., eds. *The molecular basis of blood diseases*, 2nd ed. Philadelphia. Saunders, 1994:157–206.

64. Thein SL, Al-Hakim I, Hoffbrand AV. Thalassaemia intermedia: a new molecular basis. *Br J Haematol* 1984;56:333–337.

65. Moi P, Paglietti E, Sanna A, et al. Delineation of the molecular basis of δ- and normal Hb A_2 ß-thalassemia. *Blood* 1988;72:530–533.

66. Pirastu M, Galanello R, Melis MA, et al. δ^+-Thalassemia in Sardinia. *Blood* 1983;62:341–345.

67. Heisman THJ, Reese AL, Gardiner MB, et al. The occurrence of different levels of $^{G}\gamma$ chain and of the $^{A}\gamma^{T}$ variant of fetal hemoglobin in newborn babies from various countries. *Am J Hematol* 1983;14:133–148.

68. Liebhaber SA. α Thalassemia. *Hemoglobin* 1989;13:685–731.

69. Huisman THJ, Carver MFH, Baysal E. *A syllabus of thalassemia mutations (1997)*, 1st ed. Augusta, GA: Sickle Cell Anemia Foundation, 1997:235–296.

70. Pierce HL, Kurachi S, Sofroniadore K, et al. Frequencies of thalassemia in American blacks. *Blood* 1977;49:981–986.

71. Bowden DK, Hill AV, Higgs DR, et al. Different hematologic phenotypes are associated with the leftward (-alpha 4.2) and rightward (-alpha 3.7) alpha$^+$-thalassemia deletions. *J Clin Invest* 1987;79:39–43.

72. Skogerboe KJ, West SF, Smith C, et al. Screening for α thalassemia: correlation of hemoglobin H in inclusion bodies with DNA determined genotype. *Arch Pathol Lab Med* 1992;116:1012–1018.

73. Wasi P, Na-Nakorn S, Pootrakul SN. The α-thalassemias. *Clin Haematol* 1974;3:383–410.

74. Milner PF, Clegg JB, Weatherall DJ. Haemoglobin H disease due to a unique haemoglobin variant with an elongated α-chain. *Lancet* 1971;I:729–732.

75. Clegg JB, Weatherall DJ. Hemoglobin Constant Spring, an unusual α-chain variant involved in the etiology of hemoglobin H disease. *Ann N Y Acad Sci* 1974;232:168–178.

76. Clegg JB, Weatherall DJ, Milner PF. Haemoglobin Constant Spring—a chain termination mutant? *Nature* 1971;234:337–340.

77. Pootrakul P, Winichagoon P, Fucharoen S, et al. Homozygous haemoglobin Constant Spring: a need for revision of concept. *Hum Genet* 198;59:250–255.

78. Pongsamart S, Pootrakul S, Wasi P, et al. Hemoglobin Constant Spring: hemoglobin synthesis in heterozygous and homozygous states. *Biochem Biophys Res Commun* 1975;64:681–686.

79. Merritt D, Jones RT, Head L, et al. Hb Seal Rock [(α2)142 term→glu, codon 142 TAA→GAA]: an extended α-chain variant associated with anemia, microcytosis, and α-thalassemia-2 (-3.7 kb). *Hemoglobin* 1997;21:331–344.

80. Efremov GD, Josifovska O, Nikolov N, et al. HbH-Hb Icaria disease: identification of the Hb Icaria mutation through analysis of amplified DNA. *Br J Haematol* 1990;75:250–253.

81. De Jong WW, Meera Kahn P, Bernini LF. Hemoglobin Koya Dora: high frequency of a chain termination mutant. *Am J Hum Genet* 1975;27:81–90.

82. Waye JS, Eng B, Patterson M, et al. Identification of a novel termination codon mutation (TAA→TAT, term→tyr) in the α2 globin gene of a Laotion girl with hemoglobin H disease. *Blood* 1994;83:3418–3420.

83. Kanavakis E, Traeger-Synodinos J, Papasotiriou I, et al. The interaction of α$^\circ$thalassemia with Hb Icaria: three unusual cases of haemoglobinopathy H. *Br J Haematol* 1996;92:332–335.

84. Chui DHK, Waye JS. Hydrops fetalis caused by α-thalassemia: an emerging health care problem. *Blood* 1998;91:2213–2222.

85. Weatherall DJ, Clegg JB, Boon WH. The haemoglobin constitution of infants with the haemoglobin Bart's hydrops foetalis syndrome. *Br J Haematol* 1970;18:357–367.

86. Kazazian HH, Boehm CD. Molecular basis and prenatal diagnosis of ß-thalassemia. *Blood* 1988;62:1107–1116.

87. Huisman THJ, Carver MFH, Baysal E. *A syllabus of thalassemia mutations (1997)*, 1st ed. Augusta, GA: Sickle Cell Anemia Foundation, 1997:211–234.

88. Coleman MB, Steinberg MH, Adams JG III. Hemoglobin Terre Haute arginine beta 106. A posthumous correction to the original structure of hemoglobin Indianapolis. *J Biol Chem* 1991;266:5798–5800.

89. Millard DP, Mason K, Serjeant BE, et al. Comparison of haematological features of the ß$^\circ$- and ß$^+$-thalassaemia traits in Jamaican Negroes. *Br J Haematol* 1977;36:161–170.

90. Kuczynski A. The relationship between the serum iron concentration and erythrocytic hemoglobin A_2 level. *J Med* 1971;2:136–142.

91. Steinberg MH. Review: thalassemia: molecular pathology and management. *Am J Med Sci* 1988;31:308–321.

92. Yataganas X, Fessas P. The pattern of hemoglobin precipitation in thalassemia and its significance. *Ann N Y Acad Sci* 1969;165:270–287.

93. Weatherall DJ, Clegg JB. *The thalassaemia syndromes*, 3rd ed. Oxford: Blackwell Scientific Publications, 1981:199–202.

94. Weatherall DJ, Clegg JB, Higgs DR, et al. The hemoglobinopathies. In: Scriver CR, Beaudet AL, Sly WS, eds. *The metabolic and molecular basis of inherited disease*, 7th ed. New York: McGraw-Hill, 1995:3417–3484.

95. Bollekens JA, Forget BG. δß-thalassemia and hereditary persistence of fetal hemoglobin. *Hematol Oncol Clin N Am* 1991;5:399–422.

96. Huisman THJ, Carver MFH, Baysal E. *A syllabus of thalassemia mutations (1997)*, 1st ed. Augusta, GA: Sickle Cell Anemia Foundation, 1997:211–234.

97. Weatherall DJ, Clegg JB. *The thalassaemia syndromes*, 3rd ed. Oxford: Blackwell Scientific Publications, 1981:428–431.

98. Boyer SH. The emerging complexity of genetic control of persistent fetal hemoglobin biosynthesis in adults. *Ann N Y Acad Sci* 1989;565:23–36.

99. Huisman THJ, Carver MFH, Baysal E. *A syllabus of thalassemia mutations (1997)*, 1st ed. Augusta, GA: Sickle Cell Anemia Foundation, 1997:187–210.

100. Kleihauer E, Braun H, Betke K. Demonstration von fetalem Hämoglobin in den Erythrocyten eines Blutausstrichs. *Klin Wochenschr* 1957;35:637–638.

101. Shepard MK, Weatherall DJ, Conley CC. Semiquantitative estimation of the distribution of fetal hemoglobin in red cell populations. *Johns Hopkins Hosp Bull* 1962;110:293–310.

102. Davis BH, Olsen S, Bigelow NC, et al. Detection of fetal red cells in fetomaternal hemorrhage using a fetal hemoglobin monoclonal antibody by flow cytometry. *Immunohematology* 1998;38:749–756.

103. Hoyer JD, Penz CS, Fairbanks VF, et al. A flow cytometric method for measurement of Hb F in red cells: applications in the evaluation of hereditary persistence of fetal hemoglobin (HPFH) and other conditions with elevated Hb F levels. *Blood* 1998;92(Suppl 1):40b.

104. Baglioni C. The fusion of two peptide chains in hemoglobin Lepore and its interpretation as a genetic deletion. *Proc Natl Acad Sci U S A* 1962;48:1880–1886.

105. Efremov GD. Hemoglobins Lepore and anti-Lepore. *Hemoglobin* 1978;2:197–233.

106. England JM, Fraser PM. Differentiation of iron deficiency from thalassaemia trait by routine blood-count. *Lancet* 1973;1:449–450.

107. Green R, King R. A new red cell discriminant incorporating volume dispersion for differentiating iron deficiency anemia from thalassemia minor. *Blood Cells* 1989;15:481–495.

108. Konetey-Ahulu FID, Gallo E, Lehmann H, et al. Haemoglobin Korle-Bu (ß73 aspartic acid replaced by asparagine) showing one of the two amino acid substitutions of haemoglobin C Harlem. *J Med Genet* 1968;5:107–111.

109. DeJong WWW, Went LN, Bernini LF. Haemoglobin Leiden: deletion of ß6 or 7 glutamic acid. *Nature* 1968;220:778–790.

110. Hoyer JD, Wick MJ, Snow K, et al. DNA sequence analysis of Hb Leiden [ß6(A3) or ß7 (A4)glu→0]. *Hemoglobin* 1999;23:299–302.

111. Winslow RM, Swenberg ML, Gross E, et al. Hemoglobin McKees Rocks ($\alpha_2\beta_2$145 Tyr→Term): a human "nonsense" mutation leading to a shortened ß-chain. *J Clin Invest* 1976;57:772–781.

112. Garel MC, Goossens G, Oudart JL, et al. Hemoglobin Dakar = Hemoglobin Grady: demonstration by a new approach to the analysis of the tryptic core region of the α chain and oxygen equilibrium properties. *Biochem Biophys Acta* 1976;453:459–471.

113. Hoyer JD, Wick MJ, Thibodeau SN, et al. Hb Tak confirmed by DNA analysis: not expressed as thalassemia in a Hb Tak/Hb E compound heterozygote. *Hemoglobin* 1998;22:45–52.

114. Seid-Akhaven M, Winter WP, Abramson RK, et al. Hemoglobin Wayne: a frame shift mutation detected in human hemoglobin alpha chains. *Proc Natl Acad Sci U S A* 1976;73:882–886.

115. Jandl JH. *Blood: textbook of hematology*, 2nd ed. Boston: Little, Brown, 1996:543.

116. Mears JG, Beldjord C, Benabadji M, et al. The sickle gene polymorphism in Africa. *Blood* 1981;58:599–601.

117. Bunn HF, Forget BG. *Hemoglobin: molecular genetic and clinical aspects*, 1st ed. Philadelphia: Saunders, 1986:453–501.

118. Kark JA, Posey DM, Schumacher HR, et al. Sickle-cell trait as a risk factor for sudden death in physical training. *N Engl J Med* 1987;317:781–787.

119. Jandl JH. *Blood: textbook of hematology*, 2nd ed. Boston: Little, Brown, 1996:544–555.

120. Bunn HF, Forget BG. *Hemoglobin: molecular genetic and clinical aspects*, 1st ed. Philadelphia: Saunders, 1986:502–564.

121. Reider RF, Safaya S, Gillette P, et al. Effect of ß-globin gene cluster haplotype on the hematologic and clinical features of sickle cell anemia. *Am J Hematol* 1991;36:184–189.

122. Powars DR. The natural history of sickle cell disease the first ten years. *Semin Hematol* 1975;12:276–285.

123. Francis RB Jr, Johnson CS. Vascular occlusion in sickle cell disease: current concepts and unanswered questions. *Blood* 1991;77:1405–1414.

124. Serjeant GR, Grandison Y, Lowrie Y, et al. The development of haematological changes in homozygous sickle cell disease: a cohort study from birth to 6 years. *Br J Haematol* 1981;48:533–543.

125. Davis LR. Aplastic crisis in haemolytic anaemias: the role of a parvovirus-like agent. *Br J Haematol* 1983;55:391–393.

126. Davies SC, Brozovic M. The presentation, management and prophylaxis of sickle cell disease. *Blood Rev* 1989;3:29–44.

127. Steinberg MH. Management of sickle cell disease. *N Engl J Med* 1999;340:1021–1030.

128. Castro O. Management of sickle cell disease: recent advances and controversies. *Br J Haematol* 1999;107:2–11.

129. Koshy M, Burd L, Wallace D, et al. Prophylactic red-cell transfusions in pregnant patients with sickle cell disease. A randomized cooperative study. *N Engl J Med* 1988;319:1447–1452.

130. Kaufman RE. Hydroxyurea: specific therapy for sickle cell anaemia [Editorial]. *Blood* 1992;79:2503–2506.

131. Vermylen C, Fernandez Robles E, Ninane J, et al. Bone marrow transplantation in five children with sickle cell anaemia. *Lancet* 1988;1:1427–1428.

132. River GL, Robbins AB, Schwartz SO. S-C hemoglobin: a clinical study. *Blood* 1961;18:385–416.

133. Ballas SK, Lewis CN, Noone AM, et al. Clinical, hematologic and biochemical features of Hb SC disease. *Am J Hematol* 1982;13:37–51.

134. Nagel RL, Lawrence C. The distinct pathobiology of sickle cell-hemoglobin C disease: therapeutic implications. *Hematol Oncol Clin North Am* 1991;5:433–451.

135. Bain BJ. Blood film features of sickle cell hemoglobin C disease. *Br J Haematol* 1993;83:516–518.

136. Milner PF, Miller C, Grey R, et al. Hemoglobin O Arab in four Negro families and its interaction with hemoglobin S and hemoglobin C. *N Engl J Med* 1970;283:1417–1425.

137. Kelleher JF, Park JOK, Kim HC, et al. Life-threatening SD-Los Angeles disease. *Hemoglobin* 1984;8:203–213.

138. Murray N, Serjeant BE, Serjeant GR. Sickle cell-hereditary persistence of fetal hemoglobin and its differentiation from other sickle cell syndromes. *Br J Haematol* 1988;69:89–92.

139. Jacob GF, Raper AB. Hereditary persistence of foetal haemoglobin production, and its interaction with the sickle cell trait. *Br J Haematol* 1958;4:138–149.

140. Serjeant GR, Sommereux A, Stevenson M, et al. Comparison of sickle cell-ßdeg; thalassaemia with homozygous sickle cell disease. *Br J Haematol* 1979;41:83–93.

141. Gonzalez-Redondo JM, Kutlar F, Kutlar A, et al. HbS(C)-ß$^+$-thalassemia: different mutations are associated with different levels of normal Hb A. *Br J Haematol* 1988;70:85–89.

142. Steinberg MH, Embury SH. α-Thalassemia in blacks: interactions with the sickle hemoglobin gene. *Birth Defects* 1988;23:43–48.

143. Huisman THJ. Percentages of abnormal hemoglobins in adults with a heterozygosity for an α-chain and/or a ß-chain variant. *Am J Hematol* 1983;14:393–404.

144. Schneider RG, Hightower B, Hosty TS, et al. Abnormal hemoglobins in a quarter million people. *Blood* 1976;48:629–637.

145. Bunn HF, Forget BG. *Hemoglobin: molecular, genetic and clinical aspects*, 1st ed. Philadelphia: Saunders, 1986:502–564.

146. Redetzki JE, Bickers JN, Samuels MS. Homozygous hemoglobin C disease: clinical review of 15 patients. *South Med J* 1968;61:238–242.

147. Murphy JR. Hemoglobin CC erythrocytes: decreased intracellular pH and decreased O_2 affinity-anemia. *Semin Hematol* 1976;13:177–180.

148. Bunn HF, Forget BG. *Hemoglobin: molecular, genetic, and clinical aspects*, 1st ed. Philadelphia: Saunders, 1986:422.

149. Bunn HF, Forget BG. *Hemoglobin: molecular, genetic, and clinical aspects*, 1st ed. Philadelphia: Saunders, 1986:425–427.

150. Fairbanks VF, Gilchrist GS, Brimhall B, et al. Hemoglobin E trait reexamined: a cause of microcytosis and erythrocytosis. *Blood* 1979;53:109–115.

151. Lachant NA. Hemoglobin E: an emerging hemoglobinopathy in the United States. *Am J Hematol* 1987;25:449–462.

152. Mehta BC, Agarwal MB, Verandan DJ, et al. Hemoglobin E thalassemia: a study of 16 cases. *Acta Hematol* 1980;64:201–204.

153. Orkin SH, Kazazian HH, Antonarakis SE, et al. Abnormal RNA processing due to the exon mutation of $ß^E$ globin gene. *Nature* 1982;300:768–769.

154. Reider RF, Woodbury DH, Rucknagel DL. The interaction of α-thalassemia and haemoglobin G Philadelphia. *Br J Haematol* 1976;32:159–165.

155. Rucknagel DL, Rising JA. A heterozygote for Hb $ß^S$ Hb $ß^C$ and Hb$α^G$-Philadelphia in a family presenting evidence for heterogeneity of human alpha chain loci. *Am J Med* 1975;59:53–60.

156. Charache S, Mondzac AM, Gessner V. Hemoglobin Hasharon ($α_2$147 his (CD5) $ß_2$): a hemoglobin found in low concentration. *J Clin Invest* 1969;48:834–847.

157. Albert R, Mariuzzi GM, Marinucci M, et al. Haemoglobin Hasharon in a North Italian community. *J Med Genet* 1975;12:294–296.

158. Weatherall DJ. Hemoglobin J (Baltimore) coexisting in a family with hemoglobin S. *Johns Hopkins Hosp Bull* 1964;114:1–12.

159. Clegg JB, Naughton MA, Weatherall DJ. An improved method for identification of human hemoglobin mutants: identification of $α_2ß_2$ 95 glu haemoglobin N (Baltimore). *Nature* 1965;207:945-947.

160. Rucknagel DL, Page EB, Jensen WN. Hemoglobin I: an inherited hemoglobin anomaly. *Blood* 1955;10:999–1009.

161. Fairbanks VF, Zimmerman BR. Measurement of glycosylated hemoglobin by affinity chromatography. *Mayo Clin Proc* 1983;58:770–773.

162. Ohba Y. Unstable hemoglobin. *Hemoglobin* 1990;14:353–388.

163. Carrell RW, Lehmann H, Hutchinson HE. Haemoglobin Köln (ß98 valine → methionine): an unstable protein causing inclusion body anaemia. *Nature* 1966; 210:915–916.

164. Dacie JV, Grimes AJ, Meisler A, et al. Hereditary Heinz-body anaemia. A report of studies on five patients with mild anaemia. *Br J Haematol* 1964;10:388–402.

165. Carrell RW, Kay R. A simple method for the detection of unstable haemoglobins. *Br J Haematol* 1972;23:615–619.

166. Bonaventura J, Riggs A. Hemoglobin Kansas, a human hemoglobin with a neutral amino acid substitution and abnormal oxygen equilibrium. *J Biol Chem* 1968;243:980–991.

167. Huismann THJ, Carver MFH, Efremov GD. *A syllabus of human hemoglobin variants*, 2nd ed. Augusta, GA: Sickle Cell Anemia Foundation 1998:21.

168. Gerald PS, Efron ML. Chemical studies of several varieties of Hb M. *Proc Natl Acad Sci U S A* 1961;47:1758–1767.

169. Beutler E. Methemoglobinemia and other causes of cyanosis. In: Beutler E, Lichtman MA, Coller BS, et al., eds. *Hematology*, 5th ed. New York: McGraw-Hill, 1995:654–663.

170. Priest JR, Watterson J, Jones RT, et al. Mutant fetal hemoglobin causing cyanosis in a newborn. *Pediatrics* 1988;83:734–736.

171. Gerald PS, George P. A second spectroscopically abnormal methemoglobin associated with hereditary cyanosis. *Science* 1959;129:393–394.

172. Rosatelli C, Pirastu M, Cao A. The prenatal diagnosis of thalassemia [Annotation]. *Br J Haematol* 1986;63:215–220.

173. Driscoll CM, Lerner N, Anyane-Yeboa K, et al. Prenatal diagnosis of sickle hemoglobinopathies: the experience of the Columbia University Comprehensive Center for Sickle Cell Disease. *Am J Hum Genet* 1987;40:548–558.

174. Consensus Conference. Newborn screening for sickle cell disease and other hemoglobinopathies. *JAMA* 1987;258:1205–1209.

175. Galacteros F, Kleman K, Caburi-Martin J, et al. Cord blood screening for hemoglobin abnormalities by thin layer isoelectric focusing. *Blood* 1980;56:1068–1071.

176. Jacobs S, Peterson L, Thompson L, et al. Newborn screening for hemoglobin abnormalities. A comparison of methods. *Am J Clin Pathol* 1986;85:713–715.

43

Acute Leukemias and Myelodysplastic Syndromes

Curtis A. Hanson

Serhan Alkan

- ACUTE LEUKEMIA
- ACUTE LYMPHOBLASTIC LEUKEMIA
- ACUTE MYELOGENOUS LEUKEMIA
- ACUTE BIPHENOTYPIC LEUKEMIA
- MYELODYSPLASTIC SYNDROMES

ACUTE LEUKEMIA

Part of "43 - Acute Leukemias and Myelodysplastic Syndromes"

Pathogenesis

Acute and chronic leukemias are malignancies that arise from uncontrolled clonal proliferations of hematopoietic cells. The normal cellular control mechanisms are thought to be inoperative in leukemia owing to changes in the genetic code that are responsible for regulation of cell growth and differentiation. It is often mistakenly believed that acute leukemias are hyperproliferative disorders that result in the production of new cells at a rapid pace. However, it is almost paradoxically true that leukemia cells may actually have cycling times that are several times slower than normal bone marrow cells. These leukemic cells mature slowly and incompletely and survive longer than their normal marrow counterparts because of the failure to achieve normal and final maturation. It is believed, for example, that AML cells may actually have doubling times of approximately 30 days and require 50 to 75 doubling times before there is a significant accumulation of malignant cells. This final accumulation may occur over a period of 1 to 10 years and thus makes it difficult to truly access the initiation of disease.

Leukemias undoubtedly arise as a result of multiple transformation steps, or “hits.” Significant evidence has been gathered that strongly implicates cellular oncogenes as key players in this leukemic transformation process. Oncogenes typically play an important role in normal cellular control and regulation and only display their oncogenic potential when their genetic structure or control elements are mutated, rearranged, amplified, or activated. Amplification and promotion of oncogenic activity may occur when the normal transcriptional promoter and control regions of these genes are disrupted. Various retroviruses have also been implicated as a cause or initiator of leukemia. We also know that leukemias may secondarily arise as a result of exposures to a variety of environmental and iatrogenic therapies.

Immunosuppression, for example, is known to increase the risk of leukemia and lymphoma in organ transplant patients receiving such therapy. Ionizing radiation as well as alkylating agents may also contribute to the problem of secondary leukemias.

Leukemias, in general, are defined as malignant neoplasms of the hematopoietic system arising in the bone marrow. As the bone marrow is replaced with the malignant cells, the excess malignant cells escape into the peripheral blood, hence the derivation of the name leukemia: white (leuk-)/blood (-emia). In a simplistic fashion, it is easiest to classify leukemias based on (a) the natural course of disease, i.e., acute versus chronic, and (b) the basic cell type involved (lymphoid versus myeloid). Acute leukemias are the result of a block in normal hematopoietic differentiation, leading to an accumulation of immature lymphoid or myeloid cells. This accumulation of leukemic blasts, as described above, is the result of a block in maturation and differentiation rather than an increased rapidity of differentiation. Acute leukemias are characterized by a rapidly fatal course of days to weeks, if untreated. This is in contrast to the chronic leukemias, which are typically associated with an indolent, albeit progressive, course of disease. Chronic leukemias are characterized by a proliferation of differentiated cells, as opposed to the immature blasts seen in the acute leukemias. Table 43.1 outlines the various acute and chronic leukemias; Figure 43.1 is a schematic diagram of their relationship to normal hematopoiesis. The relative incidence of the four major subgroups of leukemias is as follows: acute lymphoblastic leukemia (ALL), 10%; B-chronic lymphocytic leukemia (CLL), 30%; acute myelogenous leukemia (AML), 45%; and chronic myelogenous leukemia (CML), 15%.

TABLE 43.1. CLASSIFICATION OF LEUKEMIA

Acute leukemia
Acute lymphoblastic leukemia (ALL)
FAB: L1, L2, L3
IPh: B-precursor ALL
T-cell ALL
B-ALL (Burkitt's leukemia/lymphoma)
Acute myelogenous leukemia
FAB: M0-M7
Acute biphenotypic leukemia
Myelodysplastic syndromes (preleukemias)
Refractory anemia
Idiopathic refractory sideroblastic anemia
Refractory anemia with excess blasts
Refractory anemia with excess blasts in transformation
Chronic myelomonocytic leukemia
Myelodysplastic syndromes, NOS
Chronic leukemia
Chronic lymphoproliferative disorders (CLPD)
B-CLPD
B-Chronic lymphocytic leukemia (CLL)
Prolymphocytic leukemia (Galton's)
Hairy cell leukemia
Variants of CLL
Waldenström's macroglobulinemia
Multiple myeloma
T-CLPD
T-CLL
Tγ-Lymphoproliferative disorder
Adult T-cell leukemia/lymphoma
Sézary syndrome
Chronic myeloproliferative disorders (CMPD)
Chronic myelogenous leukemia
Polycythemia vera
Essential thrombocythemia
Idiopathic myelofibrosis
CMPD, NOS

FAB, French-American-British; IPh, immunophenotype; NOS, not otherwise specified.

Clinical Manifestations

The clinical manifestations of acute leukemia are related to three major effects: (a) the replacement of normal bone marrow elements, leading to complications resulting from anemia, thrombocytopenia, and leukopenia; (b) the complications of either tissue infiltration or leukostasis owing to a marked increase in peripheral cell counts; and (c) the release of physiologic factors that may lead to significant complications, for example, disseminated intravascular coagulation. It is likely that the pancytopenia and suppression of normal elements seen in acute leukemia are the result of a physical displacement of normal bone marrow elements; it is also quite likely that unrecognized suppression factors are involved that actively suppress normal hematopoiesis. The normal bone marrow is more than a simple factory assembly line and involves complex interactions between hematopoietic

stem cells and stromal cells, endothelial cells, and other environmental factors that lead to cellular differentiation. It is likely then, that any disturbance in this normal marrow equilibrium, such as leukemia, would be associated with marrow suppression; thus, the goal of any therapy should be to decrease and eliminate the leukemic population to allow recovery of the normal bone marrow elements.

Diagnostic Modalities

The diagnosis of acute leukemia may be a multistep process involving multiple laboratory diagnostic modalities (Table 43.2). Morphology remains the most important and significant diagnostic modality despite the impressive array of technology that is available in the clinical laboratories. The morphologic examination of the peripheral blood, bone marrow aspirate, and bone marrow biopsy specimen is an important ingredient in the morphologic evaluation and diagnosis of acute leukemia.

TABLE 43.2. DIAGNOSIS OF ACUTE LEUKEMIA

Routine morphology
Peripheral blood
Bone marrow aspirate
Bone marrow biopsy
Cytochemical stains
Immunophenotyping
Flow cytometry
Immunocytochemistry
Paraffin section immunoperoxidase
Cytogenetics
Karyotyping
In situ hybridization
Molecular diagnostics
Gene rearrangement studies
Oncogenes/translocations

Cytochemical enzymatic stains have been available since the early part of this century and have provided investigators with a wealth of knowledge over the years concerning the various subtypes of acute leukemia. This old fashioned laboratory test remains quite effective in confirming the diagnosis and classification of acute leukemia, as various enzymatic products are more likely associated with particular myeloid or lymphoid cells. The particulars of the various cytochemical stains are described in subsequent sections.

Newer diagnostic modalities have arisen that provide important biological, diagnostic, and prognostic information in the acute leukemias. Immunophenotyping with monoclonal antibodies has provided a consistent method of subclassifying the acute leukemias into either myeloid or lymphoid origin, as well as delineating various subsets within each of these two broad leukemic subgroups. This is detailed in the subsequent sections. Cytogenetic analysis has also provided a wealth of knowledge concerning the biology of leukemia as well as important diagnostic and prognostic information. Cytogenetic studies have allowed us to recognize the importance of reciprocal translocations, deletions, and other chromosomal abnormalities, which have significantly contributed to our biological knowledge of leukemia. Specific cytogenetic abnormalities are associated with particular subtypes of acute leukemia and have been shown over the past decade to have significant, independent prognostic information. Another diagnostic modality available for the evaluation of acute leukemia is molecular diagnostics. These techniques allow the evaluation of the immunoglobulin and T-cell receptor genes, as well as analyses of various oncogenes involved with particular chromosomal translocations.

ACUTE LYMPHOBLASTIC LEUKEMIA

Part of "43 - Acute Leukemias and Myelodysplastic Syndromes"

Pathogenesis and Clinical Manifestations

ALL is a clonal malignancy of lymphopoietic precursor cells, arising primarily within the bone marrow. The etiology behind the development of ALL is unknown but is not thought to be associated with radiation or exposure to toxic agents. The onset of symptoms in patients with ALL is usually acute and rapidly progressive. The clinical manifestations of ALL relate primarily to the suppression of normal bone marrow elements: weakness, fatigue, and malaise owing to anemia; fever, infection, and unresponsiveness to antibiotics owing to leukopenia; and bleeding owing to thrombocytopenia. Skeletal pain as a result of marrow expansion is also not an uncommon finding, whereas lymphadenopathy and hepatosplenomegaly are common but are usually not prominent findings. Virtually any organ may be infiltrated by the leukemic cells; involvement of the leptomeninges may result in various central nervous system (CNS) symptoms, whereas the testes are a frequent source of residual disease in young boys with ALL.

Incidence

ALL is the most common cancer in children under the age of 15 years and is the second or third leading cause of death in this age group. Approximately 80% of ALL cases occur in childhood and approximately 80% of childhood acute leukemia is of lymphocytic origin. Thus, ALL is typically considered a pediatric neoplasm and is encountered much more frequently in institutions

having a sizable pediatric population. Childhood ALL has a peak incidence at approximately 4 years of age, is slightly more common in boys than in girls, and is slightly more common in whites than in nonwhites. ALL can also occur in adulthood but has a peak incidence range between the ages of 2 and 10 years.

Laboratory Findings

Although we typically equate leukemia with disorders having markedly elevated white blood counts (WBCs), the absence of leukocytosis does not eliminate the possible diagnosis of acute leukemia. ALL is a good example of the variation in presenting WBC. Approximately 40% of ALL cases will have a presenting WBC of less than 10.0×10^9/L; approximately 40% of patients will have a WBC between 10.0 and 50.0×10^9/L. Only 20% of cases will have a significant leukocytosis exceeding 50.0×10^9/L and only one half of those patients will have a marked leukocytosis greater than 100.0×10^9/L. The differential count of the WBC in patients with ALL shows predominantly lymphoblasts along with an expected neutropenia. Differentiation to mature lymphocytes is not seen. Other laboratory abnormalities that are commonly found in patients with ALL include elevated serum lactate dehydrogenase (LDH) and serum uric acid.

Morphology and FAB Classification

ALL can be classified either morphologically or immunologically. The French-American-British (FAB) morphologic classification of acute leukemia has designated subtypes FAB-L1, -L2, and -L3 (Table 43.3). The revised FAB classification of ALL is based on weighing various criteria: nuclear-to-cytoplasmic ratio, the number of nucleoli, nuclear membrane irregularity, and cell size (Table 43.4).

TABLE 43.3. FAB CLASSIFICATION OF ACUTE LYMPHOBLASTIC LEUKEMIA

Cytologic Feature	L1	L2	L3
Size	Predominant small cells	Heterogeneous, intermediate to large cells	Large cells
Cytoplasm	Scant	Variable to moderately abundant	Moderately abundant
Nucleoli	Small/inconspicuous	≥1; prominent to large	≥1; prominent and large
Nuclear chromatin	Homogeneous and intermediate reticular	Heterogeneous with some having finely reticular chromatin	Finely reticular
Nuclear shape	Regular and round	Irregular	Regular to round
Basophilic cytoplasm	Slight to none	Slight to none	Intense
Vacuolation	Slight to none	Slight to none	Prominent; sharply punched out

FAB, French-American-British.

TABLE 43.4. REVISED FAB SCORING SYSTEM FOR L1 AND L2 VARIANTS

	Score	
Criterion	**Present**	**Absent**
High N/C ratio in 75% of cells	+1	0
Low N/C ratio in 25% of cells	-1	0
Nucleoli: 0–1 (small) in 75% of cells	+1	0
Nucleoli: ≥1 (prominent) in 25% of cells	-1	0
Irregular nuclear membrane in 25% of cells	-1	0
Large cells are 50% of total	-1	0
FAB-L1	0 to +2	
FAB-L2	-1 to -4	

FAB, French-American-British; N/C, nuclear:cytoplasmic.

The distribution of FAB-L1 and -L2 differs in children and adults. Approximately 80% to 85% of cases of ALL in children are FAB-L1 compared with 35% to 40% in adult ALL. Most of the remaining cases of ALL are classified as FAB-L2. Only 1% to 3% of ALL cases in adults and children are of the FAB-L3 subtype. This latter type of leukemia is a leukemic manifestation of Burkitt's lymphoma and is discussed in a subsequent section. FAB classification was useful for prognostic evaluation in the past; however, with present-day protocols for treatment of high-risk patients, the prognostic significance of morphologic classification is diminished. In fact, a recent proposal by the World

Health Organization (WHO) utilizes both immunophenotype and cytogenetic findings for classification of these neoplasms.

In FAB-L1 ALL, most of the leukemic cells are small with little to scant cytoplasm and absent, to inconspicuous at best, nucleoli. The chromatin pattern is clearly homogeneous and is not as finely reticular as those found in other types of leukemia. Nuclear irregularity is minimal (Fig. 43.2). In the FAB-L2 ALL, the most significant difference from FAB-L1 is that the leukemic cells are larger and more heterogeneous. FAB-L2 cells typically have variable to moderately abundant amounts of cytoplasm and importantly have one or more prominent nucleoli and a finer chromatin pattern compared with FAB-L1. Irregular nuclear shapes are commonly seen (Fig. 43.3). FAB-L3 ALL is characterized by a proliferation of large, homogeneous cells having deeply basophilic cytoplasm and sharply punched-out cytoplasmic vacuoles (Fig. 43.4). The nuclear chromatin is uniform and typically coarser than is found in FAB-L1 and FAB-L2. This type of ALL is actually the peripheralized version of Burkitt's lymphoma and thus should be considered as an acute leukemia more from a historical than a biological perspective. The WBC may be normal to slightly elevated with only a low percentage of circulating leukemic “blasts.” The bone marrow typically shows only partial replacement with the leukemic cells. Because of its association with Burkitt's lymphoma, FAB-L3 ALL is typically associated with ileocecal tumor masses in the Western countries.

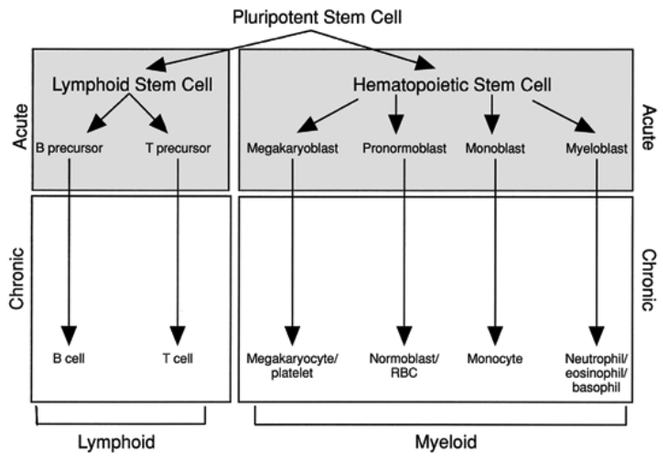


FIGURE 43.1. This schematic diagram outlines the relationship of the acute and chronic myeloid and lymphoid leukemias to normal hematopoietic and lymphoid development. The acute leukemias represent clonal expansions of immature precursor cells, whereas chronic leukemias represent proliferations of differentiated and mature lymphoid and hematopoietic elements.

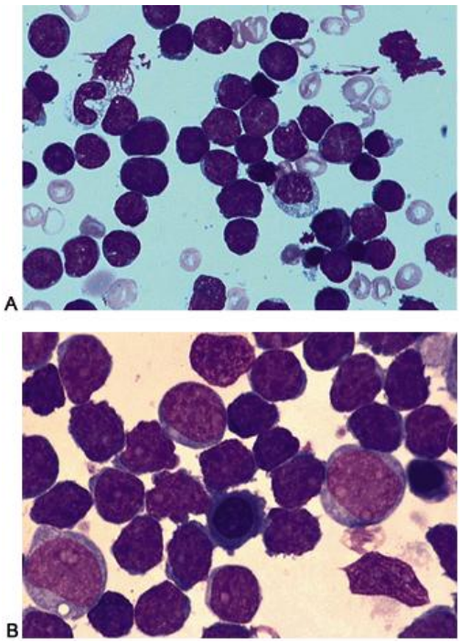


FIGURE 43.2. Acute lymphoblastic leukemia L1. **A:** Bone marrow with L1 blasts; **B:** bone marrow with primarily L1 blasts and occasional L2 blasts.

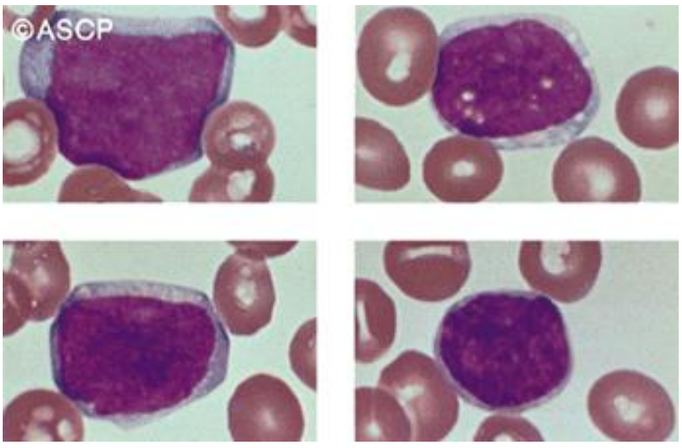


FIGURE 43.3. Acute lymphoblastic leukemia bone marrow examples of L2 blasts.

Morphologic Variants

Uncommon morphologic variants of ALL have been recognized. A “hand-mirror” cell variant of ALL has been described in which the cytoplasm loses its spherical shape and appears to have a handle. The elongated cytoplasmic handle has been described as the result of cellular motility or locomotion and has been regarded by some as a poor prognostic sign in ALL. This has not been widely accepted and is likely to be more of a laboratory artifact rather than having true biological significance.

Granular ALLs have also been described, characterized by the presence of azure granules within the cytoplasm of the leukemic lymphoblasts. These granules are reported to be negative with myeloperoxidase (MPO). This subset of ALL is not considered to have significant clinical or prognostic importance. However,

recognition of this variant is important in avoiding diagnostic problems that might arise in differentiating ALL from AML.

Shen and colleagues described a group of childhood ALL that presented with significant residual myeloid activity. These cases were referred to as “ALL with left shift.” Clinically, these patients had a longer duration of complete remission that appeared to be the result of either a smaller tumor load or lesser endogenous suppression of normal hematopoiesis. Diagnostically this may create a problem because the marrow is only partially replaced by the leukemic blasts and may be easily confused with AML unless appropriate diagnostic tests are performed.

It has been recognized for approximately 20 years that some cases of ALL may present with significant hypereosinophilia. These patients may have ALL present simultaneously with the hypereosinophilia, present initially as hypereosinophilia to be followed with a diagnosis of ALL, or present as ALL and relapse with a concurrent ALL and hypereosinophilic picture. The eosinophils may occasionally be slightly dysplastic in appearance, but no other evidence of myeloid differentiation is typically identified. Complications related to hypereosinophilia have been described in some patients, including endocarditis and respiratory complications. It has been postulated that the eosinophilia is secondary to release of eosinophilic growth factors from the leukemic cells. A t(5;14) (q31;q32) translocation has been identified in some cases of ALL with hypereosinophilia.

Cytochemistry

The role of cytochemical staining in the diagnosis of ALL is more relevant to exclude the diagnosis of AML rather than to find supportive evidence for the diagnosis of ALL (Table 43.5). Historically, the periodic acid–Schiff (PAS) reaction has classically been considered a diagnostic stain for ALL. However, only 40% to 60% of ALL cases show positivity with PAS, and PAS positivity can occasionally be found in AML. Thus, PAS lacks both sensitivity and specificity for ALL. Nonetheless, PAS positivity in ALL is characterized by large chunks or blocks of PAS-staining cytoplasmic material (Fig. 43.5). This can be distinguished from the diffuse PAS reaction seen in granulocytic cells. This enzyme stains primarily the glycogen present within the cytoplasm of leukemic blasts. Thus, PAS positivity in an acute leukemia should be looked at as suggestive, but certainly not diagnostic, of an ALL. Acid phosphatase (ACP) has been evaluated as a possible marker for T-cell ALL. T lymphocytes and most T lymphoblasts show focal, punctate, perinuclear positivity with ACP. However, not all T-cell ALLs show this pattern of staining, and a significant number of B-precursor ALL show positivity with ACP, thus limiting its utility as a stain to differentiate between T-cell and non-T-cell ALL. Myeloblasts may also show a focal staining with ACP. Oil red O, which stains lipid material, is an excellent marker for ALL-L3 or Burkitt's leukemia/lymphoma. The oil red O stain distinctly stains the vacuoles that are seen in this subtype of leukemia/lymphoma.

TABLE 43.5. CYTOCHEMICAL STAINING IN ACUTE LYMPHOBLASTIC LEUKEMIA

Stain	L1	L2	L3
MPO	-	-	-
SBB	-	-	-
CAE	-	-	-
NSE	-[a]	-[a]	-
PAS	+(70%)	+(70%)	-
ACP	-/+[b]	-/+[b]	-
MGP	-	-	+
ORO	+	+	+
TdT	+	+	-

[a] Faint positivity may be seen.

[b] Most T-cell acute lymphoblastic leukemia and some B-precursor acute lymphoblastic leukemia, will be positive.

MPO, myeloperoxidase; SBB, Sudan black B; CAE, chloroacetate esterase; NSE, nonspecific esterase; PAS, periodic acid–Schiff; ACP, acid phosphatase; MGP, methyl green pyronine; ORO, oil red O; TdT, terminal deoxynucleotidyl transferase.

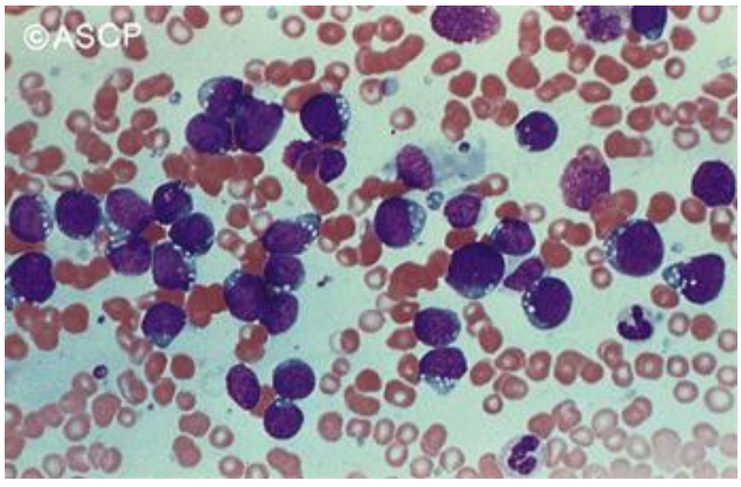


FIGURE 43.4. Acute lymphoblastic leukemia L3-Burkitt's leukemia/lymphoma.

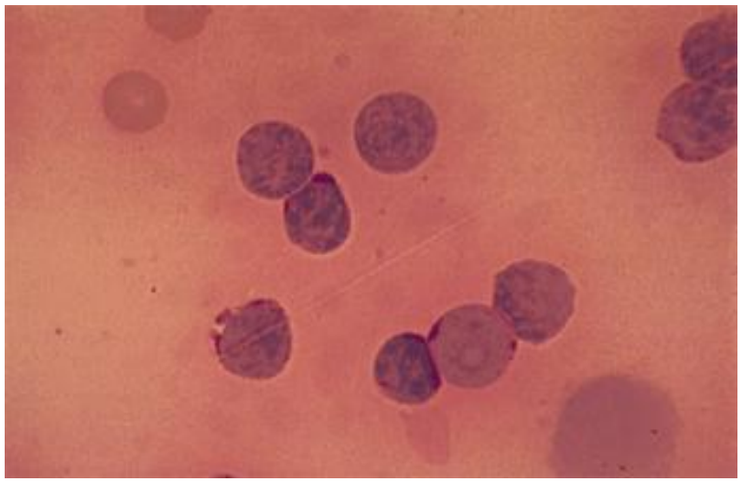


FIGURE 43.5. Periodic acid–Schiff stain of acute lymphoblastic leukemia L1.

Terminal deoxynucleotidyltransferase (TdT) is a DNA polymerase that contributes to the recombination heterogeneity seen in immunoglobulin and T-cell receptor gene rearrangements. This enzyme is active during lymphoblast development and is expressed in 95% to 99% of ALL cases. ALL-L3 lacks this enzyme, as it biologically represents a maturing lymphocyte rather than a true leukemic blast. Unfortunately, 5% to 10% of AMLs also express this nuclear antigen, thus limiting the absolute lineage specificity of the TdT assay. TdT may be detected by either

immunofluorescence, immunoperoxidase, or enzyme immunoassay methods (Fig. 43.6). TdT can be identified in ALLs of either T- or B-precursor immunophenotypes.

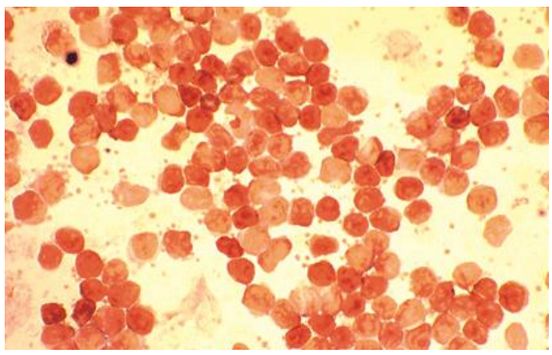


FIGURE 43.6. Terminal deoxynucleotidyltransferase immunoperoxidase stain of a B-precursor acute lymphoblastic leukemia.

Immunophenotype

The most consistent and effective means of classifying ALL is based on immunophenotyping data (Table 43.6). Approximately 80% to 85% of ALLs can be classified as malignant counterparts of bone marrow B-precursor cells. Ten percent to 15% of ALL will be identified as T-cell ALL, thus having immunologic characteristics of immature T cells or thymocytes. The remaining 1% to 3% of ALLs will immunologically represent mature B cells having surface immunoglobulin and correspond to the FAB-L3 subgroup or Burkitt's leukemia/lymphoma.

TABLE 43.6. THE ROLE OF IMMUNOPHENOTYPING IN ACUTE LEUKEMIA

Distinction between ALL and AML
Identification of B-precursor ALL, T-cell ALL, and B-cell ALL subgroups
Diagnosis of acute megakaryoblastic leukemia (FAB-M7)
Recognition of acute biphenotypic leukemia

ALL, acute lymphoblastic leukemia; AML, acute myelogenous leukemia; FAB-M7, French-American-British subclassification of acute myelogenous leukemia.

Historical Terminology

The terminology used in subclassifying ALLs has at best been confusing for those involved with diagnostic hematology (Table 43.7). This growth in confusion has paralleled our increased sophistication in immunophenotyping and represents an evolution of knowledge concerning these leukemias.

TABLE 43.7. TERMINOLOGY USED FOR ACUTE LYMPHOBLASTIC LEUKEMIA

B-Precursor ALL[a]
T-Cell ALL[a]
B-Cell ALL[a]
Non-T, non-B-cell ALL
Non-T-cell ALL
Common ALL
CALLA-positive ALL
Pre-B-cell ALL
Pre-pre-B-cell ALL
Pro-B-cell ALL

[a] Currently preferred terminology.

ALL, acute lymphoblastic leukemia; CALLA, common acute lymphoblastic leukemia antigen.

CD10 (CALLA)

The breakthrough in subclassifying ALLs came with the development of a polyclonal heteroantiserum against a case of non-T, non-B-cell ALL; this antibody was called anti-common ALL antigen, or anti-CALLA (CD10). Patients with CALLA-positive ALL were found to have a much better prognosis than those with either B- or T-cell ALL. This development was considered a major breakthrough for the role of immunophenotyping in evaluating acute leukemias. These CALLA-positive ALLs made up the majority, but not all, of the non-T-, non-B-cell ALLs. Because of their reactivity with the CALLA antiserum (and subsequently with a monoclonal antibody against CALLA), these cases were called “common” ALLs. After this development, several B lymphocyte–associated differentiation antigens were recognized by a variety of monoclonal antibodies. This monoclonal antibody explosion led to a rapidly increasing knowledge of lymphoid development and the realization that virtually all the CALLA-positive, non-T-, non-B-cell ALLs were indeed of early B-cell lineage. Although other terminologies have been and are used to describe immunologic subgroups of ALL, we refer to those leukemias reactive with B cell–associated antibodies as B-precursor ALL.

B-Precursor ALL

A tremendous diversity of immunophenotypes can be identified in B-precursor ALL. Practical experience with immunophenotyping finds that the diagnosis of ALL is actually quite simple, to the relief of the diagnostician. However, it should also be recognized that there is only limited correlation between the FAB subtypes and immunologic results (Fig. 43.7).

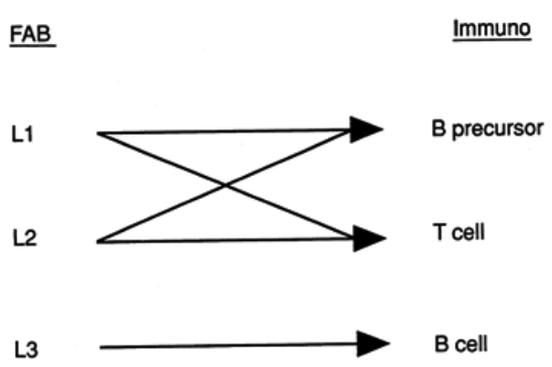


FIGURE 43.7. Correlation of FAB and immunophenotype in acute lymphoblastic leukemia.

The complexity of the various immunologic subgroups recognized within the B-precursor ALLs, however, has increased as the number of monoclonal antibodies to immature B cells has grown. It has been postulated that the different stages of leukemic expression are representative of stages of normal pre–B-cell development and that, by inference, leukemias arise by clonal proliferation of cells arrested at a particular stage of normal maturation.

The CD19 antigen is recognized as an early pan–B-cell marker and is present in more than 95% of ALL cases. Intracytoplasmic localization of the CD22 antigen has also been found at an early stage of normal B-cell differentiation; surface CD22 does not appear until later in immature B-cell development. The expression of CALLA or CD10 antigen occurs at the next level of development, followed by CD24. The fluorescence intensity of CD10 diminishes as the B-cell precursor undergoes maturation and CD20 expression begins.

CD10 is thought to be lost before surface CD22 expression. As can be gleaned from Fig. 43.8 and Fig. 43.9, most ALL cases will show expression of HLA-DR, CD19, and CD10 with either CD34 or CD20 typically being found. Some important exceptions and subgroups need to be recognized and are discussed further. Lack of CD10 expression appeared to be associated with a poorer prognosis.

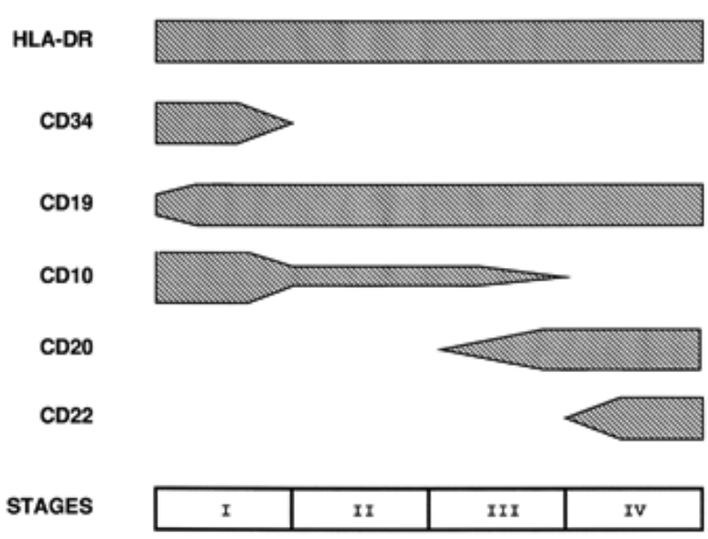


FIGURE 43.8. Proposed immunophenotypic stages of B-precursor lymphoid development. The narrower bars indicated decreased fluorescent intensity expression of that particular antigen (for example, CID 10).

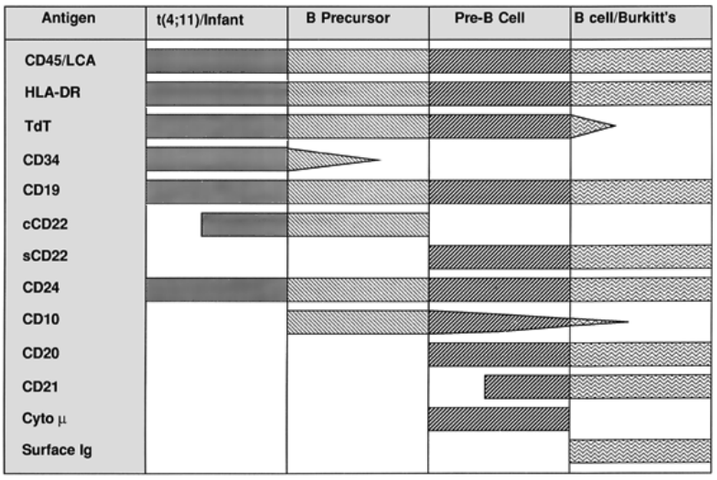


FIGURE 43.9. Immunophenotypes of B-precursor and B-cell acute lymphoblastic leukemia (ALL). The various antigens are listed in the left column. Four types of B-precursor and B-cell ALL are listed in the respective columns. cCD22 and sCD22, cytoplasmic and surface expression, respectively; Cyto μ, cytoplasmic expression of the immunoglobulin μ heavy chain.

The earliest stages of recognizable B-cell development are characterized by positivity with HLA-DR, TdT, CD34, and CD19. This early B-lymphoid immunophenotype (with or without CD7) has been commonly associated with leukemias having translocations involving 11q23, in particular the t(4;11) translocation (Fig. 43.9). These t(4;11) leukemias commonly present in infancy, have a marked leukocytosis, and have an extremely short survival without response to chemotherapy. In addition, many of these leukemias will have a monocytic component at the time of diagnosis or relapse and are frequently associated with biphenotypic features. All these findings suggest that 11q23 leukemias arise very early in B-lymphocyte differentiation. Although earlier immunologic stages of B-cell differentiation are theoretically possible, characterized by HLA-DR, TdT, CD34, and CD7 expression and perhaps cytoplasmic expression of B-lineage antigens, only rare leukemias have been definitively identified with these findings.

In the latter stages of B-cell development, CD34 expression is absent, CD10 has diminished in intensity, and CD20 and/or CD22 can be identified for the first time. Cytoplasmic μ heavy chain is detectable at this latter stage and typifies the so-called pre–B-cell ALL (Fig. 43.8 and Fig. 43.9). No intact surface immunoglobulin can be identified. The terminology pre-B cell is used as an indication that these cells have evidence of B-cell commitment, i.e., cytoplasmic immunoglobulin heavy chain, but without an intact immunoglobulin molecule being identified. This subgroup constitutes 10% to 15% of the B-precursor ALL group and is commonly associated with a t(1;19) translocation.

The CD34 antigen also needs to be discussed in relation to the B-precursor ALL group. As mentioned earlier, CD34 is a hematopoietic progenitor cell antigen present very early in B-cell development. As the mature B cell goes through maturation and undergoes immunoglobulin heavy chain rearrangement, CD34 expression is lost. Most reports have described CD34 positivity in 50% to 60% of cases of B-precursor ALL. CD34 expression may also provide important prognostic information in ALL.

B-Cell ALL: Burkitt's Leukemia

The final and most differentiated type of ALL has detectable surface immunoglobulin. This corresponds to the original B-cell ALL group and morphologically represents ALL, FAB-L3, or what has been called Burkitt's leukemia/lymphoma; TdT reactivity is usually absent in such neoplasms. B-cell ALL will show positivity with CD21 in addition to the previously described B cell–associated monoclonal antibodies: HLA-DR, CD19, CD20, CD22, and CD24 (Fig. 43.9). CD10 expression is found in one third to one half of B-cell ALL cases and is of much weaker intensity than that found in the typical B-precursor ALL.

T-Cell ALL

T-cell ALL (T-ALL) accounts for 10% to 15% of all ALL cases. It is usually associated with a marked leukocytosis; a mediastinal mass is found in more than 50% of cases; male patients are favored;

and median age is approximately 14 years. These patients have a high incidence of CNS relapse and a poorer survival than patients with B-precursor ALL. This survival difference has led some to conclude that immunophenotype may provide prognostic information. However, multivariate analysis has not shown that a T-cell phenotype is independent of WBC and age, both of which typically represent high risk factors in T-ALL.

The morphologic features of T-ALL are not uniquely specific for this subtype, but some findings need to be noted. The T-ALL cases are more likely to have a FAB-L1 appearance, having scant cytoplasm and relatively dense chromatin. A high mitotic rate is typically found and nuclear convolution may be evident. This latter finding, however, is neither a sensitive nor a specific feature of T-ALL and cannot be reliably used to differentiate this immunophenotypic subgroup from B-precursor ALL.

The anti-CD7, CD5, and CD2 monoclonal antibodies are the most sensitive markers that react with the leukemic blasts of T-ALL. In general, these three antigens are expressed in virtually all cases of genuine T-ALL (Fig. 43.10). Reactivity with other T-cell monoclonal antibodies in T-ALL in general correspond to intrathymic patterns of expression. The CD3 antigen has generally been reported in less than one third of cases of T-ALL and T-lymphoblastic lymphoma, as evaluated by flow cytometric techniques. In contrast, several studies observed CD3 expression in 95% of T-ALL/lymphoblastic lymphoma cases utilizing cryostat sections. This difference in results is related to the cytoplasmic detection of CD3 in cryostat sections. Several studies have now shown that cytoplasmic CD3 is present much earlier in thymic differentiation before the surface expression of CD3 and before the expression and rearrangement of the T-cell receptor gene. These data confirm that CD3 positivity is not relegated to the final stages of thymocyte development and can be a useful marker of early T-cell differentiation if a cytoplasmic method of detection is utilized. Cytoplasmic CD3 must therefore be considered as an essential antigen to be looked for if other markers of T-cell differentiation are equivocal.

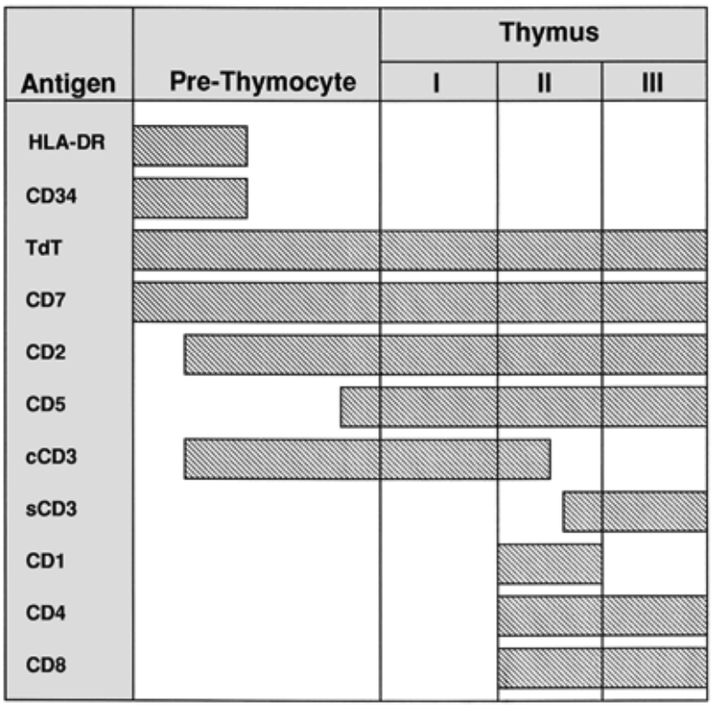


FIGURE 43.10. Correlation of immunophenotype and thymic stage of development in T-cell ALL. The respective antigens are listed in the left column and the various stages of thymic development in the rest of the table. cCD3 and sCD3, cytoplasmic and surface expression of the CD3 molecule, respectively.

Although CD7 has been promulgated as being the earliest marker of T-cell differentiation, one must exercise great caution in using this marker in isolation as evidence of T-cell differentiation. As is discussed later, 5% to 10% of AML cases show expression with this marker. Likewise, the sole expression of CD2 without any other T cell-associated marker should lead one to exercise caution before diagnosing the case as definite T-ALL because this marker can be identified in as many as 5% of AML cases. Thus, if one identifies only one of the pan–T-cell markers listed above in a case of acute leukemia, it is essential that a wide battery of both T-cell markers, especially cytoplasmic CD3, and myeloid markers, including megakaryoblast-associated glycoproteins, be used to confidently distinguish between a T-cell and a myeloid process. Obviously, morphologic, cytochemical, and clinical findings are essential in confirming this distinction. Although CD1, surface CD3, CD4, and CD8 provide interesting information regarding the corresponding stage of thymic development of T-ALL, not much is gained from a diagnostic viewpoint, as these antigens are almost always expressed in conjunction with CD2, CD5, and CD7.

Other antigens not typically associated with T cells can be identified in T-ALL. Essentially all cases express the common leukocyte antigen CD45. CD10 (anti-CALLA) positivity is usually found in approximately 10% to 20% of cases of T-ALL.

HLA-DR expression has also been found in 10% to 40% of T-ALL. Leukemic blasts in this disorder are also strongly positive for TdT. CD34, a hematopoietic progenitor cell antigen, can be found in 10% to 20% of all T-ALL cases. This is in sharp contrast to 50% to 60% positivity found in B-precursor ALL. CD11b, CD11c, and CD15 are all antigens that are more typically associated with myeloid processes but can also be found in 20% to 60% of T-ALL cases. Thus, one cannot use these antigens as strict evidence of a myeloid component.

Another interesting finding (but currently of uncertain clinical significance) is the recognition of different subtypes of T-ALL based on the type of T-cell antigen receptor present on the leukemic blasts. More than 95% of peripheral T-cell lymphocytes express the α/ß heterodimer T-cell receptor, with the remainder expressing the γ/δ T-cell receptor. A higher percentage of T cells within epithelial locations express the γ/δ receptor than the α/ß receptors. Within the thymus, γ/δ rearrangement is thought to occur before the α/ß rearrangement process. Several investigators have now reported T-ALLs that express the γ/δ receptor. These have been identified primarily based on the expression of surface CD3 without CD4 or CD8. Although the status of the antigen receptor T-ALL has provided us with important biological information, no clinical or prognostic difference has been found between T-ALL with α/ß receptor and T-ALL with γ/δ receptor. It is uncertain whether the identification of this subgroup has any clinical or prognostic importance.

Lymphoblastic Lymphoma

Malignant lymphoma, lymphoblastic type, is a high-grade lymphoma that occurs predominantly in older children and young adults; males are more frequently affected, and a mediastinal mass is often found. This neoplasm is often morphologically characterized by convoluted nuclei with fine chromatin and a high mitotic rate. Both precursor T-cell and B-cell ALL may present primarily as extramedullary tumors. The distinction between these disorders is usually based on the presentation of disease: tissue presentation generally indicates lymphoma, whereas marrow or blood presentation indicates leukemia. Most lymphoblastic lymphomas are T cell in origin and are reactive with multiple T cell–associated antibodies. This lymphoma can be distinguished from other types of lymphoma based on TdT positivity.

Cytogenetics

Cytogenetic analysis may provide the most important prognostic information for childhood ALL. Several studies have now shown that by using high-resolution banding techniques, more than 90% of ALL patients had demonstrable clonal chromosomal abnormalities. Large international studies have shown that children and adults with ALL can be prognostically stratified according to the type of chromosomal abnormalities. Hyperdiploidy (more than 50 chromosomes) has been consistently shown to be a favorable prognostic feature and is an independent prognostic predictor of therapeutic response. In contrast, three particular translocations in ALL have been associated with very poor prognoses, short survivals, and poor responses to induction chemotherapy. These include the Philadelphia chromosome or t(9;22), t(4;11), and t(8;14). The Philadelphia chromosome occurs in approximately 5% of childhood ALL and as many as 20% to 25% of adult ALL. Some of these may represent the blast crisis

phase of chronic myelogenous leukemia, although molecular studies clearly indicate that many of these are *de novo* presentations of acute leukemia. ALL patients with the t(9;22) translocation typically are older, have marked leukocytosis, and have a B-precursor immunophenotype.

The t(4;11) translocation or any ALL involving 11q23 has been reported in as many as 5% of cases of ALL. These cases typically have a FAB-L2 morphology, may frequently be associated with biphenotypia, and may occasionally show lineage switch to a monoblastic component (Fig. 43.11). The t(4;11) leukemias are most commonly identified in infant ALL (less than 2 years of age), usually present with marked leukocytosis, and have a median survival of less than 1 year. Splenomegaly is commonly identified. The immunophenotype of this subgroup of ALL shows HLA-DR, CD34, and CD19 positivity; CD10 and CD20 are typically negative.

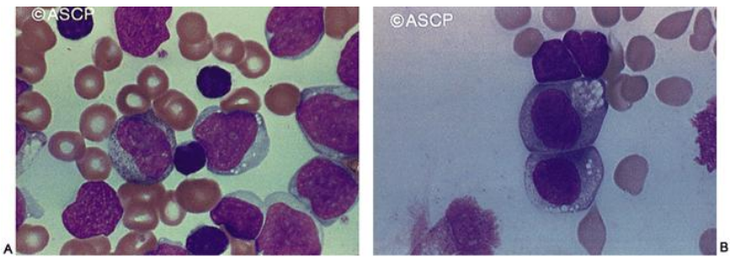


FIGURE 43.11. An acute leukemia with a t(4;11) translocation demonstrating both lymphoid (L2) **(a)** and monocytic **(B)** features.

The t(12;21)(p13;q22) translocation was recognized with advent of molecular based testing. Approximately 25% of childhood ALL is associated with this “cryptic” chromosome rearrangement. It is usually not recognized by routine cytogenetic studies but easily identified by fluorescent *in situ* hybridization and polymerase chain reaction. One of the most significant implications of this finding is its association with a clinically favorable prognosis.

The t(8;14) translocation and its variants, t(2;8) and t(8;22), are always found in the B-ALL (surface immunoglobulin-positive) and FAB-L3 morphology subtype, i.e., the Burkitt's leukemias/lymphomas. These patients have short survivals and are characterized by surface immunoglobulin expression.

The t(1;19)(q23;p13), E2A-PBX fusion, is found in 25% of ALL patients with cytoplasmic Ig positivity. It was considered to have a poor prognosis in the past because it was resistant to antimetabolite based regimens. However, recent protocols for high-risk ALL appear to improve treatment response.

Several clinical, hematologic, and laboratory criteria have been used to determine prognosis in patients with ALL (Table 43.8). Clinical, hematologic, and cytogenetic findings are essential components for predicting therapeutic outcomes of patients with ALL. Major determinants for the prognosis of ALL include the initial WBC, the degree of mediastinal involvement, and age. Children from 2 to 8 years of age do better than those younger or older, and infants less than 1 year of age have a particularly poor prognosis. In children between the ages of 2 and 8 years having WBCs less than 10.0 × 10^9/L, a 90% cure rate can be achieved. Patients of any age presenting with WBCs greater than 50.0 × 10^9/L have the shortest survivals, and those with intermediate WBCs show intermediate survivals. The previously discussed cytogenetic factors have also been shown to be prognostically independent of age and WBC and clearly add to the overall prognostic database.

TABLE 43.8. PROGNOSTIC GROUPS OF CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Good prognosis (20%)	WBC <10.0 × 10^9/L 2–8 yr of age
Average prognosis (60%)	WBC <10.0 × 10^9/L <2 or 8 yr of age or WBC 10.0–50.0 × 10^9/L Any age
Poor prognosis (20%)	WBC 50.0 × 10^9/L

WBC, white blood cells.

ACUTE MYELOGENOUS LEUKEMIA

Part of "43 - Acute Leukemias and Myelodysplastic Syndromes"

Pathogenesis

The AMLs are a heterogeneous group of malignancies originating in the hematopoietic, or myeloid, stem cell. The neoplastic proliferation seen in AML consists of myeloblasts or partially differentiated myeloid cells. This failure to show complete maturation by the neoplastic clone results in an accumulation of immature precursor cells and gradual replacement of normal bone marrow elements. To restate an earlier contention, AML is not a disorder of rapidly proliferating cells but rather an accumulation of incompetent, long-surviving cells. The various subgroups or variants of AML basically reflect maturation arrest at various points in hematopoietic/myeloid differentiation. Thus, referring to Fig. 43.1, one can rationalize the existence of the various morphologic subtypes of AML based on whether the granulocytic,

monocytic, erythroid, or megakaryocytic arm of myeloid differentiation is involved.

The exact cause of this leukemia is obviously unknown. However, various risk factors for the development of AML are known, including chemical exposures, such as to benzene; alkylating-agent chemotherapy; ionizing radiation; preceding myelodysplastic syndromes (MDSs); aplastic anemia; paroxysmal nocturnal hemoglobinuria; syndromes with chromosome instability, such as Fanconi's anemia, ataxia-telangiectasia, and Bloom's syndrome; Down's syndrome; and rare forms of familial inheritance.

Clinical Manifestations

The most common clinical findings of AML at presentation are very similar to those described in patients with ALL. Symptoms are primarily related to increasing suppression of normal bone marrow elements leading to anemia, thrombocytopenia, and neutropenia. Splenomegaly is mild to moderate. Lymphadenopathy can be found in as many as one third of AML patients but is not as uniform or pronounced as in ALL. Thymic involvement is exceedingly uncommon in AML. Cutaneous infiltration, however, is relatively common, occurring in as many as 10% of AML patients. This and other forms of extramedullary involvement, such as the gingival tissue, are most commonly associated with those leukemias having a monocytic component. The terms chloroma, granulocytic sarcoma, myeloblastoma, and extramedullary myeloid tumor all refer to the collection of leukemic cells in an extramedullary organ. These extramedullary tumors may occur in soft tissues, the orbital region, ovaries, testes, gastrointestinal tract, breast, skin, and other sites. CNS involvement in AML is not as common as in ALL but must still be considered as a possible complication of disease.

Incidence

AML is primarily a disease of adults. It constitutes approximately 45% of all leukemias but more than 80% of all adult acute leukemias and less than 20% of childhood acute leukemia.

Laboratory Findings

In AML, the WBC is elevated in more than half of patients at the time of diagnosis. Pancytopenia is a constant feature of this leukemia. The number of blasts in the peripheral blood varies according to the subtype of AML and depends on the degree of maturation and differentiation that characterizes that particular leukemic subtype. As many as 10% of patients will have WBC counts that are greater than 100.0×10^9/L, with some patients approaching 500.0×10^9/L. Patients with WBCs greater than 50.0 to 100.0×10^9/L are at risk of developing complications owing to hyperleukocytosis. Leukophoresis is mandated in these patients and is an essential part of the therapeutic process. Serum uric acid level is typically elevated in approximately two thirds of patients with AML. Serum LDH level is typically not higher than in ALL.

Morphology and FAB Classification

The French-American-British (FAB) group on acute leukemia has morphologically subclassified the AMLs into seven major subtypes: FAB-M1 to FAB-M7 (Table 43.9). The FAB classification of AML basically relies on the degree of granulocytic, monocytic, erythroid, and megakaryocytic differentiation. This is based on the morphologic appearance of cells, the number of leukemic blasts, and cytochemical findings with myeloid-associated enzymatic stains. Important variants of these seven subtypes also exist. The National Cancer Institute has proposed revised standards for the classification of AML, including another subtype based on immunophenotyping (FAB-M0). The various subtypes of AML are listed in Table 43.9. The incidences of the AML subgroups are listed in Table 43.10.

TABLE 43.9. FAB CLASSIFICATION OF ACUTE MYELOGENOUS LEUKEMIA

M0	Acute leukemia, undifferentiated; myeloid immunophenotype
M1	Acute myelogenous leukemia (AML) without maturation
M2	AML with maturation
M3	Acute promyelocytic leukemia (APL)
M3v	Hypogranular variant of APL
M4	Acute myelomonocytic leukemia
M4e	M4 with eosinophilia
M5a	Acute monoblastic leukemia
M5b	Acute monocytic leukemia
M6	Acute erythroleukemia
M7	Acute megakaryoblastic leukemia

FAB, French-American-British.

TABLE 43.10. INCIDENCE OF ACUTE MYELOGENOUS LEUKEMIA ACCORDING TO FAB SUBTYPE

FAB Subtype	Incidence (%)
M0	1–2
M1	18
M2	28
M3	8
M4	27
M5	10
M6	4
M7	5

FAB, French-American-British classification.

The importance of the FAB classification of AML lies not in its ability to provide important prognostic information to the patient but rather its ability to provide precise and consistent classification of AML subtypes. The ability to reproduce the various FAB subgroups morphologically and cytochemically is well proven and allows knowledge interaction and consistency between institutions and physicians.

The basic definition of an AML is a leukemia that has more than 30% myeloblasts in the bone marrow. However, in the recent WHO classification, the required percentage of blasts for diagnosis of AML is decreased to 20% in the bone marrow and peripheral blood. Myeloblasts may be grouped into three different subtypes: I, II, and III. Type I myeloblasts are undifferentiated blasts that have no evidence of granulocytic differentiation, i.e., they contain no cytoplasmic granulation or Auer rods (Fig. 43.12). Auer rods are absolutely specific for a leukemic myelogenous process and consist of abnormally fused primary granules. Auer rods are needle- to fusiformlike eosinophilic rods that are found in the cytoplasm of leukemic myeloid cells and stain red to purple with Wright-Giemsa stain.

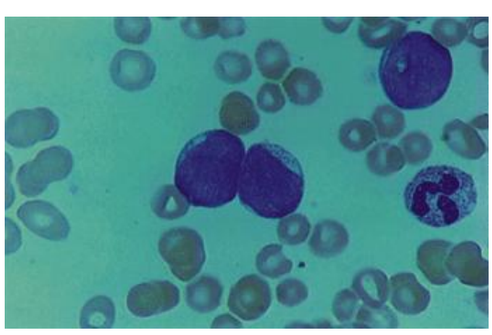


FIGURE 43.12. Type I (left cell) and II (right cell) myeloblasts.

Type II blasts are leukemic blasts that have relatively few cytoplasmic granules (Fig. 43.12). It is too rigid to specify the exact number of these granules; however, most morphologists would define type II blasts as having fewer than eight to 12 granules per cell. The type III blasts are cells intermediate between type II blasts and promyelocytes. Obviously, this distinction may be quite difficult to make at times, but it is generally relegated to cells that have the nuclear characteristics of blasts (fine chromatin and prominent nucleoli) but with more granules than in the type II blasts. The cytoplasm retains its basophilic appearance.

The National Cancer Institute's expanded definition of blasts is listed in Table 43.11 and includes the type I, II, and III blasts discussed previously. Also included in their basic definition of blasts are the abnormal promyelocytes seen in FAB-M3, the monoblasts and promonocytes of FAB-M4 and FAB-M5, and the megakaryoblasts of FAB-M7. Erythroblasts, including early pronormoblasts, are not included in the definition of a leukemic blast. Specific criteria for each subtype is summarized in Table 43.12

TABLE 43.11. EXPANDED BLAST DEFINITION: NATIONAL CANCER INSTITUTE

Type I blast: no granules
Type II blast: few granules
Type III blast: multiple granules, central nucleolus, fine chromatin
Abnormal promyelocytes of M3
Monoblasts/promonocytes of M5
Megakaryoblasts of M7
Erythroblasts should *not* be included

TABLE 43.12. CLASSIFICATION OF ACUTE MYELOGENOUS LEUKEMIA

FAB Subtype	Diagnostic Criteria
M0	No morphologic evidence of differentiation
	Negative cytochemical staining with MPO/SBB/NSE
	Positivity with myeloid markers (CD13, CD33, etc.) and negative lymphoid markers
M1	≥90% blasts in bone marrow
	<10% of marrow showing granulocytic differentiation
	>3% of blasts with MPO or SBB positivity
M2	≥20% blasts and <90% blasts in marrow
	≥10% of marrow showing granulocytic differentiation
	≤20% monocytic cells
M3	≥20% of marrow or abnormal promyelocytes
M3v	Same as M3 except composed of hypogranular variants
M4	≥20% blasts
	≤80% myeloblasts and granulocytic precursors in marrow
	20% monocytic cells in marrow (morphology/NSE stain/serum lysozyme)
M4e	Same as M4 with increased number of atypical/immature marrow eosinophils
M5a	≥80% monoblasts
M5b	≥80% of monoblasts/promonocytes/monocytes
	<80% monoblasts in marrow
M6	≥50% nucleated RBC
	Prominent dyserythropoiesis
	≥20% myeloblasts in nonerythroid cells
M7	≥20% blasts
	Identification of megakaryoblasts by ultrastructural cytochemistry or by immunophenotyping

MPO, myeloperoxidase; SBB, Sudanblack b; NSE, nonspecific esterase; RBC, red blood cells.

FAB-M0: AML Without Maturation

The FAB-M0 group of AML is myeloid leukemia with a minimal differentiation that has a demonstrable myeloid lineage by immunophenotyping or ultrastructural studies. These AMLs consist of small- to intermediate-size leukemic blasts without any evidence of granulocytic differentiation (Fig. 43.13). Cytochemical staining with MPO, Sudan black B (SBB), and nonspecific esterase (NSE) is negative. Thus, from a strict morphologic and cytochemical analysis, these cases would have been classified as ALL in the past. However, by flow cytometric immunophenotyping, these cases show reactivity with one or more myeloid-associated markers, such as CD117 (C-kit) CD13, CD33, or CD15, while lacking any expression of lymphoid-associated markers.

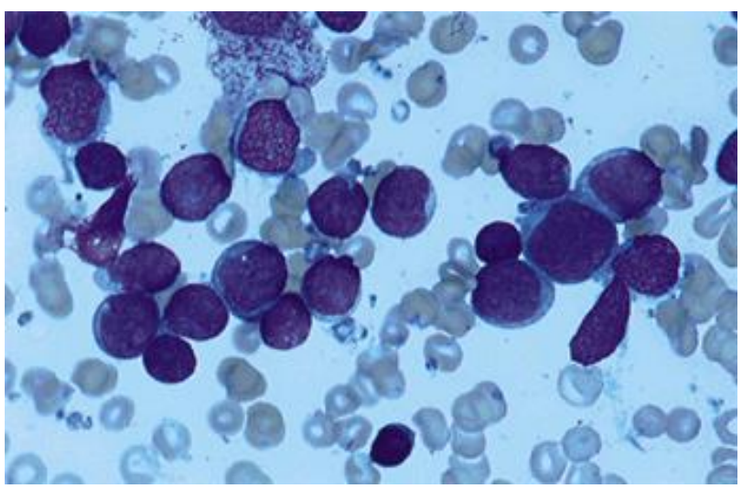


FIGURE 43.13. Acute myelogenous leukemia M0.

FAB-M1: AML Without Maturation

In these AMLs without maturation, there is minimal evidence of cytoplasmic granulation and minimal numbers of Auer rods. The blasts are intermediate in size with finely reticular chromatin, small amounts of grayish-blue cytoplasm, and typically one or more prominent nucleoli (Fig. 43.14). Basic FAB criteria

for AML-M1 include the following: (a) the sum of total blasts is greater than 90% of the bone marrow cells; (b) less than 10% of bone marrow cells shows evidence of granulocytic differentiation at or beyond the promyelocyte stage; and (c) at least 3% of the leukemic blasts demonstrate MPO and/or SBB positivity. The differential diagnosis in this disorder includes ALL-L2, AML-M5a, and AML-M7. Immunophenotyping and cytochemical staining are usually necessary to make the distinction between these subtypes of leukemias.

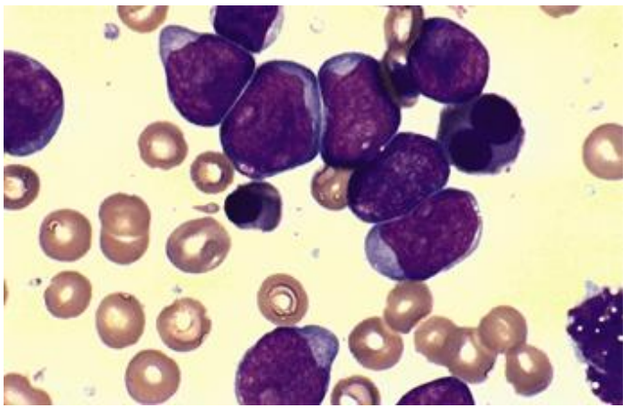


FIGURE 43.14. Acute myelogenous leukemia M1.

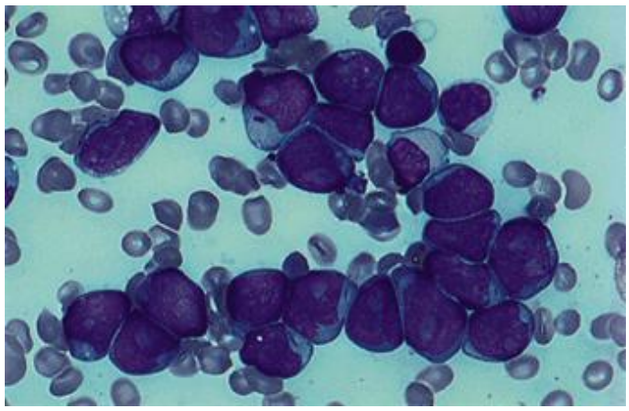


FIGURE 43.15. Acute myelogenous leukemia M2 with numerous single Auer rods.

FAB-M2: AML with Maturation

The AMLs with maturation are the most common subtype of AML. These leukemias show clear evidence of differentiation at or beyond the promyelocyte stage, and Auer rods are commonly identified (Fig. 43.15). These cells usually have more cytoplasm than is found in FAB-M1 leukemic cells, and fewer undifferentiated blasts are seen. From a diagnostic viewpoint, this subtype of AML is easy to diagnose and creates little diagnostic consternation. The basic FAB criteria for FAB-M2 include the following: (a) the sum of blasts is 20% or greater but less than 90% of the bone marrow cells; (b) more than 10% of the bone marrow cells shows evidence of granulocytic differentiation; and (c) monocytic cells constitute fewer than 20% of the bone marrow cells. It should also be noted that dysplastic features may be identified in the granulocytic, erythroid, and/or megakaryocytic series. Frank panmyelosis is uncommon but can be seen in some cases. Although most bone marrow specimens are hypercellular and show frank replacement of all bone marrow elements, some patients, mainly the elderly, may have bone marrows that are moderately hypocellular.

The differential diagnosis of FAB-M2 would include a leukemoid reaction, a MDS that does not meet the criteria of AML, and possibly other types of AML having a granulocytic component such as FAB-M3, FAB-M4, or possibly FAB-M6.

A specific bone marrow chromosome abnormality, t(8;21), has been observed in some cases of FAB-M2. The percentage of patients with the t(8;21) translocation has varied from 30% to less than 5% of FAB-M2s in various reports in the literature. Patients having AML-M2 with a t(8;21) are believed to have a good prognosis. The morphologic features of FAB-M2 with a t(8;21) include larger myeloblasts, easily identified and numerous Auer rods, and a distinctive dysmyelopoiesis in the developing granulocytic cells. This dysmyelopoiesis has been described as a “crushed” orange granularity in the cytoplasm of the granulocytic cells. These patients usually aberrantly express CD 19 and/or CD56.

FAB-M3: Acute Promyelocytic Leukemia

Acute promyelocytic leukemia (APL) can be diagnosed when 20% or more of the bone marrow cells are abnormal promyelocytes. These promyelocytes have abnormally dense and heavy granulation. One of the characteristic features of APL are the so-called faggot cells, which are cells that contain multiple Auer rods that may be bundled, intertwined, or fused together (Fig. 43.16 and Fig. 43.17). The granules of the promyelocyte are larger and darker staining than normal and at times may be so numerous as to obscure nuclear borders. Intensely basophilic cytoplasm may be present in some cells. The nuclear features, which are frequently ignored or even obscured by the granulation, typically have a monocytoid, bilobed, or kidney-bean shape. It is uncommon for this type of AML to have a significant percentage of blasts. In a large series of APLs, the median blast count was only 12%. These patients are typically leukopenic at presentation and are clinically characterized by disseminated intravascular coagulation (DIC) and bleeding.

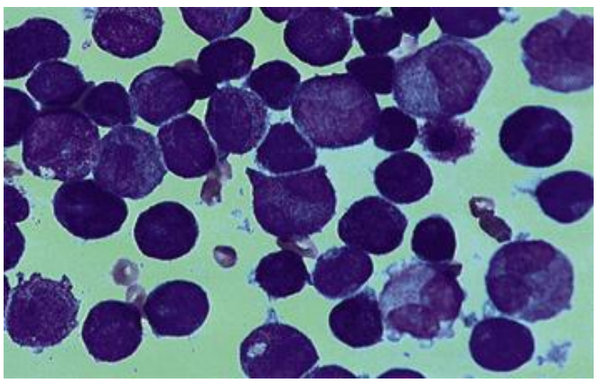


FIGURE 43.16. Acute myelogenous leukemia M3 (acute promyelocytic leukemia).

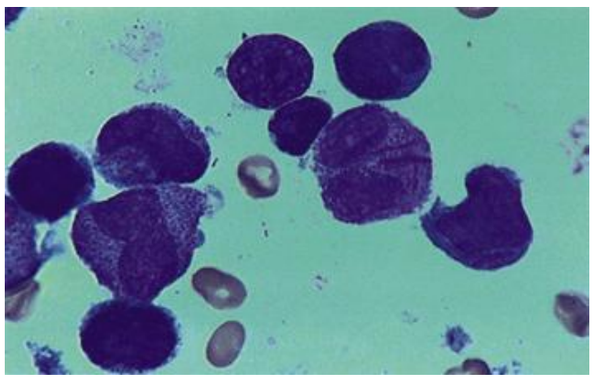


FIGURE 43.17. Acute myelogenous leukemia M3 (acute promyelocytic leukemia Faggot cell).

Approximately 20% of APLs are a variant form of FAB-M3

and are designated as microgranular or hypogranular APL (M3v). The leukemic cells in this FAB-M3 variant are characterized by sparse and/or fine granulation and a strikingly irregular nuclear shape (Fig. 43.18). Their identity as abnormal promyelocytes may be obscured by the scarcity of granulation and the nuclear shape. Cells containing multiple Auer rods are usually present, but they may be extremely difficult to identify because they are certainly less abundant than in the typical hypergranular FAB-M3. The variant form of APL has the same incidence of DIC described previously but will more likely present with leukocytosis at the time of diagnosis. The obvious differential diagnostic problem relates to the confusion with acute myelomonocytic (M4) or acute monocytic (M5b) leukemia.

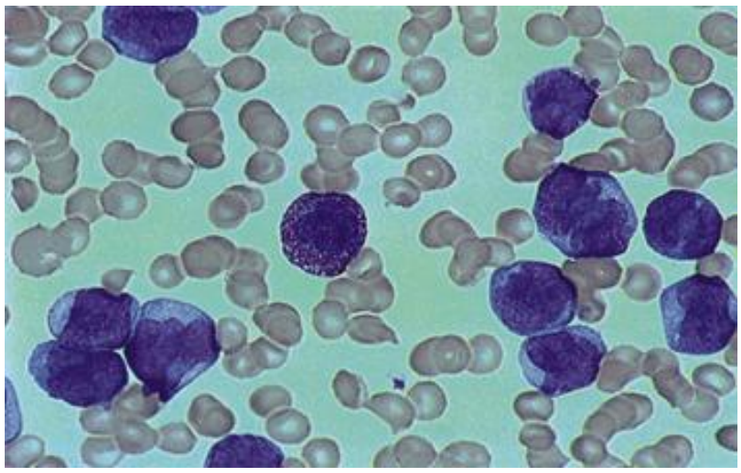


FIGURE 43.18. Acute myelogenous leukemia M3 variant (hypogranular acute promyelocytic leukemia).

APL may occur at any age but is most common in young adults, with a median age of diagnosis of 35 to 40 years. The most outstanding clinical feature associated with APL is the high frequency of DIC. This is owing to the release of procoagulant factors from the cytoplasmic granules, which leads to activation of the coagulation cascade. In many patients, there are severe DIC and hemorrhage before or during induction of therapy, when the malignant cell contents are lysed and released. Hemorrhage is the cause of death in many patients. Thus, it is essential that this subgroup of AML be accurately diagnosed rapidly so that appropriate chemotherapy and supportive care can be initiated. If the DIC and hemorrhage are adequately controlled and treated, patients with this subgroup of AML probably have the best prognostic outcome of any AML group. Recognition of this particular subgroup is especially important because a specific regimen including all-*trans* -retinoic acid (ATRA) causes leukemic cells to overcome maturation arrest.

The differential diagnosis of APL includes other types of AML with granulocytic components, such as FAB-M2, FAB-M4/M5, and benign agranulocytosis with a promyelocyte arrest. In cases of benign agranulocytosis, the platelet count and hemoglobin are generally normal, in contrast to the pancytopenia seen in APL. In addition, in agranulocytosis, the bone marrow is not hypercellular and Auer rods are obviously not identified. The major difficulty in the diagnosis of APL is the distinction of microgranular APL (M3v) from FAB-M4 or M5b. Evaluation with cytogenetic and molecular studies typically resolves this issue. Demonstration of t(15;17)(q22;q12-21) or t(11;17) is diagnostic of APL and essential for diagnosis. The immunophenotypic studies in typical APL include myeloid phenotype (CD13, CD33) and lack of HLADR, CD34.

FAB-M4: Acute Myelomonocytic Leukemia

Acute myelomonocytic leukemia and FAB-M2 are the most common AMLs, together accounting for approximately two thirds of all AML cases. The FAB-M4 subgroup of AML is probably the most difficult subgroup of AML in which to perform a reliable differential count, as this is a very heterogeneous-appearing leukemia. Both granulocytic and monocytic differentiation are present in varying proportions in the bone marrow (Fig. 43.19). It may be difficult to identify the granulocytic and monocytic component, as hybrid cells clearly exist in this type of leukemia. The criteria for the diagnosis of FAB-M4 includes the following elements: (a) the sum of all blasts is greater than 20%; (b) the sum of myeloblasts and granulocytic precursor cells account for less than 80% of the bone marrow cells; (c) more than 20% of the bone marrow cells are of the monocytic lineage as demonstrated by morphology, NSE cytochemical stain, or elevated serum lysozyme level (three times normal). Thus, granulocytic and monocytic precursors coexist in proportions varying reciprocally from 20% to 80%. Auer rods may also be identified in approximately half of the cases. Because of the difficulty in the morphologic identification of both promonocytes and hybrid

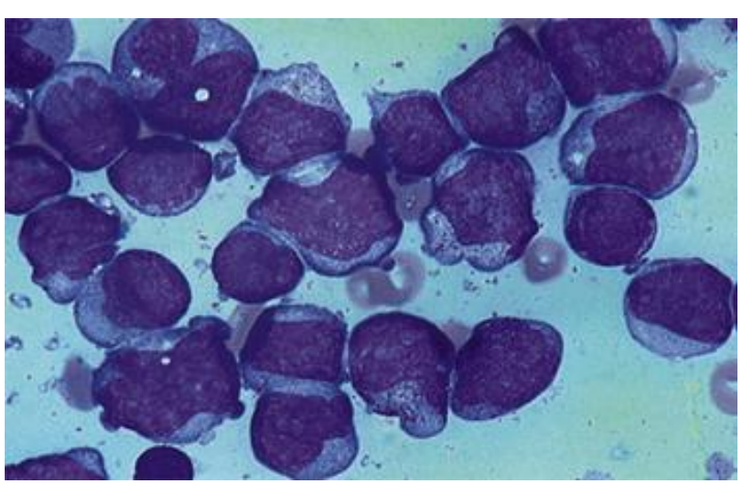


FIGURE 43.19. Acute myelogenous leukemia M4.

granulocytic/monocytic cells in bone marrow, additional diagnostic criteria utilizing NSE and elevated serum lysozyme levels have been used. Organomegaly, lymphadenopathy, and other sites of tissue infiltration may commonly be encountered.

A variant of FAB-M4 exists, which has been called FAB-M4 with eosinophilia (M4e). Criteria for diagnosis include the usual diagnostic criteria for FAB-M4 and an increased number of atypical, immature bone marrow eosinophils (Fig. 43.20). The immature and atypical eosinophils contain an abundance of large basophilic-staining granules in addition to the large, red eosinophilic granules that characterize a mature eosinophil. Abnormalities of chromosome 16, including inversion of 16(p13;q22) or deletion of 16q22, are consistently identified in this variant. This type of leukemia has a high rate of remission after the initial induction of therapy compared with other types of AMLs, and diagnosis of this variant is considered a good prognostic sign.

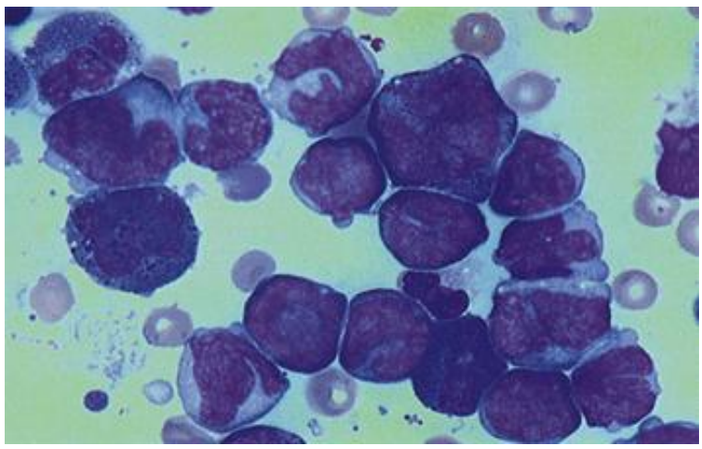


FIGURE 43.20. Acute myelogenous leukemia M4 with eosinophilia.

Morbidity owing to an increased incidence of CNS relapse with leptomeningeal infiltration and pulmonary involvement, however, may occur in this subtype of AML.

FAB-M5a: Acute Monoblastic Leukemia and FAB-M5b: Acute Monocytic Leukemia

FAB-M5a is characterized by a predominance of monoblasts that are large and have relatively abundant cytoplasm. Some azurophilic granules may be identified in the cytoplasm that are MPO negative. The nucleus is typically round to oval with finely reticular chromatin and prominent nucleoli. Typically, the nucleus is displaced to one side with an ample amount of cytoplasm wrapping around the nucleus (Fig. 43.21). The sole criterion for diagnosis of FAB-M5a is the existence of more than 80% monoblasts in the bone marrow. This differs slightly from the criteria for FAB-M5b, in which more than 80% of the marrow cells are monoblasts, promonocytes, or monocytes but less than 80% of the marrow cells are monoblasts. In other words, FAB-M5b shows more differentiation than FAB-M5a. Nuclear folding and irregularity are common in FAB-M5b, and more azurophilic granulation is identified than in FAB-M5a (Fig. 43.22). Both FAB-M5a and FAB-M5b are associated with a high incidence of extramedullary infiltration; DIC may also develop in patients with FAB-M5, second in incidence only to FAB-M3 among classes of AML. NSE cytochemical stains are positive in the FAB-M5 leukemias. Auer rods may be seen in a small percentage of monoblasts but are certainly much less frequent than in the granulocytic types of AML. AML-M5a is more commonly diagnosed in the pediatric age group.

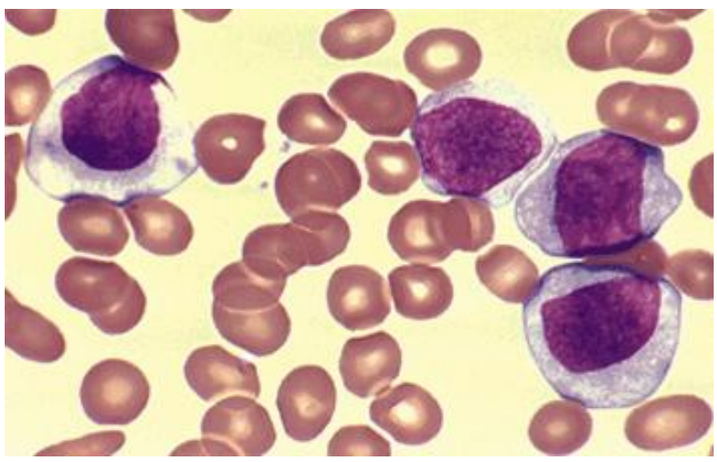


FIGURE 43.21. Acute myelogenous leukemia M5a.

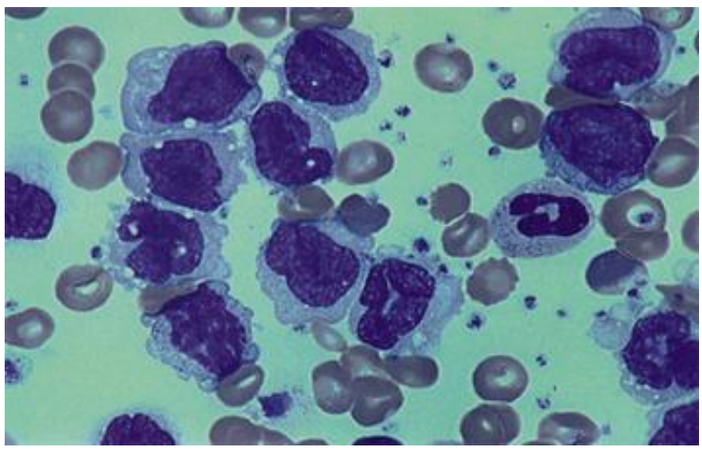


FIGURE 43.22. Acute myelogenous leukemia M5b.

FAB-M6: Acute Erythroleukemia (Di Guglielmo's Syndrome)

Acute erythroleukemia is a relatively uncommon variant of AML and may have multiple presenting appearances. One form,

which has previously been called erythemic myelosis, is characterized by bizarre and markedly atypical megablastoid changes accompanied by extreme erythroid hyperplasia within the bone marrow (Fig. 43.23). Few, if any, myeloblasts can be identified. Normal granulocytic and megakaryocytic precursors are not identified and are replaced by giant, multinucleated, and markedly dysplastic erythroblasts. In other cases of FAB-M6, the marrow contains more differentiated, albeit dysplastic, erythro-blasts at the time of presentation along with a definite population of granulocytic cells, including myeloblasts (Fig. 43.24). In 1985, the following criteria were established by the FAB group for the diagnosis of erythroleukemia in determining blast percentage: (a) 50% or more of all nucleated bone marrow cells must be erythroblasts; (b) dyserythropoiesis is prominent; (c) 20% or more of the nonerythroid cells in the bone marrow are myeloblasts.

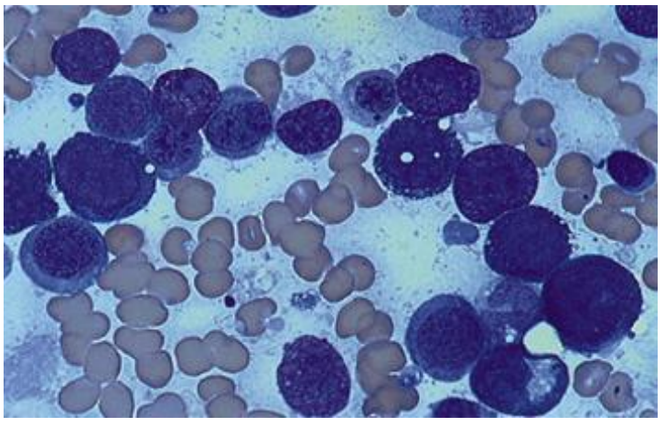


FIGURE 43.23. Acute myelogenous leukemia M6 (erythemic myelosis).

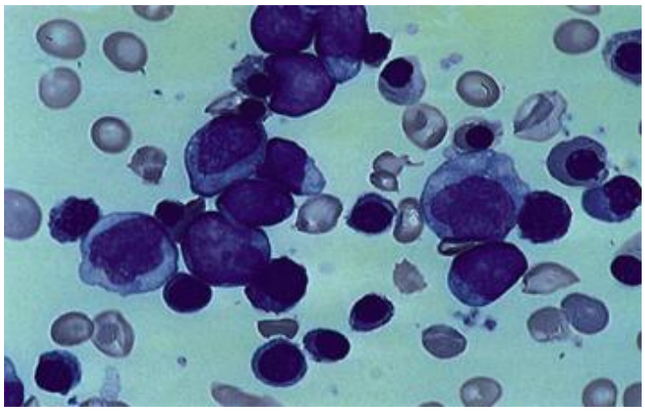


FIGURE 43.24. Acute myelogenous leukemia M6.

Erythroleukemia will frequently evolve into other types of AML such as FAB-M1, M2, or M4. Auer rods may be identified in the myeloblasts in these cases. The leukemic erythroblasts frequently contain a “frothy” or “foamy” vacuolation in the cytoplasm. Progression of the disease is frequently marked by an increase in myeloblasts and a decrease in erythroblasts. A striking erythroblastemia may be identified in the peripheral blood. It is quite common for erythroleukemia to evolve from MDSs or as a secondary leukemia in patients who have received prior radiation and/or alkylating-agent chemotherapy. The differential diagnosis for erythroleukemia includes B_{12}/folate deficiency, heavy metal intoxication (such as arsenic), drug effects (such as with antineoplastic agents or chloramphenicol), congenital dyserythropoietic syndromes, MDSs, and potentially other types of AML. The dysplastic erythroblasts are typically PAS positive, which reflects a cytoplasmic maturation defect. This cytochemical finding is not restricted to leukemia and can be identified in benign disorders, such as ß thalassemia, iron deficiency, sideroblastic anemia, and heavy metal intoxication.

FAB-M7: Acute Megakaryoblastic Leukemia

Acute megakaryoblastic leukemia has only recently been added to the FAB classification. The diagnostic criteria for diagnosis includes (a) more than 20% blasts in the bone marrow and (b) definitive identification of megakaryoblastic involvement by a platelet peroxidase reaction by electron microscopy or reactivity with megakaryocyte-specific monoclonal antibodies. The differential diagnosis of megakaryoblastic leukemia includes ALL-L2, AML-M0, AMLM1, and AML-M5a. As can be seen from this differential diagnostic list, there is a morphologic heterogeneity to FAB-M7. The blasts may vary from small- to medium-sized to large, bizarre blasts typically having high, nuclear:cytoplasmic ratios (Fig. 43.25). The nuclear chromatin may be dense and homogeneous or fine and reticular. There is typically scanty basophilic cytoplasm, which may or may not be vacuolated. Indeed, the degree of basophilia may be comparable with that found in erythroleukemia, reflecting the close developmental relationship between erythroid and megakaryoblastic precursors. An irregular cytoplasmic border may be noted in some of the megakaryoblasts resembling pseudopods. Very fine granulation can be identified in the cytoplasm of some blasts. Intermediate forms between undifferentiated blasts and definitive micromegakaryocytes may also be seen (Fig. 43.26). Some cases may show little if any differentiation and resemble the lymphoblasts of FAB-L2.

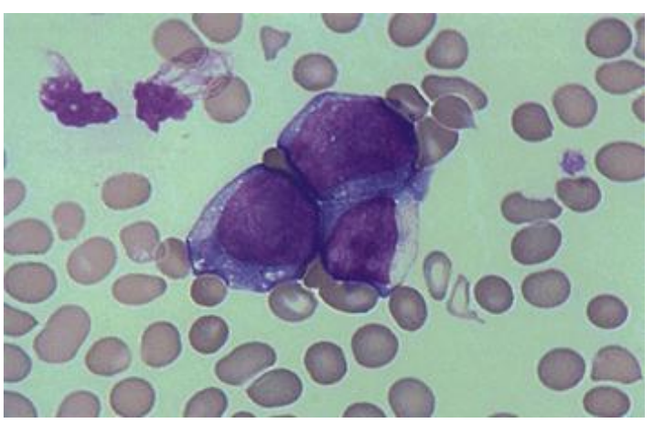


FIGURE 43.25. Acute myelogenous leukemia M7.

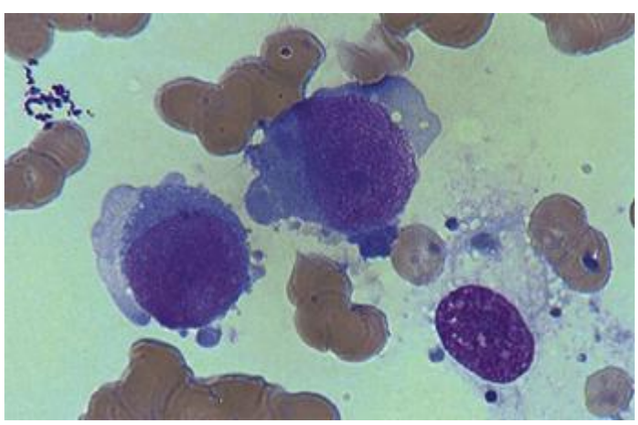


FIGURE 43.26. Acute myelogenous leukemia M7.

The diagnosis of FAB-M7 relies on multiple criteria including peripheral blood and bone marrow aspirate smear morphology, bone marrow trephine section histology, routine cytochemical stains, immunophenotyping, and ultrastructural studies (Table 43.13). In some cases, the peripheral blood and bone marrow aspirate morphology may be sufficient to make the diagnosis,

whereas in other cases, more sophisticated diagnostic modalities may be needed to confirm the diagnosis of FAB-M7. Some cases of FAB-M7 have morphologic features suggestive of a megakaryoblastic process, including the presence of circulating micromegakaryocytes, atypical platelets, or the presence of myelofibrosis and atypical megakaryocytes in the bone marrow sections (Fig. 43.27).

TABLE 43.13. DIAGNOSIS OF ACUTE MEGAKARYOBLASTIC LEUKEMIA

Aspirate morphology
Trephine section histology
Routine cytochemistry
Immunophenotyping
Ultrastructure

Cytochemically, megakaryoblasts show no reactivity with MPO or SBB. A diffuse, coarse granular positivity is typically seen with PAS, not to be confused with the chunky staining seen in ALL (Fig. 43.28). Megakaryoblasts show no reactivity with α-naphthyl butyrate esterase but can manifest strong reactivity with the acetate substrate of NSE (Fig. 43.29). This latter discrepancy between the butyrate and acetate substrates of NSE can be a very useful diagnostic feature of FAB-M7. Immunophenotyping with monoclonal antibodies reactive with platelet glycoprotein (Gp) IIb/IIIa or Gp IIIa (CD41 and CD61) has provided a more sensitive and reproducible method of detecting megakaryoblasts. These antibodies are discussed in more detail in the subsequent section on immunophenotyping.

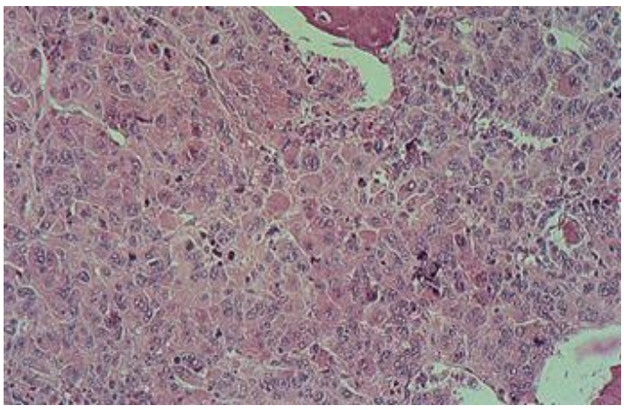


FIGURE 43.27. Acute myelogenous leukemia M7 (bone marrow biopsy).

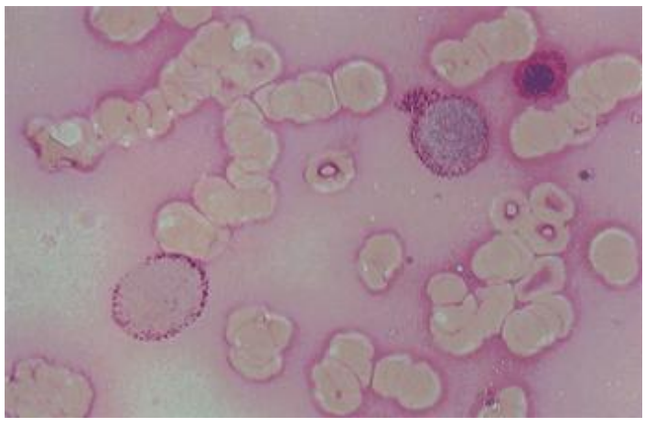


FIGURE 43.28. Acute myelogenous leukemia M7: periodic acid–Schiff cytochemical stain.

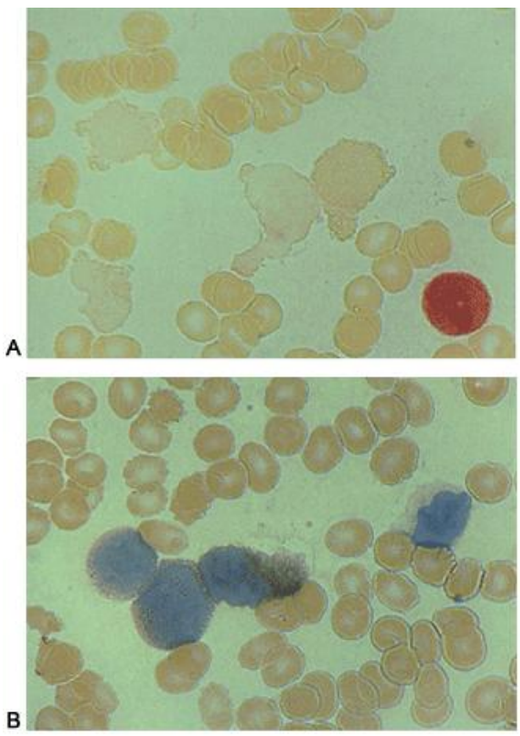


FIGURE 43.29. Acute myelogenous leukemia M7 nonspecific esterase stain. **A:** butyrate; **B:** acetate substrates.

The standard for the diagnosis of FAB-M7 is the demonstration of CD41/CD61 or platelet peroxidase by ultracytochemistry with electron microscopy. Ultrastructural peroxidase activity is found in the nuclear envelope and endoplasmic reticulum but is absent from the granules and Golgi of the leukemic megakaryoblasts. This pattern is distinctly opposite that found in myeloblasts by ultracytochemistry.

Acute megakaryoblastic leukemia is commonly associated with patients that have Down's syndrome, therapy-related acute leukemias, and blast crises of chronic myeloproliferative disorders. Clinical features of FAB-M7 are variable and can occur in both children and adults. A marked leukocytosis is relatively uncommon in M7 and typically is not associated with extramedullary involvement, lymphadenopathy, or hepatosplenomegaly.

Acute Myeloid Leukemia, Not Otherwise Specified

Some of the cases of AML not fitting in the above groups are placed in this category. Rare cases of AML with basophilic differentiation as commonly demonstrated by ultrastructural studies

are called acute basophilic leukemia. The basophilic differentiation is usually detected by toluidine blue or electron microscopy. Extramedullary myeloid tumors, so-called granulocytic sarcoma, are also included in this category. The third group is acute panmyelosis with myelofibrosis, which presents with marked pancytopenia and panmyeloid proliferation. Because they usually show an increased number of atypical megakaryocytes, differentiation from AML-M7 may be difficult in some cases. In fact, some consider this entity as AML-M7 with marked myelofibrosis. The lack of splenomegaly and prominent hematopoietic dysplasia help to distinguish this disorder from chronic myeloproliferative disorders.

Cytochemistry

Table 43.14 shows the distribution of cytochemical staining in AML. In general, it can be noted that the AMLs with a granulocytic component will stain with MPO, SBB, and chloroacetate esterase (CAE) (Fig. 43.30 and Fig. 43.31), and those with a monocytic component will react with NSE and ACP (Fig. 43.31 and Fig. 43.32). Megakaryoblastic staining was previously discussed.

TABLE 43.14. CYTOCHEMICAL STAINING IN ACUTE MYELOGENOUS LEUKEMIA

Stain	M1	M2	M3	M4	M5	M6	M7
MPO	+	+	+	+	-	+/-	-
SBB	+	+	+	+	-	+/-	-
CAE	+/-	+	+	+	+/-	+/-	-
NSE-b	-	-	-	+	+	-	-
NSE-a	-	-	-	+	+	-	+
ACP	-	-	-	+	+	-	+
PAS	-/+	-/+	-	-	+/-	+	+
TdT[a]	-	-	-	-	-	-	-

[a] Between 10% and 20% of AML will be TdT positive.

MPO, myeloperoxidase; SBB, Sudan black B; CAE, chloroacetate esterase; NSE-b, nonspecific esterase butyrate; NSE-a, nonspecific esterase acetate; ACP, acid phosphatase; PAS, periodic acid–Schiff; TdT, terminal deoxynucleotidyltransferase.

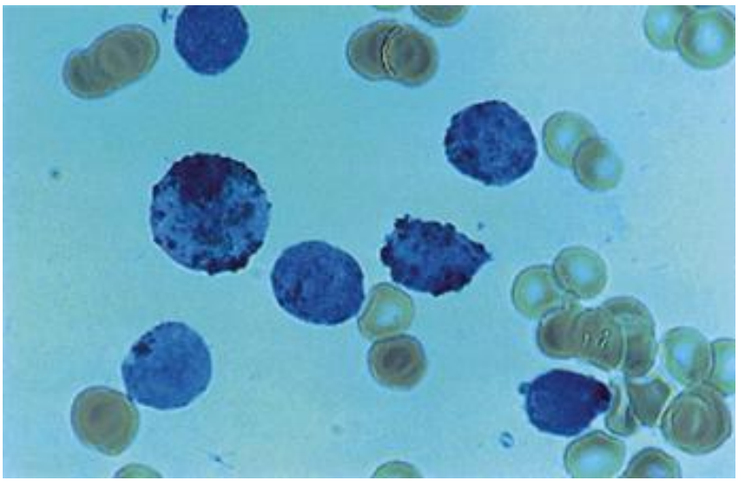


FIGURE 43.30. Myeloperoxidase stain of acute myelogenous leukemia M2.

MPO is an enzyme present within the primary granules of the granulocytic series, including granulocytes, eosinophils, and some basophils. Monocytes also contain peroxidase granules but are present in very small numbers and never pose a serious problem when interpreting such a stain. Lymphoid and erythroid cells do not stain with MPO. We have found MPO to be the most sensitive marker for granulocytic differentiation; however, MPO stains will fade with time; for this reason SBB is also commonly used and is a good alternative to MPO. SBB stains intracellular lipids that are within the primary granules of the granulocyte series. This is easily detected within myeloblasts and the rest of the granulocyte series. Lymphoblasts, however, may also, show an occasional rare, small positive granule. Thus, one must interpret SBB with some caution. CAE, or specific esterase/Leder stain, is confined to the primary granules of differentiated granulocytes. Myeloblasts are typically lacking in this enzyme. Monocytes, lymphocytes, and erythroid cells are negative with this enzyme stain. CAE is not as sensitive as MPO or SBB.

NSE is a particularly useful stain for the monocyte series. Little, if any, staining of the granulocytic series will be found. T lymphocytes and T lymphoblasts typically show a distinct punctate, perinuclear positivity that is easily distinguished from monocyte staining. The butyrate esterase gives a clearer and more definitive stain compared with the acetate substrate of NSE and is the preferred monocyte-associated enzyme stain. It gives a deep, brick-red stain that is easy to interpret and distinguish

from background staining, as opposed to the faint positivity seen with the acetate substrate of NSE. The acetate NSE activity can also be found in megakaryocytes, platelets, and some basophils and plasma cells. Fluoride sensitivity is used by some laboratories to help differentiate between different NSE-positive cells. The NSE reactivity found in monocytes is blocked by fluoride and is partially blocked in megakaryocytes. NSE staining found in lymphocytes and lymphoblasts, however, is not sensitive to fluoride treatment. Occasional cases of B-precursor ALL also show a faint blush in the Golgi area of such lymphoblasts; this is also fluoride resistant.

As mentioned in the ALL section, PAS may stain some cases of AML, thus limiting its diagnostic use as a specific ALL marker. TdT reactivity can also be found in as many as 20% of AML cases, although it is usually dimmer and positive in a lower percentage of cells than in typical ALL.

Immunophenotype of AML

The diagnosis of AML is usually a straightforward diagnosis made based on morphology and cytochemical staining. Immunophenotyping of potential cases of AML should be done for the following reasons (Table 43.6): (a) to distinguish AML from ALL (the most practical reason to analyze a potential AML) and (b) to provide correlation with the FAB morphologic subtype of AML.

The two situations that commonly pose the most difficult problem for diagnosing an AML are leukemias with no or minimal evidence of differentiation (FAB-M0 and FAB-M1) and those of megakaryoblastic differentiation (FAB-M7). Pan-myeloid monoclonal antibodies, CD13, CD33, and CD117, have been universally used to distinguish AML from ALL on an immunologic basis. Overall, 90% to 98% of AMLs will react with these markers. Moreover, these reagents have appeared to be highly specific for myeloid cells and are expressed in all FAB morphologic subtypes of AML. Some FAB-M0 and FAB-M1 AMLs tend to express CD33 only, suggesting that those leukemias are of earlier myeloid development than other AML. In addition, acute megakaryoblastic leukemias (FAB-M7) may not express either CD13 or CD33 or tend to express CD13 and/or CD33 in a lower percentage of blasts with weaker intensity than in more classic AML.

The use of a panel of myeloid-associated monoclonal antibodies is necessary to access any possible correlations between an immunophenotype and the FAB morphologic classification (Fig. 43.33). Reports in the literature have demonstrated that myeloid monoclonal antibody reactivity may correspond to broad categories of morphologic differentiation: myeloblastic (FAB-M1 and M2), promyelocytic (FAB-M3), monocytic (FAB-M4 and M5), erythroid (FAB-M6), and megakaryocytic (FAB-M7). FAB-M1 and M2 acute myeloid leukemias react with the pan-myeloid markers CD13, CD33, and CD117. CD15 also reacts with most cases of FAB-M1 and M2. Monocyte-associated

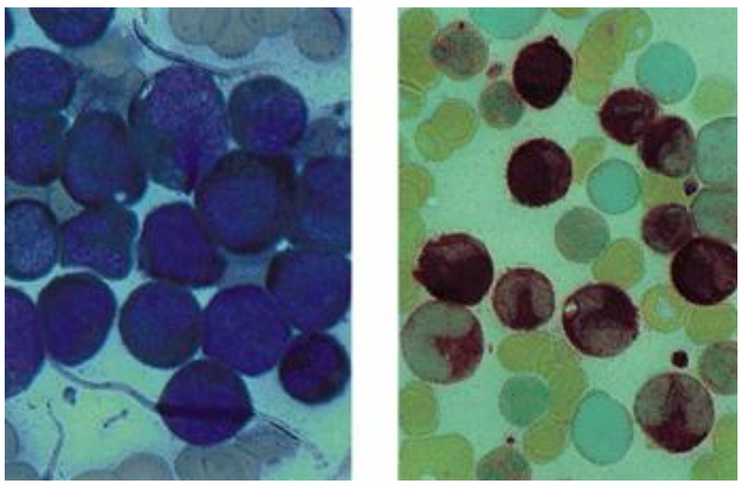


FIGURE 43.31. Sudan black B **(left)** and nonspecific esterase (butyrate) **(right)** stain of acute myelogenous leukemia M4.

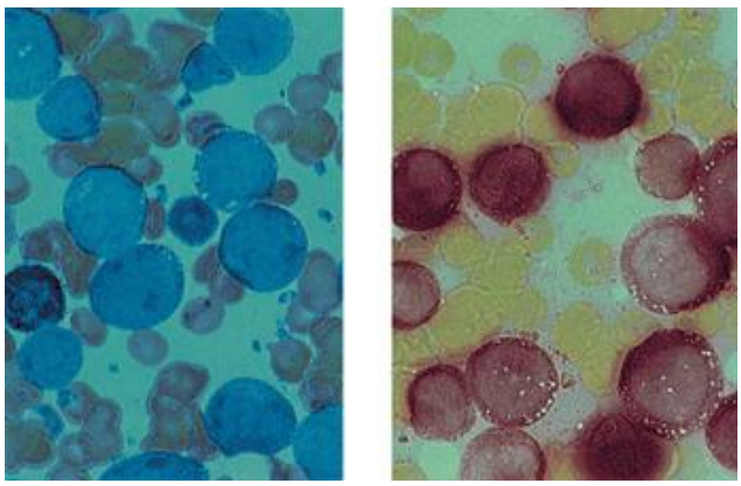


FIGURE 43.32. Nonspecific esterase (butyrate) stain of acute myelogenous leukemia M5a.

monoclonal antibodies (CD11b, CD14, CD36, and CD64) are usually nonreactive in cases of AML M1 and M2; CD14, however, has been detected in as many as 15% of such AML cases. Monoclonal antibodies more restricted to the mature stages of granulocytic differentiation, such as CD10, CD16, and CD24, do not react with the myeloblasts in these leukemias.

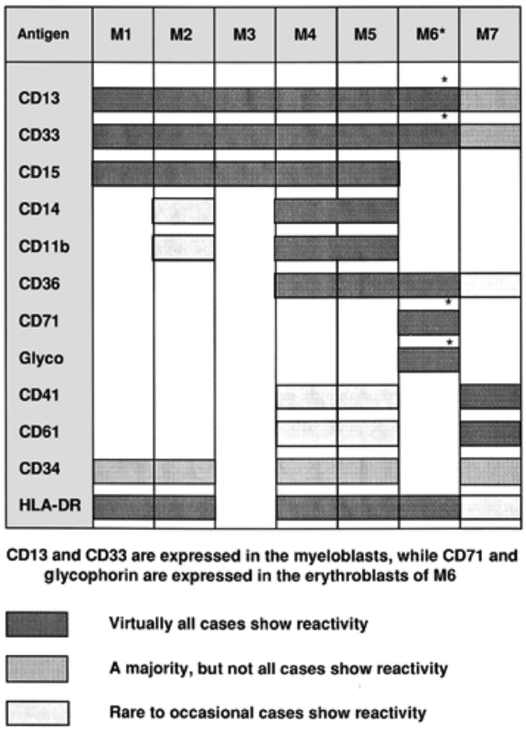


FIGURE 43.33. Correlation of FAB (French-American-British classification) subtype of acute myelogenous leukemia with immunophenotype.

HLA-DR expression is a common finding in AML. This reflects the normal distribution of HLA-DR or myeloid precursor cells up to, but not including, the promyelocyte stage. The lack of reactivity of HLA-DR antibodies in most cases of APL is consistent with the observation that the normal promyelocyte is HLA-DR negative. However, occasional genuine cases of APL may be HLA-DR antigen positive. The hypogranular variant of APL has a pattern of reactivity similar to that of the typical granular form. HLA-DR negativity is not exclusive to APL because as many as 10% of AMLs (FAB-M1 and M2) may not express the HLA-DR antigen. Hence, although HLA-DR negativity is common in APL, it does not appear to be a 100% sensitive and specific diagnostic feature of this unique subtype of AML.

The monocytic leukemias (FAB-M4 and M5) are also reactive with the pan-myeloid monoclonal antibodies, CD13, CD15, and CD33, and uniformly express the HLA-DR antigen. CD11b, CD11c, CD14, CD36, and CD64 antigens are also displayed by the majority of acute myelomonocytic and monocytic leukemias. CD14 expression is thought to be acquired at some point after the development of the monoblast and before maturation to the monocyte stage. This suggests that CD14 reactivity would be expected to be most prominent in the FAB-M4 and FAB-M5b leukemias. However, CD14 is commonly found in almost all cases of monoblastic leukemias (FAB-M5a). Interestingly, CD14 has recently been identified in a significant number of B-cell, non-Hodgkin's lymphomas, confirming that this antigen is not myeloid restricted.

No specific and sensitive marker of early erythroid development is currently available for the diagnosis of erythroleukemia (FAB-M6). The transferrin receptor CD71 is reactive with some myeloblasts in addition to all erythroblasts in the erythroleukemias. However, labeling of blasts with CD71 can also be observed in other types of AML and may be detected in some cases of ALL as well. Antiglycophorin antibodies appear to react only with cells that are morphologically obvious as erythroblasts, beginning at the basophilic normoblast stage, and thus will not label the early pronormoblasts. This finding is consistent with studies showing that these glycophorin antibodies label only the post-CFU-E stages of erythrocyte development.

The diagnosis of acute megakaryoblastic leukemia by immunophenotyping with monoclonal antibodies directed against megakaryoblast-associated antibodies has allowed a rapid and easier method of diagnosing this poorly recognized subtype of acute leukemia. Monoclonal antibodies against platelet glycoprotein IIb/IIIa or IIIa (CD41 and CD61) are routinely used in the diagnosis of acute megakaryoblastic leukemia (FAB-M7). Although factor VIII–related antigen activity may be observed in rare cases of megakaryoblastic leukemia, the reactivity of megakaryoblasts with platelet glycoprotein IIb/IIIa or IIIa is more sensitive than labeling for factor VIII–related antigen.

Although immunophenotyping is generally accepted as the standard for diagnosing acute megakaryoblastic leukemia, it must be realized that some false positivity can be seen with these antibodies. This is most generally seen in myeloid leukemias that have a monocytic component. The increased reactivity of these megakaryoblast-associated glycoproteins appears to be nonspecific background labeling, reaction with platelet membrane components adhering to monocytic cells. In general, it has been our experience that the expression of these megakaryocyte-associated glycoproteins in the monocytic leukemias has been of much weaker intensity and fails to give a distinct uniform population on the fluorescence histograms but rather a “smearing” of intensities. Our policy has been to interpret a CD61-positive immunophenotype with some caution and to correlate the findings closely with morphologic and cytochemical findings. Immunocytochemical detection of these glycoproteins on glass-slide smears may also be of value in determining whether the positivity seen is truly on the leukemic cell or merely attributable to nonspecific adherence.

Other antigens not typically associated with myeloid cells can also be identified in AML. CD34 can be found in 30% to 40% of all AML and has been associated with a poorer prognosis than CD34-negative AML. The CD34 antigen can be useful in evaluating a specimen containing a mixture of myeloblasts and differentiated myeloid or lymphoid cells. CD38 has classically been used as an early thymocyte marker but is now recognized to be present on activated T cells as well as being an excellent plasma cell marker. Not surprisingly, it can also be identified in AML. CD4 is the receptor for HLA class II antigen that characterizes the T-helper subset of T lymphocytes. This antigen is not restricted to T cells and can be identified in monocytes, monocyte precursors, and monocytic leukemias. As previously discussed, CD7 is a pan–T-cell antigen that can be found in 5% to 10% of all AML cases; CD7 can be found in all subtypes of AML. Although some have suggested that the presence of CD7 reflects an early bone marrow precursor, other data do not support this concept. CD45, the common leukocyte antigen, is found in virtually all cases of AML.

In summary, immunophenotyping with monoclonal antibodies in AML is at its best when trying to distinguish between ALL and AML. One must use great caution in using the immunophenotype to aid in the FAB subclassification in these leukemias. Although the immunophenotype may suggest a relationship to a particular FAB subgroup, one must still rely on morphology, cytochemistry and cytogenetic findings to make the final diagnosis. The lack of DR expression suggests the possibility of an APL, expression of CD14, CD11b, and CD64 are associated with monocytic leukemias, and CD71 and glycophorin antibodies certainly react with erythroid cells. The final area in which immunophenotyping has been of great diagnostic utility is the identification of megakaryoblast-associated antigens for the diagnosis of acute megakaryoblastic leukemia.

ACUTE BIPHENOTYPIC LEUKEMIA

Part of "43 - Acute Leukemias and Myelodysplastic Syndromes"

Considerable interest has developed over the past decade in the acute leukemias that have immunophenotypes that do not fit into the accepted sequence of normal lymphoid or myeloid cell development. Perhaps the most controversial and confusing of these immunologically aberrant cases have been the so-called acute biphenotypic leukemias. A plethora of terminologies (hybrid, mixed-lineage, bilineal, biclonal, biphenotypic) has added

to the confusion in this area. Such phenotypes imply that either malignant transformation leads to aberrant gene expression (“lineage infidelity”) or that coexpression of lineage-specific markers occurs normally in the differentiation of hematopoietic cells (lineage “promiscuity”). The clinical significance of biphenotypia in acute leukemia has not been uniformly determined or accepted. Such studies have uniformly lacked consistent guidelines for classification. Some studies have suggested that the biphenotypic leukemias have a worse clinical prognosis than nonbiphenotypic ALL or AML, whereas other reports have not confirmed that a clinical significance exists. The difficulty in evaluating biphenotypic leukemias may be demonstrated by looking at the reported frequency of biphenotypic leukemia in the literature. This frequency has varied from less than 1% of acute leukemia cases in some studies to almost 50% of acute leukemia cases in others. This widely variable incidence range cannot be rationally explained solely based on the clinical variability between different population groups. Indeed, this variability, along with the lack of consensus as to the clinical or biological significance of biphenotypia in acute leukemia is undoubtedly owing to several contributing factors: the use of inconsistent and variable diagnostic criteria, the inevitable subjectiveness of immunophenotypic interpretation, and the lack of a standard and uniform panel of monoclonal antibodies utilized in immunophenotypic analyses.

Figure 43.34 is a schematic drawing outlining the diverse groups that compose what we have called biphenotypic leukemia. Leukemias with the Philadelphia chromosome or 11q23 translocations have clearly been associated with biphenotypic processes. Identification of other immunologic and cytogenetic subgroups will be essential in further characterizing and understanding these intriguing leukemias and the cellular counterparts from which they arise.

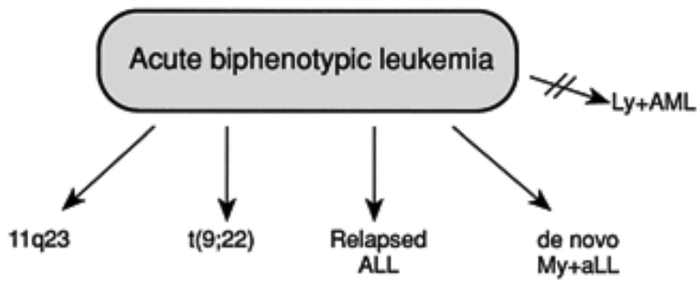


FIGURE 43.34. Proposed schema of subtypes of acute biphenotypic leukemia. Biphenotypic leukemias are broadly represented by cases having an 11q23 translocation, a t(9;22) translocation, relapsed acute lymphoblastic leukemia (ALL), and *de novo* myeloid antigen–positive ALL (My + ALL). Cases having lymphoid antigen–positive acute myelogenous leukemia (Ly + AML) probably do not represent true acute biphenotypic leukemia.

MYELODYSPLASTIC SYNDROMES

Part of "43 - Acute Leukemias and Myelodysplastic Syndromes"

The MDSs are a heterogeneous group of bone marrow disorders generally characterized by cytopenias and morphologic abnormalities of the erythroid, granulocytic, and megakaryocytic cell lines within a hypercellular bone marrow. Bone marrow myeloblasts may be increased but do not reach the 20% blast level necessary for a diagnosis of AML. Although MDS does not meet the morphologic criteria for the diagnosis of AML, distinction from AML may be difficult and a proportion of cases will ultimately evolve to overt AML. Patients who do not evolve into AML frequently will experience complications related to chronic transfusion or bone marrow failure.

MDS usually occurs in individuals older than the age of 50, although genuine cases of MDS may rarely be seen in the pediatric age group. In general, however, the diagnosis of MDS in children should be made very cautiously. Patients with MDS typically present with symptoms of variable duration related to the degree of the underlying cytopenias.

Historical terms that have been used to describe MDS include preleukemia, early leukemia, smoldering leukemia, subacute leukemia, atypical leukemia, and hematopoietic dysplasia. The MDSs do not include disorders that are known to predispose an individual to the development of acute leukemia, such as chromosomal breakage syndromes. Although the MDSs are classified as distinctly different from the chronic myeloproliferative disorders, there are certainly individual cases that have features of both MDS and chronic myeloproliferative diseases. Such cases may present with neutrophilia and granulocytic hyperproliferation together with striking dysplastic features or thrombocytosis and dysplastic neutrophils and megakaryocytes.

Diagnosis

A diagnosis of MDS should be considered in a setting of unexplained cytopenias and ineffective hematopoiesis in an adult patient. Most MDSs will show quantitative and qualitative evidence of dyserythropoiesis, dysgranulopoiesis, and dysmegakaryopoiesis. An iron stain for ringed sideroblasts and the percentage of blood and bone marrow myeloblasts are essential in accurately classifying the MDS. These and other features are listed in Table 43.15 and are described below.

TABLE 43.15. MORPHOLOGIC FEATURES OF MYELODYSPLASIA

Dyserythropoiesis
Megaloblastoid changes
Multinucleation
Nuclear/cytoplasmic asynchrony
Nuclear karyorrhexis
Ringed sideroblasts
Dysgranulopoiesis
Pseudo Pelger-Huet changes
Hypogranulation
Abnormal granules
Basophilia
Decreased or absent myeloperoxidase
Decreased leukocyte alkaline phosphatase
Dysmegakaryopoiesis
Large, bizarre platelets
Hypogranular platelets
Bizarre megakaryocytes
Uninuclear megakaryocytes

Dyserythropoiesis

Dyserythropoiesis may include megaloblastoid changes, irregular nuclear shapes and forms, multinucleation, internuclear

bridging, abnormal chromatin, ringed sideroblasts, or nuclear karyorrhexis. Dyssynchrony between nuclear and cytoplasmic development is probably the most reliable and significant evidence of dyserythropoiesis (Fig. 43.35). For example, a red cell precursor showing evidence of hemoglobinization in the cytoplasm, while still maintaining an earlier or open chromatin pattern, would be described as dyssynchronous. The peripheral blood smear may show significant anisopoikilocytosis and occasional nucleated red blood cells. The term ringed sideroblasts has been used to describe erythroblasts that contain cytoplasmic iron that wraps around at least one third of the nucleus. This iron is located in the mitochondria and results in ineffective erythropoiesis and subsequent marrow destruction of red cells and precursors. Ringed sideroblasts are not found in normal marrow and are easily distinguished from normal sideroblasts, which have scattered cytoplasmic iron granules.

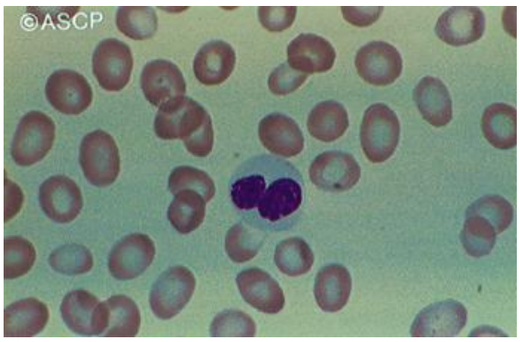


FIGURE 43.35. Dyserythropoiesis.

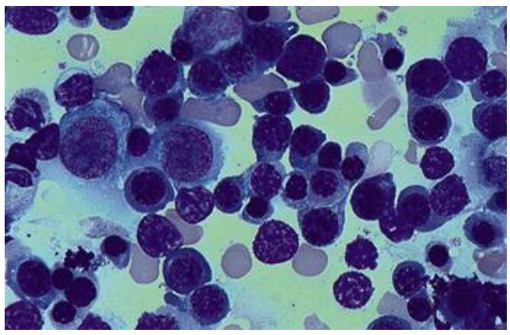


FIGURE 43.36. Dysgranulopoiesis (pseudo–Pelger-Huét change).

Dysgranulopoiesis

Dysgranulopoiesis can be effectively summarized as morphologic evidence of hypolobulation or hyposegmentation of the nucleus and/or hypogranulation of the cytoplasm (Fig. 43.36). The hyposegmentation of the nucleus, i.e., the presence of only two nuclear lobes, has been called pseudo–Pelger-Huét change. The determination of this hyposegmentation must be made on a well-prepared blood or bone marrow smear, and care must be taken to distinguish that two lobes are not overlapping each other. Likewise, hypogranulation is best identified on a Wright-Giemsa stain smear. Smears stained with only Wright stain may sometimes appear faint, leading to a mistaken diagnosis of hypogranulation. In addition, granulocyte enzymatic stains, such as MPO, SBB, and leukocyte alkaline phosphatase, may be negative if this dysplastic hypogranulation is marked. Occasionally, one may see abnormally dense chromatin or ringed nuclei, both of which represent uncommon forms of dysgranulopoiesis. Abnormal granulation and basophilia are other manifestations of dysgranulopoiesis.

Dysmegakaryopoiesis

Abnormalities of peripheral blood platelets and bone marrow megakaryocytes are common in the MDSs. Dysplastic platelets on the peripheral blood smear may appear as either hypogranular platelets or as large, bizarre platelet forms. Examination of the bone marrow aspirate smear and/or bone marrow biopsy specimen may show megakaryocytes with separated nuclear lobes instead of a lobulated nucleus, large bizarre nuclei with dysplastic features, open nuclear chromatin, high nuclear:cytoplasmic ratios, and large accumulations, or sheets, of megakaryocytes in the bone marrow biopsy (Fig. 43.37). Abnormally small megakaryocytes with single, round nuclei have been described in the MDSs having a chromosomal deletion of the long arm of chromosome 5 (5q-syndrome).

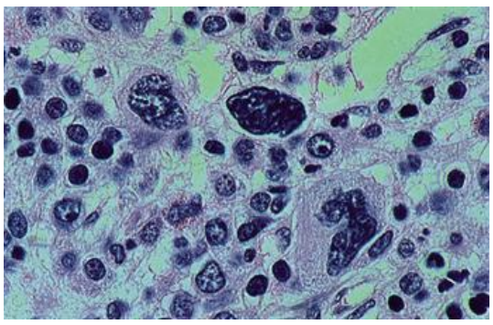


FIGURE 43.37. Dysmegakaryopoiesis.

FAB Classification

The FAB classification scheme for MDS is listed in Table 43.16. However, the recent proposal by WHO redefined some of the categories in the FAB classification (Table 43.17). One of the most important changes was decreasing the blast count from 30% to 20% for diagnosis of AML. The high-grade myelodysplastic process is also consolidated in one category, refractory anemia with excess blasts (RAEB), with two types based on blast count. Chronic myelomonocytic leukemia is designated as myelodysplastic/myeloproliferative syndromes because they typically have features of both disorders. Finally, WHO recognizes the two categories as a separate entity: 5q-syndrome and refractory cytopenia with multilineage dysplasia.

TABLE 43.16. FAB CLASSIFICATION OF MYELODYSPLASTIC SYNDROMES

Refractory anemia
Idiopathic refractory sideroblastic anemia
Refractory anemia with excess blasts (RAEB)
RAEB in transformation
Chronic myelomonocytic leukemia
Myelodysplastic syndrome, unclassified (NOS)
Therapy-related myelodysplasia

NOS, not otherwise specified.

TABLE 43.17. PROPOSED WORLD HEALTH ORGANIZATION CLASSIFICATION OF MYELODYSPLASTIC SYNDROMES

Refractory anemia
Refractory anemia with ringed sideroblasts
Refractory cytopenia with multilineage dysplasia
Refractory anemia with excess blasts
Type 1: 5% to 9% blasts in blood or marrow
Type 2: 10% to 19% blasts in blood or marrow
5q-syndrome
Therapy-related myelodysplastic syndrome
Myelodysplastic syndrome, unclassified

It should also be noted that secondary or therapy-related MDS is also a newly recognized distinct clinicopathologic entity occurring in patients after exposure to radiation and/or alkylating-agent chemotherapy for primary neoplastic or nonneoplastic disorders.

Refractory Anemia

This subtype of MDS is basically a chronic, unexplained anemia occurring in adult patients older than the age of 50. It may also be referred to as refractory cytopenia, as it is not purely restricted to anemia (Table 43.18). Classically, patients with refractory anemia show erythroid hyperplasia in the bone marrow accompanied by a low reticulocyte count in the peripheral blood. This combination is referred to as ineffective erythropoiesis and is a characteristic hallmark of this type of MDS. Ringed sideroblasts are not increased (less than 15%) in these cases, and it is important to note that minimal to no dysplastic changes are identified in any cell line; no increase in myeloblasts are seen. This subgroup of MDS consists mainly of patients who have unexplained anemias despite erythroid hyperplasia. Obviously, other causes of chronic anemia must be ruled out, including vitamin B_{12}/folate deficiency, drug or toxin exposure, congenital dyserythropoietic anemias, aplastic anemia, or other uncommon causes of anemia.

TABLE 43.18. FAB CLASSIFICATION IN MYELODYSPLASTIC SYNDROMES

Refractory anemia
Chronic unexplained anemia/refractory cytopenia
Reticulocytopenia
No blasts in peripheral blood
Bone marrow normocellular or hypercellular
Erythroid hyperplasia/mild dyserythropoiesis
Minimal to no dysplastic changes
<5% myeloblasts
Ringed sideroblasts are not increased
Idiopathic refractory sideroblastic anemia
Findings similar to those described for refractory anemia
<5% myeloblasts
Dysplastic features are uncommon
Erythroid hyperplasia; hypercellular bone marrow
Dyserythropoiesis
>15% ringed sideroblasts
Chronic anemia, transfusion dependent
Dimorphic peripheral blood (hypochromic and normochromic)
Ineffective erythropoiesis
10% evolve to AML
Refractory anemia with excess blasts
Chronic clinical course
Pancytopenia
All three cell lines are dyspoietic
Bone marrow failure is common
Hypercellular bone marrow
<5% myeloblasts in blood
5% to 20% myeloblasts in bone marrow
No Auer rods
Ringed sideroblasts may be present
30% to 40% evolve to AML
Refractory anemia with excess blasts in transformation
20% to 30% myeloblasts in bone marrow
Or 0 to 30% myeloblasts with Auer rods
Or >5% myeloblasts in blood
Same clinical features as refractory anemia with excess blasts
60% to 80% evolve to AML
Chronic myelomonocytic leukemia
Males more than females
Hepatomegaly/splenomegaly in 30% to 50%
Anemia, thrombocytopenia
Leukocyte count is variable
Monocytes >2 × 10^9/L
Hypercellular marrow
Significant myelodysplastic changes may not be found
Increased serum/urine lysozyme
Easily confused with AML-M4 or M5b

AML, acute myelogenous leukemia.

Idiopathic Refractory Sideroblastic Anemia

The morphologic features of idiopathic refractory sideroblastic anemia (IRSA) are similar to those described for refractory anemia except for the presence of 15% or more ringed sideroblasts. The criteria for IRSA are listed in Table 43.18. The pathologic ringed sideroblasts have iron granules located in a perinuclear location, reflecting mitochondrial iron deposition (Fig. 43.38). It is important to rule out hereditary and secondary causes of sideroblastic anemia, including hereditary sideroblastic anemia, exposure to drugs (e.g., antituberculosis agents, chloramphenicol), alcohol-induced sideroblastic anemia, chronic lead poisoning, and various other diseases, including some autoimmune disorders. These patients have a significant erythroid hyperplasia but do not have increased numbers of myeloblasts. Dysplastic features are uncommon in granulocytic and megakaryocytic precursors and, when present, may portend a more aggressive clinical course. Most patients become transfusion dependent, and only 10% to 15% of cases of IRSA progress to an AML. The peripheral blood smear classically shows a dimorphic blood picture with both normochromic, normocytic red blood cells and hypochromic, microcytic red blood cells (Fig. 43.39). This dimorphism

reflects the heterogeneity of erythroid differentiation in the bone marrow. This dimorphic blood picture is quite distinct and should always raise the possibility of IRSA when encountered.

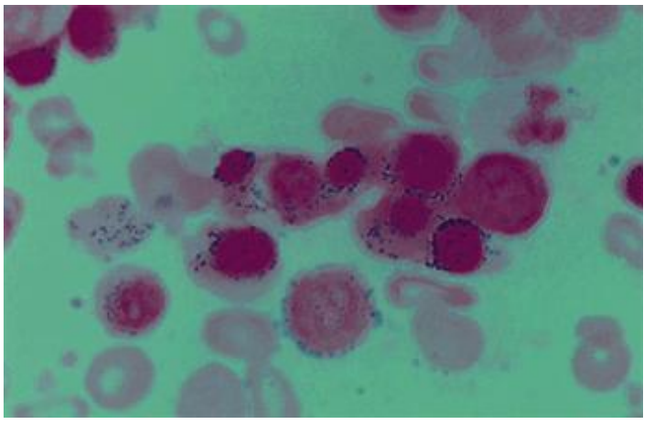


FIGURE 43.38. Idiopathic refractory sideroblastic anemia with ringed sideroblasts.

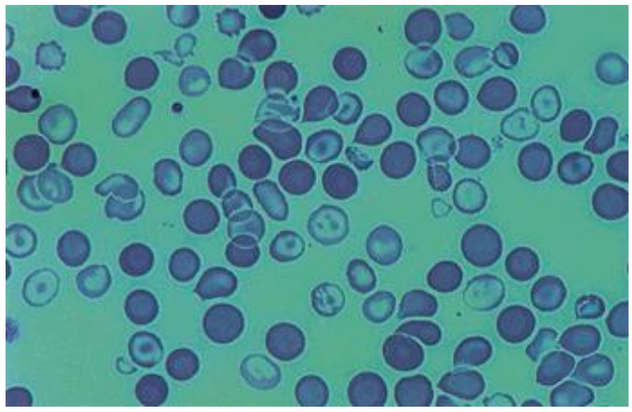


FIGURE 43.39. Dimorphic peripheral blood smear in idiopathic refractory sideroblastic anemia.

Refractory Anemia with an Excess of Blasts

Refractory anemia with an excess of blasts (RAEB) is typically associated with pancytopenia and significant dysplastic features in the granulocytic series. Erythroid and megakaryocytic dysplasia may also be found. The basic diagnostic criterion for RAEB is a hypercellular bone marrow containing between 5% and 20% myeloblasts. The criteria for the diagnosis of RAEB are listed in Table 43.18. It is important to note that a significant number of patients in this subgroup of MDS will eventually evolve into an overt AML. Ringed sideroblasts may be identified; however, the increased myeloblasts allow for distinction from IRSA.

RAEB is separated into two subgroups based on the blast count. Type 1 is characterized by a blast count between 5% and 9% in the blood or bone marrow, and type 2 includes patients with a blast count of 10% to 19%. It is determined that patients with a blast count greater than 10% usually have a more aggressive clinical course and are more likely to develop overt leukemia. The category of RAEB in transformation (RAEB-t) in the FAB classification was eliminated by the WHO classification.

Secondary/Therapy-Related MDS

Therapy-related MDS is a form of MDS occurring in patients previously treated with chemotherapy and/or radiation for neoplastic or nonneoplastic disease. The incidence of a secondary MDS or AML has been variously estimated at 2% to 10% of patients treated as such and may be related to the amount, duration, and repetition of exposure to the therapeutic agent and the age of the individual involved. Long-term alkylating-agent therapy and the type II topoisomerase inhibitors (epipodophyllotoxins) are the most commonly implicated in dysplastic transformation. Topoisomerase inhibitors particularly are associated with abnormalities of chromosome 11q23. The median onset of secondary MDS is approximately within 5 years of therapy, with a range of approximately 2 to 15 years. Approximately one third of these patients will present with a MDS before the evolution to a frank AML. Typically, these patients present with unexplained cytopenias and observed dysplastic abnormalities in the blood and bone marrow. Panmyelosis with dysplastic features in all myeloid lines are typically found. The prognosis of patients with therapy-related MDS is abysmal, with a median survival of less than 6 months.

5q-Syndrome

These patients usually present with macrocytosis and thrombocytosis. Bone marrow evaluation typically shows erythro-blastopenia, megakaryocytic hyperplasia with unique morphology (hypolobulated mononuclear smaller size). These patients have a much better prognosis compared with the other MDS categories but usually become transfusion dependent.

Cytogenetics in MDS

Several recurring chromosome defects are found in patients with *de novo* MDS, including complete or partial loss of chromosome 5 or 7, loss of the long arm of chromosome 5 or 7, trisomy 8, or multiple, miscellaneous complex chromosome defects.

Prognosis

There are varying reports of the frequency that MDS evolves to overt AML. Data in the literature regarding the evolution to AML vary according to the subtype of MDS (Table 43.19). Generally, refractory anemia and IRSA show the lowest incidence of transformation to AML, whereas RAEB shows a much higher incidence of leukemic progression.

TABLE 43.19. PROGNOSTIC FACTORS IN MYELODYSPLASTIC SYNDROMES

Low risk
Refractory anemia with normal platelets, WBC
Numerous ringed sideroblasts
Normal CFU capacity
Absence of chromosomal abnormalities
High risk
Severe pancytopenia
Refractory anemia with excess blasts in transformation
Low CFU capacity
Chromosomal abnormalities

WBC, white blood cells; CFU, colony-forming unit.

The survival of patients with MDS varies, with some patients surviving many years and others only a few months. Death may be owing to bone marrow failure if progression to AML does not occur. The age of patients, the degree of cytopenia, the percentage of blasts, the degree of abnormal chromosome abnormalities, and *in vitro* culture changes have all been used to predict the survival of patients with MDS.

SUGGESTED READINGS

Alessandrino EP, Orlandi E, Brusamolino E, et al. Chronic myelomonocytic leukemia: clinical features, cytogenetics, and prognosis in 30 consecutive cases. *Hematol Oncol* 1985;3:147–155.

Arthur D, Bloomfield CD, Lindquist LL, et al. Translocation 4;11 in acute lymphoblastic leukemia: clinical characteristics and prognostic significance. *Blood* 1982;59:96–99.

Bain B, Catovsky D, O'Brien M, et al. Megakaryoblastic leukemia presenting as acute myelofibrosis: study of four cases with platelet peroxidase reaction. *Blood* 1981;58:206–213.

Belanger R, Gyger M, Perreault C, et al. Bone marrow transplantation for myelodysplastic syndromes. *Br J Haematol* 1988;69:29–33.

Bennett JM, Catovsky D, Daniel HT, et al. Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol* 1982;51:189–199.

Bennett IM, Reed CE. Acute leukemia cytochemical profile: diagnostic and clinical implications. *Blood Cells* 1975;1:101–113.

Bennett JM, Catovsky D, Daniel MT, et al. Proposals for classification of acute leukemias French-American-British (FAB) Cooperative Group. *Br J Haematol* 1976;33:451–458.

Bennett JM, Catovsky D, Daniel MR, et al. A variant form of hypergranular promyelocytic leukemia (M3). *Br J Haematol* 1980;44:169–170.

Bennett JM, Catovsky D, Daniel MT, et al. Proposed revised criteria for the classification of acute myeloid leukemia. *Ann Intern Med* 1985;103:626–629.

Bennett JM, Catovsky D, Marit-Therese D, et al. Criteria for the diagnosis of acute leukemia of megakaryocyte lineage (M7). *Ann Intern Med* 1985;103:460–462.

Bennett JM, Catovsky D, Daniel M-T, et al. Proposed revised criteria for the classification of acute myeloid leukemia. *Ann Intern Med* 1985;103:620–625.

Bennett JM, Catovsky D, Daniel M-T, et al. Proposal for the recognition of minimally differentiated acute myeloid leukaemia (AML-M0). *Br J Haematol* 1991;78:325–329.

Berger R, Bernheim A, Daniel M-T, et al. Cytologic characterization and significance of normal karyotypes in t(8;21) AML. *Blood* 1982;59:171–178.

Bloomfield CD, Brunning RD. FAB-M7: acute megakaryoblastic leukemia-beyond morphology. *Ann Intern Med* 1985;103:450–452.

Borowitz MJ, Shuster JJ, Civin CI, et al. Prognostic significance of CD34 expression in childhood B-precursor acute lymphocytic leukemia: a pediatric-oncology group study. *J Clin Oncol* 8:1389–1398.

Boultwood J, Lewis S, Wainscoat JS. The 5q-syndrome. *Blood* 1994;84:3253–3260.

Breton-Gorius J, Reyes F, Duhamel G, et al. Megakaryoblastic acute leukemia: identification of the ultrastructural demonstration of platelet peroxidase. *Blood* 1978;51:45–60.

Catovsky D, de Salvo Cardullo L, O'Brien M, et al. Cytochemical markers of differentiation in acute leukemia. *Cancer Res* 1981;11:4824–4832.

Cave H, Van Der Werff ten Bosch J, Suciu S, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. *N Engl J Med* 1998;339:591–598.

Cazzola M, Barosi G, Gobbi PG, et al. Natural history of idiopathic refractory sideroblastic anemia. *Blood* 1988;71:305–312.

Chessells JM, Swansbury GJ, Reeves B, et al. Cytogenetics and prognosis in childhood lymphoblastic leukaemia: results of MRC UKALL X. *Br J Haematol* 1997;99:93–100.

Coiffier B, Adeleine P, Viola JJ, et al. Dysmyelopoietic syndromes. A search for prognostic factors in 193 patients. *Cancer* 1983;52:83–90.

Coiffier B, Adeleine P, Gentilhomme O, et al. Myelodysplastic syndromes: a multiparametric study of prognostic factors in 336 patients. *Cancer* 1987;60:3029–3032.

Copelan EA, McGuire EA. The biology and treatment of acute lymphoblastic leukemia in adults. *Blood* 1995;85:1151–1168.

d'Onofrio G, Mancini S, Tamburrini E, et al. Giant neutrophils with increased peroxidase activity: another evidence of dysgranulopoiesis in AIDS. *Am J Clin Pathol* 1987;87:584–591.

Delacretaz F, Schmidt P, Piguet D, et al. Histopathology of myelodysplastic syndromes. The FAB classification (proposals) applied to bone marrow biopsy. *Am J Clin Pathol* 1987;87:180–186.

Faderl S, Kantarjian HM, Talpaz M, et al. Clinical significance of cytogenetic abnormalities in adult acute lymphoblastic leukemia. *Blood* 1998;91:3995–4019.

Fenaux P, Morel P, Lai JL. Cytogenetics of myelodysplastic syndromes. *Semin Hematol* 1996;33:127–138.

Foucar K, McKenna RW, Bloomfield CD, et al. Therapy-related leukemia: a panmyelosis. *Cancer* 1979;43:1285–1296.

Foucar K, Langdon RM, Armitage JO, et al. Myelodysplastic syndromes. A clinical and pathologic analysis of 109 cases. *Cancer* 1985;56:553–561.

Gassman W, Loffler H, Thiel E, et al. Morphological and cytochemical findings in 150 cases of T-lineage acute lymphoblastic leukaemia in adults. *Br J Haematol* 1997;97:372–382.

Golomb HM, Rowley JD, Vardiman JW, et al. Microgranular acute promyelocytic leukemia: a distinct clinical, ultrastructural, and cytogenetic entity. *Blood* 1980;55:253–259.

Greaves MF, Hariri G, Newman RA, et al. Selective expression of the common acute lymphoblastic leukemia (gp 100) antigen on immature lymphoid cells and their malignant counterparts. *Blood* 1983;61:628–639.

Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* 1997;89:2079–2088.

Groupe Francais de Morphologie Hematologique. French Registry of acute leukemia and myelodysplastic syndromes. Age distribution and hemogram analysis of the 4496 cases recorded during 1982-1983 and classified according to FAB criteria. *Cancer* 1987;60:1385–1394.

Haas OA, Gadner H. Pathogenesis, biology, and management of myelodysplastic syndromes in children. *Semin Hematol* 1996;33:225–235.

Haferlach T, Löffler H, Nickenig C, et al. Cell lineage specific involvement in acute promyelocytic leukaemia (APL) using a combination of May-Grünwald-Giemsa staining and fluorescence in situ hybridization techniques for the detection of the translocation t(15;17)(q22;q12). *Br J Haematol* 1998;103:93–99.

Harris NL, Jaffe ES, Diebold J, et al. The World Health Organization classification of Hematological Malignancies Report of the Clinical Advisory Committee Meeting, Airlie House, Virginia, November 1997. *Mod Pathol* 2000;13:193–207.

Hurwitz CA, Mirro J. Mixed lineage leukemia and asynchronous antigen expression. *Hematol Oncol Clin North Am* 1990;4:767–794.

Janssen JWG, Buschle M, Layton M, et al. Clonal analysis of myelodysplastic syndromes: evidence of multipotent stem cell origin. *Blood* 1989;73:248–254.

Juneja SK, Imbert M, Sigaux F, et al. Prevalence and distribution of ringed sideroblasts in primary myelodysplastic syndromes. *J Clin Pathol* 1983;36:566–569.

Juneja SK, Imbert M, Jouault H, et al. Haematological features of primary myelodysplastic syndromes (PMDS) at initial presentation: a study of 118 cases. *J Clin Pathol* 1983;36:1129–1135.

Kaffe S, Hsu LYF, Hoffman R, et al. Association of 5q and refractory anemia. *Am J Hematol* 1978;4:269–272.

Kantarjian HM, Keating MJ, Walters RS, et al. Therapy-related leukemia and myelodysplastic syndrome: clinical, cytogenetic, and prognostic features. *J Clin Oncol* 1986;4:1748–1757.

Kersey JH. Fifty years of studies of the biology and therapy of childhood leukemia. *Blood* 1997;90:4243–4251.

Knuutila S. Lineage specificity in haematological neoplasms. *Br J Haematol* 1997;96:2–11.

Koeffler HP. Introduction: myelodysplastic syndromes. *Semin Hematol* 1996;33:87-94.

Lewis DS, Thompson M, Hudson E, et al. Down's syndrome and acute megakaryoblastic leukemia. *Acta Haematol* 1983;70:236–242.

Lambertenghi-Deliliers G, Orazi A, Luksch R, et al. Myelodysplastic syndrome with increased marrow fibrosis: a distinct clinico-pathological entity. *Br J Haematol* 1991;78:161–166.

Licht JD, Chomienne C, Goy A, et al. Clinical and molecular characterization of a rare syndrome of acute promyelocytic leukemia associated with translocation (11;17). *Blood* 1995;85:1083–1094.

Liu PP, Hajra A, Wijmenga C, et al. Molecular pathogenesis of the chromosome 16 inversion in the M4Eo subtype of acute myeloid leukemia. *Blood* 1995;85:2289–2302.

Longacre T, Foucar K, Crago S, et al. Hematogones: a multiparameter analysis of bone marrow precursor cells. *Blood* 1989;73:543–552.

McKenna RW, Brynes R, Nesbit M, et al. Cytochemical profiles in acute lymphoblastic leukemia. *Am J Pediatr Hematol Oncol* 1979;1:263–275.

McKenna RW, Bloomfield CD, Dick F, et al. Acute monoblastic leukemia: diagnosis and treatment of 10 cases. *Blood* 1978;45:481–494.

McKenna RW, Parkin J, Bloomfield C, et al. Acute promyelocytic leukemia. Study of 39 cases with identification of a hyperbasophilic microgranular APL variant. *Br J Haematol* 1982;50:201–214.

McKenna RW. A multifaceted approach to diagnosis and classification of acute leukemia. *Arch Pathol Lab Med* 1991;115:328–330.

Mufti GJ, Stevens JR, Oscier DG, et al. Myelodysplastic syndromes: a scoring system with prognostic significance. *Br J Haematol* 1985;59:425–433.

Nand S, Godwin JE. Hypoplastic myelodysplastic syndrome. *Cancer* 1988;62:958–964.

Neame PB, Soamboonsrup B, Browman GP, et al. Classifying acute leukemia by immunophenotyping: a combined FAB-immunologic classification of AML. *Blood* 1986;68:1355–1362.

Nucifora G, Rowley JD. AML1 and the 8;21 and 3;21 translocations in acute and chronic myeloid leukemia. *Blood* 1995;86:1–14.

Park DJ, Koeffler HP. Therapy-related myelodysplastic syndromes. *Semin Hematol* 1996;33:256–273.

Peterson LC, Parkin JL, Arthur DC, et al. Acute basophilic leukemia: a clinical morphologic and cytogenetic study of eight cases. *Am J Clin Pathol* 1991;96:160–170.

Pui CH, Behm FG, Crist WM. Clinical and biologic relevance of immunologic marker studies in childhood acute lymphoblastic leukemia. *Blood* 1993;82:343–362.

Pui CH, Ribeiro PC, Hancock ML, et al. Acute myeloid leukemia in children treated with epipodophyllotoxins for acute lymphoblastic leukemia. *N Engl J Med* 1991;325:1682–1687.

Ribera J, Cervantes F, Rozman C. A multivariate analysis of prognostic factors in chronic myelomonocytic leukemia according to the FAB criteria. *Br J Haematol* 1987;65:307–311.

Rosati S, Anastasi J, Vardiman J. Recurring diagnostic problems in the pathology of the myelodysplastic syndromes. *Semin Hematol* 1996;33:111–126.

Ross FM, Hamilton M, Cook MK, et al. A myelodysplastic syndrome with eosinophilia associated with a break in the short arm of chromosome 16. *Leukemia* 1987;1:680–681.

Schneider DR, Picker LJ. Myelodysplasia in the acquired immune deficiency syndrome. *Am J Clin Pathol* 1985;84:144–152.

Second MIC Cooperative Study Group. Morphologic, immunologic and cytogenetic (MIC) working classification of AML. *Cancer Genet Cytogenet* 1988;30:1–15.

Second International Workshop on Chromosomes in Leukemia (1979). Cytogenetic, morphologic and clinical correlations in acute nonlymphocytic leukemia with t(8q-;21q+). *Cancer Genet Cytogenet* 1980;2:99–102.

Second International Workshop on Chromosomes in Leukemia (1979). Morphologic analysis of acute promyelocytic leukemia (M3) and t(8;21) cases. *Cancer Genet Cytogenet* 1980;2:97–98.

Shurtleff SA, Buijs A, Behm FG, et al. TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia* 1995;9:1985–1989.

Teerenhovi L, Lintula R. Natural course of myelodysplastic syndromes-Helsinki experience. *Scand J Haematol* 1986;36:102–106.

Tenen DG, Hromas R, Licht JD, et al. Transcription factors, normal myeloid development, and leukemia. *Blood* 1997;90:489–519.

Third MIC Cooperative Study Group. Recommendations for a morphologic, immunologic, and cytogenetic (MIC) working classification of the primary and therapy related myelodysplastic disorders. *Cancer Genet Cytogenet* 1988;32:1–10.

Tricot G, De Wolf-Peeters C, Hendrickx B, et al. Bone marrow histology in myelodysplastic syndromes. I. Histological findings in myelodysplastic syndromes and comparison with bone marrow smears. *Br J Haematol* 1984;57:423–430.

Tricot G, De Wolf-Peeters C, Vlietinck R, et al. Bone marrow histology in myelodysplastic syndromes. II. Prognostic value of abnormal localization of immature precursors in MDS. *Br J Haematol* 1984;58:217–225.

Tricot G, Vlietinck R, Boogaerts MA, et al. Prognostic factors in the myelodysplastic syndromes: importance of initial data on peripheral blood counts, bone marrow cytology, trephine biopsy and chromosomal analysis. *Br J Haematol* 1985;60:19–32.

Tuzuner N, Cox C, Rowe JB, et al. Hypocellular myelodysplastic syndromes (MDS): new proposals. *Br J Haematol* 1995;91:612–617.

Uckun FM, Sather HN, Gaynon PS, et al. Clinical features and treatment outcome of children with myeloid antigen positive acute lymphoblastic leukemia: a report from the Children's Cancer Group. *Blood* 1997;90:28–35.

Uckun FM, Sensel MG, Sun L, et al. Biology and treatment of childhood T-lineage acute lymphoblastic leukemia. *Blood* 1998;91:735–746.

Vardiman JW, Head D. Society for hematopathology: the myelodysplastic syndromes (MDS) and related disorders. *Mod Pathol* 1999;12:101–106.

Venditti A, Del Poeta G, Buccisano F, et al. Prognostic relevance of the expression of TdT and CD7 in 335 cases of acute myeloid leukemia. *Leukemia* 1998;12:1056–1063.

Zeleznik-Le NJ, Nucifora G, Rowley JD. The molecular biology of myeloproliferative disorders as revealed by chromosomal abnormalities. *Semin Hematol* 1995;32:201–219.

44

Chronic Lymphoproliferative Disorders, Immunoproliferative Disorders, and Malignant Lymphoma

Fred Randolph Dick

The lymphoproliferative disorders represent a spectrum of diseases with a broad morphologic, functional, and clinical diversity. The classification in Table 44.1 is based on the following: diseases classified as the *chronic lymphoproliferative disorders* predominantly affect the blood and bone marrow with varying degrees of tissue involvement. Chronic lymphoproliferative disorders include neoplastic proliferations of peripheral B-cells, T-cells, and NK-cells (1). Although these processes are classified as "chronic" lymphoproliferative disorders, some of them do not follow an indolent course. The *immunoproliferative disorders* are predominantly bone marrow-based diseases with varying degrees of tissue involvement. They are neoplasms of B-cells at the terminal stage of functional development, and thus are characterized by production of a monoclonal immunoprotein. The *malignant lymphomas* are predominantly tissue-based disorders; however, they also may have blood and marrow involvement, especially late in the course of disease. The morphology, functional cell of origin, and clinical course of malignant lymphomas is highly variable. The *acute lymphoblastic leukemias* will be discussed in a separate chapter. They are diseases of precursor B-cells and T-cells. Similar to the chronic lymphoproliferative disorders, they are predominantly blood- and bone-marrow–based disorders, but follow an "acute" course if untreated.

TABLE 44.1. CLASSIFICATION OF LYMPHOPROLIFERATIVE DISORDERS

Chronic Lymphoproliferative Disorders
Peripheral B-cell
Chronic lymphocytic leukemia
Prolymphocytic leukemia
Hairy-cell leukemia
Peripheral T-cell
Chronic lymphocytic leukemia
Prolymphocytic leukemia
Large granular lymphocyte leukemia
Sezary syndrome
Adult T-cell leukemia/lymphoma
Immunoproliferative Disorders
Plasma cell myeloma
Plasmacytoma
Waldenström's macroglobulinemia
Heavy chain disease
Benign monoclonal gammopathy
Amyloidosis
Malignant Lymphomas
Non-Hodgkin's lymphoma
B-cell lymphomas
T-cell and NK-cell lymphomas
Hodgkin's disease
Postransplant lymphoproliferative disorders
Acute Lymphocytic Leukemia (derived from precursor B- and T-cells) will be discussed in a separate chapter.

CHRONIC LYMPHOPROLIFERATIVE DISORDERS OF B-CELL TYPE

Part of "44 - Chronic Lymphoproliferative Disorders, Immunoproliferative Disorders, and Malignant Lymphoma"

Chronic lymphoproliferative disorders of B-cell type vary widely in their clinical course, ranging from less than 3 years for prolymphocytic leukemia (PLL), to 7 years or more for most subtypes of chronic lymphocytic leukemia and hairy-cell leukemia. Morphologically, they show a "mature" appearance with condensation of nuclear chromatin; however, some element of morphologic "immaturity" may be present.

Chronic Lymphocytic Leukemia

Chronic lymphocytic leukemia (CLL) is a common neoplastic disease characterized by proliferation of small, morphologically mature lymphocytes (1, 2 and 3). CLL involves the blood and bone marrow, and also frequently the lymph nodes, spleen, and liver. It is a disease with a male predominance and a median age of approximately 65, being rarely seen in individuals less than 40 years of age (4). It usually follows an indolent course, and many patients are diagnosed by finding an elevated lymphocyte count incident through evaluation for another disease. Most cases of CLL are of B-cell type, and the following discussion applies to B-cell CLL. T-cell CLL, which is very rare, is biologically distinctly different from B-cell CLL and will be discussed in a subsequent section.

Diagnostic Features

Unlike most other lymphoproliferative diseases, there is no cytologic feature of the individual cell in CLL that is diagnostic. Thus, the diagnosis rests on finding increased numbers of nearly morphologically normal-appearing lymphocytes.

To diagnose CLL, there should be a peripheral blood lymphocytosis greater than 5,000/mm^3 and a marrow lymphocytosis greater than 30% (3). At diagnosis, most patients have a white blood cell count (WBC) of greater than 15,000/mm^3 with a median count of about 30,000 to 40,000/mm^3. Some cases may have a WBC well over 100,000/mm^3. Because CLL is a clonal process, there is frequently a monotonous look-alike appearance to the lymphocytes on smears (Fig. 44.1). Cytoplasm usually is small to moderate in amount, pale blue, and agranular; the nuclear chromatin frequently shows exaggerated chromatin clumping compared with a normal lymphocyte.

The bone marrow in CLL usually aspirates freely, and marrow sections may range from normocellular to hypercellular (Fig. 44.2). When the marrow is less extensively infiltrated, the lymphocytes in the marrow may be present in focal aggregates (Fig. 44.2A) or may be diffusely intermixed with residual normal marrow elements (Fig. 44.2B). This latter pattern is called an interstitial infiltrative pattern. Frequently, a mixed focal and interstitial pattern may be seen. As the marrow becomes more extensively infiltrated, focal or interstitial patterns are gradually lost, and the infiltrate diffusely obliterates the marrow space (Fig. 44.2C). As illustrated in Fig. 44.2D, lymphocytes in CLL

on sections are indistinguishable from normal mature lymphocytes.

When the blood and marrow counts are near the lower limits described above, and the diagnosis is in question, immunologic cell surface marker studies may be helpful in establishing the diagnosis and in ruling out post-splenectomy lymphocytosis, a chronic reactive lymphocytosis (5) or other chronic lymphoproliferative disorders. Clonality, as evidenced by a weakly reacting surface immunoglobulin with a restricted light chain, and CD5 and CD23 positivity are characteristic of CLL (Table 44.2) (6, 7 and 8). In cases with a low WBC in the range of 5,000 to 15,000/mm^3, it may be easier to make the diagnosis using peripheral blood immunology than with bone marrow examination, because the bone marrow in these cases also may show a low percentage of lymphocytes, and characteristic spreading focal aggregates may be absent. If immunologic marker studies are available, immunologic evaluation of the peripheral blood is the method of choice to confirm the diagnosis, especially if therapy is not contemplated (6).

TABLE 44.2. SPECIAL STUDIES IN THE DIAGNOSIS AND DIFFERENTIAL DIAGNOSIS OF CHRONIC LYMPHOPROLIFERATIVE DISORDERS OF B-CELL TYPE

SIg	CD5	CD20	CD23	CD25	CD11c	FMC7	CD79b	TRAP	
CLL	weak	+	weak	+	-/+	-/+	-	-	-/+
PLL	strong	-/+	+	-	-/+	-	+	+	-/+
HCL	strong	-	+	-	+	+	+	-	+

SIg, surface immunoglobulin

TRAP, tartarate resistant acid phosphatase

PLL, prolymphocytic leukemia

HCL, hairy-cell leukemia

-/+, Typically negative with occasional exceptions

Morphologic Variability And Differential Diagnosis

Although most cases of CLL will fit the description above, there are some features that may lead to an altered appearance. These are listed in Table 44.3 and illustrated in Fig. 44.3.

TABLE 44.3. MORPHOLOGIC VARIABILITY IN B-CELL CHRONIC LYMPHOCYTIC LEUKEMIA

Smears
Downey II appearance
Prolymphocytes
Nuclear contour irregularity (folds, clefts)
Hairy cytoplasm
Sections
Prolymphocytes and growth centers
Clefted cells
Reticulin fibrosis

Up to 15 of cases of CLL may have numerous cells with a morphologic appearance similar to Downey II-like reactive (atypical) lymphocytes (Fig. 44.3A) (9). These cells are more readily identified on rapidly dried, thinly smeared blood films than on marrow smears. In CLL, these Downey II-like lymphocytes usually have neither cytoplasmic granules nor the marked morphologic variability seen in reactive processes. The importance of the finding of Downey II-like lymphocytes in CLL is to recognize it as a variant and not report the cells as “atypical/reactive” lymphocytes. Because CLL is a disease of individuals more than 30 to 40 years of age, CLL should be considered in a reactive-appearing lymphocytosis in older patients.

Many cases of CLL will have rare cells with the appearance of prolymphocytes, and occasional cases may have numerous prolymphocytes at presentation (Fig. 44.3B). Lymphocytes with nuclear contour irregularity in the form of folding or clefting also can be seen in CLL. In addition, there may be other atypical cells with larger than normal nuclear size or abundant eccentric cytoplasm suggesting plasmacytoid morphology. When greater than 10% prolymphocytes are present or >15% clefted, large or plasmacytoid lymphocytes are present, the CLL is diagnosed as “atypical CLL” (10, 11). The prolymphocyte in CLL as illustrated in Fig. 44.3B has partly dispersed, partly clumped chromatin, increased cytoplasmic basophilia, and a prominent single nucleolus with perinucleolar chromatin clumping. Prolymphocytes in CLL are morphologically similar to the prolymphocytes of PLL (PLL)(see later, Fig. 44.7). However, in CLL there may be more variability than in PLL because of the presence of large blastlike prolymphocytes called reticular lymphoblasts or paraimmunoblasts (9, 12). As shown in Table 44.4, more than 25% of cases of CLL at diagnosis will have greater than 5% prolymphocytes on smears (12). When the percentage of prolymphocytes exceeds 10%, the case is said to represent an atypical CLL with prolymphocytes (CLL/PL), and when the percentage

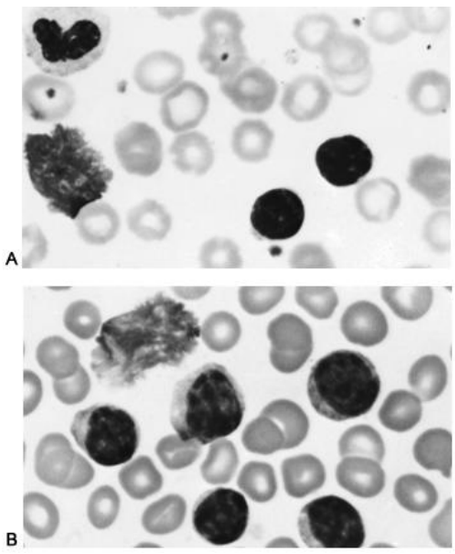


FIGURE 44.1. Chronic lymphocytic leukemia (CLL) lymphocytes on blood smear. Lymphocytes almost as small as erythrocytes (**A**) are rarely seen in CLL. Lymphocytes in CLL are more typically medium sized like those shown in **B**. Note the exaggerated chromatin clumping in some of the lymphocytes. Wright's stain 1200X. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:325–357).

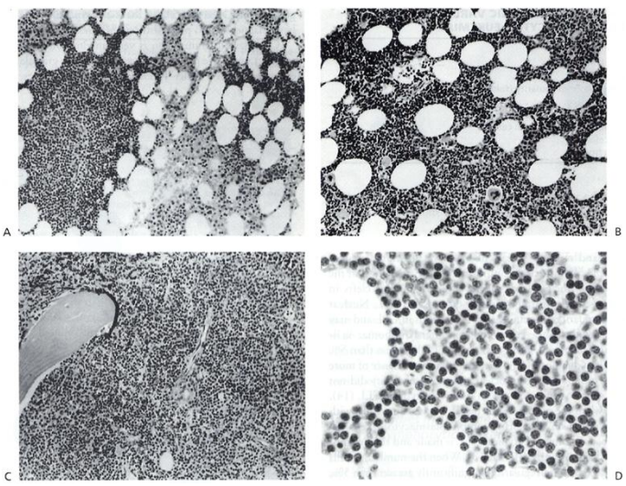


FIGURE 44.2. Chronic lymphocytic leukemia (CLL) bone marrow sections. The focal pattern shown in **A** is characterized by aggregates that have spreading of lymphocytes into and surrounding fat cells at the periphery. This spreading characteristic is especially evident in the aggregate in the upper right field of **A**. An interstitial pattern shown in **B** is characterized by diffuse infiltration of lymphocytes throughout the marrow space with many residual fat cells and admixed normal marrow elements. Extensive diffuse replacement of the marrow is shown in **C**. On high power examination in **D**, note minimal nuclear contour irregularity, clumped chromatin, uniformity, and absence of mitotic activity. Hematoxylin and eosin, 120X (**A,B,C**), 600X (**D**). (**A**, From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:325–357).

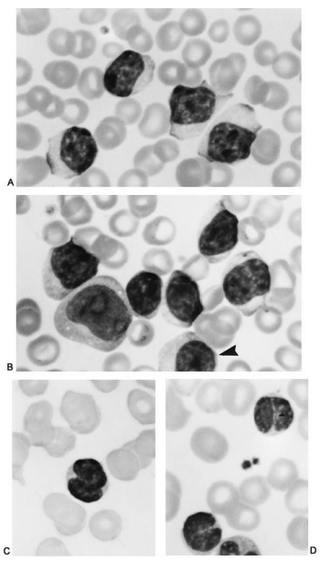


FIGURE 44.3. Morphologic variability in the peripheral blood of chronic lymphocytic leukemia (CLL). **A:** Downey II-like lymphocytes are large lymphocytes with abundant cytoplasm and peripheral basophilia. **B:** Prolymphocytes in CLL. Note the size gradation in this field from small lymphocytes to medium-sized lymphocytes to a typical prolymphocyte (arrow) to a large blastlike prolymphocyte. **C, D:** Nuclear contour irregularity in lymphocytes in CLL. Note the folded nucleus in **C** and the clefted nuclei in **D**. Clefted cells in CLL have more cytoplasm than the small cleaved cells of follicle-center lymphoma. Wright's stain 1200X. (**A** and **D**, From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:325–357).

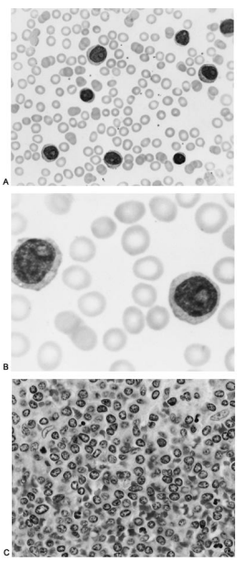


FIGURE 44.7. A, B: Typical prolymphocytes in the blood of prolymphocytic leukemia. **C:** Prolymphocytes on sections in prolymphocytic leukemia may have the uniform appearance illustrated here, or may show more variability in cell size with admixed mature lymphocytes and blastlike prolymphocytes. Wright's stain 1200X (**A, B**). Hematoxylin and eosin, 600X (**C**). (**A** and **B**, From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:325–357)

of prolymphocytes at diagnosis exceeds 55%, an alternate diagnosis of a more aggressive disease, PLL, should be considered (10, 11, 12 and 13).

TABLE 44.4. PERCENTAGE OF PROLYMPHOCYTES ON INITIAL DIAGNOSTIC FILMS OF CLL[a]

Percentage of Prolymphocytes	Number of Cases	Percentage of Cases
< 5	198	74
6–10	40	15
11–20	24	9
> 20	4	2
Total CLL cases	266	100

CLL, Chronic lymphocytic leukemia

[a] Data from 266 cases of CLL with peripheral blood lymphocytes >15,000/mm^3 seen at the University of Iowa Hospitals and Clinics from 1960 to 1975. Differentials were done on initial diagnostic bone marrow and/or peripheral blood smears, and the percentage of prolymphocytes is per 100 lymphoid cells. Marker studies were performed on only a small percentage of the cases and some T-cell CLLs may be included. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984; 1: 325–357)

The cells with nuclear contour irregularity that can be seen in the blood and marrow smears of CLL are illustrated in Fig. 44.3C and Fig. 44.3D. This nuclear contour irregularity usually is in the form of subtle nuclear folds; however, sharp nuclear clefts in some cells also are consistent with the diagnosis of CLL. Nuclear contour irregularity also can be seen on sections in CLL and may be confused with mantle-cell or follicle-center lymphoma. As illustrated in Table 44.5, 85% of cases of CLL have less than 5% of such cells at diagnosis (12). In one study, the presence of more than 5% lymphocytes with nuclear contour irregularity did not appear to alter the prognosis of an otherwise typical CLL (14). However, when the number of clefted cells plus larger cells with eccentric cytoplasm suggesting lymphoplasmacytoid cells exceeds 15%, a diagnosis of atypical CLL is made and the prognosis is altered significantly (10, 11, 13). When the number of cells with nuclear contour irregularity is significantly greater than 5%, an alternative diagnosis of a leukemic phase of a non-Hodgkin's lymphoma such as mantle-cell lymphocytic lymphoma, follicle-center lymphoma, Sézary syndrome, or a chronic lymphoproliferative disorder of T-cell type should be considered. The distinction between an atypical CLL with increased nuclear contour irregularity and a leukemic phase of mantle-cell lymphocytic lymphoma may be very difficult, but flow cytomery may be helpful (see Table 44.23 in the discussion of lymphomas) (1, 6, 7, 8, 15). Leukemic mantle-cell lymphoma like CLL is CD5 positive but is usually CD23 negative. Cyclin D1 on section material also may be helpful (16). Follicle-center lymphoma and splenic lymphoma with villous lymphocytes (SLVL) when presenting in a leukemic also may be confused with CLL but they are both CD5 and CD23 negative (1, 8).

TABLE 44.5. PERCENTAGE OF LYMPHOCYTES WITH IRREGULAR NUCLEAR CONTOUR ON INITIAL DIAGNOSTIC FILMS OF CLL[a]

Percentage Irregular Nuclear Contour	Number of Cases	Percentage of Cases
< 5	225	85
6–10	30	11
11–20	11	4
Total CLL cases	266	100

CLL, chronic lymphocytic leukemia

[a] Data from 266 cases of CLL with peripheral blood lymphocytes >15,000/mm^3 seen at the University of Iowa Hospitals and Clinics from 1960 to 1975. Differentials done on initial diagnostic bone marrow and/or peripheral blood smears. The percentage irregular nuclear contour is per 100 lymphoid cells. Marker studies were performed on only a small percentage of the cases and some T-cell CLLs may be included. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:325–357)

TABLE 44.23. IMMUNOLOGIC FINDINGS USEFUL IN SUBCLASSIFYING NON-HODGKIN'S LYMPHOMAS IN THE BONE MARROW

Lymphoma Type	Characteristic Phenotype
B-Cell	
Small lymphocytic	Clonal Surface Ig+ (weak), CD5+, CD23+, CD43+
Lymphoplasmacytoid	Clonal Cytoplasmic Ig+ (in plasmacytic cells), CD5+/-, CD23-
Mantle cell	Clonal Surface Ig+ (stronger), CD5+, CD23-, Cyclin-D1+
Marginal zone B cell	Clonal Surface Ig+, CD20+, CD11c+/-, CD43-/+, CD5-, CD23-
Follicle center	Clonal Surface Ig+ (intense), CD5-, CD23-, CD43-, CD10+
Diffuse large B cell	Clonal Surface Ig+, CD20+
Lymphoblastic	Precursor B-cells without SIg or CD20, but with CD10 (CALLA), CD19 and TDT
Burkitt's	Clonal Surface Ig+, CD20+
T-Cell	
Peripheral T cell	T-cells (of mature phenotype) with CD2, CD5, CD4 or CD8
Lymphoblastic	Precursor T-cells with TDT

Some cases of CLL will show a mixture of atypical features (Downey II-like cells, prolymphocytes, and cells with nuclear contour irregularity), giving a very heterogeneous appearance to a disease that is typically said to be characterized by monotony of cell type. A combination of the data from Table 44.4 and Table 44.5 (12) indicate that 45 of 266 cases (17%) had greater than 10% prolymphocytes and/or cells with nuclear contour irregularity. The diagnosis of atypical CLL in these cases still is tenable, although it may be necessary to use immunologic marker studies or lymph-node biopsy to exclude other diagnoses. It may be useful to enumerate the number of prolymphocytes and cells with nuclear contour irregularity on peripheral blood differentials of patients with chronic lymphoproliferative disorders, because greater than 10% to 15% of these cell types in CLL appear to have diagnostic and prognostic significance.

Cytoplasmic granulation is very uncommon in B-cell CLL, and if granules are present, a diagnosis of large granular lymphocytosis, discussed under T-cell and NK-cell disorders, or a reactive

process should be considered. Immunologic marker studies should assist in the differential diagnosis.

Occasionally, the cells in cases of CLL (as well as PLL or low-grade lymphomas) will show artifactual hairy cytoplasmic projections (Fig. 44.4). A repeat smear or bone-marrow biopsy may help solve the problem; however, a tartrate-resistant acid phosphatase (TRAP) stain or immunologic marker analysis also may be helpful to distinguish CLL from hairy-cell leukemia or splenic lymphoma with villous lymphocytes.

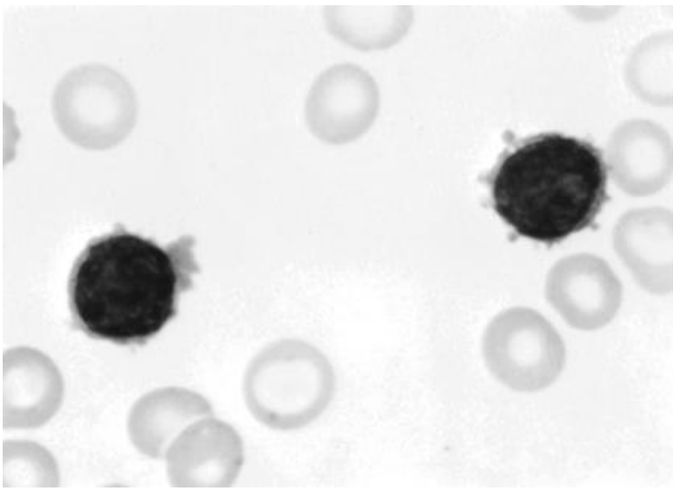


FIGURE 44.4. Artifactual hairy cytoplasmic projections in peripheral blood cells from chronic lymphocytic leukemia. The cell at the right also has a prolymphocytic appearance. Wright's stain 1200X.

Growth centers are frequently present in the lymph nodes of patients with CLL (17, 18) and less often on bone-marrow sections (illustrated in Fig. 44.5). Growth centers are thought to be the sites of proliferation of cells in CLL because many of the cells in growth centers are large and have basophilic cytoplasm, dispersed chromatin, nucleoli, and increased mitotic activity. These cells on sections are prolymphocytes that correspond to the prolymphocytes seen on blood and marrow smears. Prolymphocytes also can be diffusely distributed throughout the marrow or be admixed with small mature cells with clefted nuclei. This may lead to confusion with lymphoma on section material.

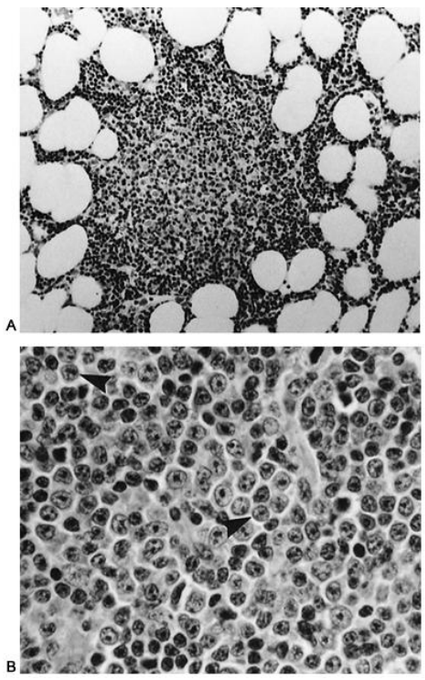


FIGURE 44.5. A: Growth center in the bone marrow of chronic lymphocytic leukemia. Note the central pallor of the aggregate resulting from focal increase in prolymphocytes. **B:** Increased concentration of prolymphocytes from a growth center. Note the gradation in size from mature lymphocytes to prolymphocytes with partly dispersed chromatin (arrowheads), to large blastlike prolymphocytes with dispersed chromatin, corresponding to the cells on smears illustrated in Fig. 3B. Hematoxylin and eosin, 120X (**A**), 600X (**B**). (**A**, From Dick, FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:325–357)

Rare cases of CLL will aspirate poorly and show increased marrow reticulin (19). This finding should not cause confusion with hairy-cell leukemia, which characteristically shows increased reticulin. A TRAP stain and immunophenotyping may assist in the diagnosis if cytology and histology are not characteristic of either diagnosis.

Special Studies In The Diferential Diagnosis of Chronic Lymphocytic Leukemia

Immunologic marker analysis as outlined above in Table 44.2 and in Table 44.23 in the discussion of lymphomas below, is a useful adjunct to the diagnosis of CLL, especially in cases with low WBC or in the differential diagnosis for other lymphoproliferative disorders (1, 6, 7 and 8, 15, 20, 21, 22 and 23).

High-resolution agarose gel electrophoresis and immunofixation have shown that about 42% of patients with CLL will have a small monoclonal serum immunoglobulin, usually IgM or free light chain type (24). An additional 14% will have only free light chains in the urine. In 5% or fewer cases of CLL, the amount of monoclonal protein is large enough to be identified on conventional serum protein electrophoresis (2). A monoclonal protein on serum protein electrophoresis in CLL may be associated with a worsened prognosis, and if the WBC is only mildly elevated, the presence of a significant IgM monoclonal paraproteinemia should stimulate a morphologic reevaluation of the case for plasmacytoid differentiation to rule out Waldenström's macroglobulinemia, which is discussed in a later section.

Of special interest is the development of a warm autoimmune hemolytic anemia in about 15% of patients with CLL (2). In addition, depression of normal immunoglobulins is seen in CLL, leading to immune deficiency (2).

Over half of patients with CLL will have a clonal chromosomal abnormality after culture with B-cell mitogens. The most common finding is trisomy 12 (10, 11, 25, 26). Survival is adversely affected in patients with trisomy 12. In addition, atypical CLL is correlated with trisomy 12. DNA content studies are nearly normal in a majority of cases (27, 28), and antigen-receptor gene rearrangement studies show clonal rearrangement, as

would be expected of a B-cell neoplasm (29). Rarely, the additional finding of a T-cell antigen receptor rearrangement may be seen in B-cell CLL (30).

Change with Progression and Transformation

Although CLL is a disease of the mature lymphocyte, it does not always remain morphologically mature or remain functional at the early B-cell stage. The various transformations and secondary tumors that can arise in CLL are shown in Table 44.6.

TABLE 44.6. TRANSFORMATION AND SECONDARY TUMORS IN CLL

Prolymphocytic transformation
Cleaved-cell transformation
Large-cell lymphoma (Richter's syndrome)
Plasma-cell myeloma
Acute lymphoblastic leukemia
Hodgkin's disease
Acute nonlymphocytic leukemia

There may be a progressive increase in prolymphocytes resulting in morphology similar to atypical CLL or even PLL (see Fig. 44.3B) (13, 31, 32). Increasing numbers of clefted cells may accompany an increase in prolymphocytes, and rarely, increased nuclear clefting may be the predominant feature (Fig. 44.3C, Fig. 44.3D). When there is a sustained progressive change in morphology towards atypical CLL with increase in the percentage of atypical cells, patient survival and response to therapy are adversely affected (13).

The cell of CLL has the ability to undergo a dramatic morphologic transformation to a cell that is much larger than a prolymphocyte, and is morphologically typical of large-cell lymphoma (17, 32, 33). There can be marked variability in this transformation from one case to the next, as illustrated in Fig. 44.6. The term “Richter's syndrome” has been applied to this transformation. Large-cell lymphoma arising in CLL frequently occurs at an extra marrow site; however, the bone marrow and occasionally the blood may be involved. In rare instances, lymphoma with characteristics indistinguishable from Hodgkin's disease may occur in CLL (35, 36).

CLL may rarely present with concomitant plasma-cell myeloma or may subsequently undergo apparent transformation to plasma-cell myeloma, with production and secretion of immunoglobulin (37). Transformation to an entity typical of acute lymphocytic leukemia is rare (38).

When another morphologic type of lymphoproliferative disorder arises in CLL, it does not automatically imply that that tumor has arisen from the same clone; it may be a new tumor. This is most likely the case for Hodgkin's disease and myeloproliferative disorders arising in CLL (39). Special studies such as chromosome analysis or gene probe analysis showing the exact pattern in both tumors will confirm that they are derived from the same clone. The appearance of a new pattern observed by these techniques suggests that either a new tumor has developed or clonal evolution of the original tumor has occurred (33, 34, 38).

Staging, Therapy, and Prognostic Features

The National Cancer Institute-Sponsored Working Group recommends a simplified scheme for the staging of CLL (Table 44.7) (3). A more complicated staging scheme was developed by Rai in 1975, which is still used by some clinicians (3).

Morphologic features also have been shown to determine prognosis in CLL. As stated in the previous section, increasing numbers of prolymphocytes on smears alters prognosis. In addition, patients with more than 15,000 prolymphocytes per cubic millimeter at presentation do significantly worse than those with 15,000 or fewer prolymphocytes per cubic millimeter independent of the percentage of prolymphocytes (13). The infiltrative pattern of marrow involvement in CLL has been correlated with survival. Patients with a focal or interstitial infiltrate (i.e., less extensive involvement) have been shown to have a better prognosis and a lower stage of disease than patients with diffuse extensive marrow involvement (40). Also, as mentioned in the

previous section, chromosomal abnormalities are associated with prognosis (10, 26).

CLL is essentially a noncurable disease; thus, therapy is directed at treatment of symptoms and palliation. Patients with early-stage disease and low WBC count may be observed without therapy. If therapy is necessary because of massive adenopathy, splenomegaly, anemia, decreased platelet count, or significantly increasing white cell count, single-agent chemotherapy is used. As the disease becomes nonresponsive, multiagent therapy is instituted (2, 3).

Prolymphocytic Leukemia

PLL is an uncommon form of leukemia that is closely related to CLL (13, 41, 42 and 43). Similar to CLL, it predominantly affects males and has a median age of onset in the sixth to the seventh decade. It differs from CLL however, in that there is a predominance of prolymphocytes. PLL frequently presents with a very high WBC count, anemia, thrombocytopenia, and splenomegaly; it differs from CLL in that adenopathy is less conspicuous in PLL. Clinically, PLL is more aggressive than CLL, with a median survival of approximately 2 to 3 years. A majority of cases of PLL are of B-cell type. PLL of T-cell type will be discussed in a later section.

Diagnostic Features

The morphology of the cell in the peripheral blood of PLL is quite characteristic. Prolymphocytes are large lymphoid cells with abundant, moderately blue cytoplasm (Fig. 44.7). The nuclear chromatin is partly condensed and partly open with a reticular pattern. A single prominent nucleolus is present and frequently is accentuated by perinucleolar chromatin clumping. The exact percentage of prolymphocytes in the blood needed to confirm a diagnosis of PLL is not well defined; however, more than 55%, with the remainder being more mature lymphocytes, is a proposed minimum (13). The bone marrow is frequently extensively infiltrated. When the marrow is not extensively involved, the distribution may be focal and/or interstitial, similar to CLL (42, 44). On high-power magnification (Fig. 44.7C), the lymphoid cells have the appearance of prolymphocytes in CLL; however, more uniformity usually is seen than in atypical CLL with increased prolymphocytes. Cases of PLL that do not have markedly elevated WBCs have been reported (45). These may be cases of PLL that are diagnosed early in the course of the disease. When the diagnosis is in question, immunologic marker studies may be of help. B-cell PLL shows strongly reacting SIg with a restricted light chain compared with the weak-reacting SIg of CLL. If additional evidence is needed, PLL frequently is CD5 negative, CD23 negative, and is positive with FMC7 and CD79b (Table 44.2) (1, 6, 7, 20).

Morphologic Variability and Differential Diagnosis

In some cases of PLL, the peripheral blood prolymphocytes may show considerable numbers of smaller lymphoid cells with more condensed chromatin and prominent nucleoli. This variant needs to be distinguished from CLL, which also may occasionally show evident but not large nucleoli. Blastlike prolymphocytes with completely dispersed chromatin also may be a prominent feature in some cases of PLL, suggesting a diagnosis of acute

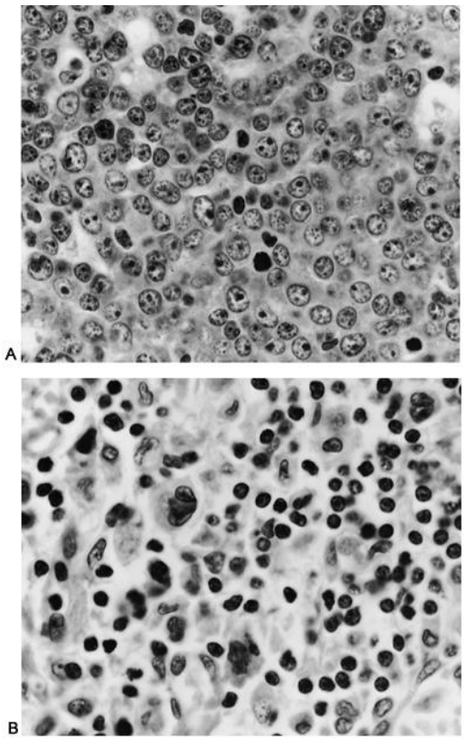


FIGURE 44.6. Large-cell lymphomas (Richter's syndrome) arising in chronic lymphocytic leukemia (CLL). **A:** Cells in this lymphoma from a lymph node biopsy are the size of blastlike prolymphocytes. Note the marked uniformity of this tumor compared to the tumor in **B**. **B:** This lymphoma from a patient with CLL shows marked pleomorphism of the large cells (with admixed residual CLL lymphocytes). This form of Richter's syndrome can be mistaken for Hodgkin's disease. Hematoxylin and eosin, 600X. (B., From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In Koepke JA, ed Laboratory Hematology. New York: Churchill Livingstone 1984;1:325-357)

TABLE 44.7. CLINICAL STAGING IN CLL

Stage	Criteria	Percentage of Total	Survival (yrs)
A	Less than three sites of involvement[a]	55	[b]
B	Three or more sites of involvement	30	7
C	Anemia (< 10 gm/dL) or thrombocytopenia (< 100 × 10^9/L)	15	2

[a] Sites of involvement include spleen, liver, cervical nodes, axillary nodes, and inguinal nodes.

[b] Survival for stage A is not significantly different from the age matched general population.

(Adapted from Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:325–357)

leukemia. Granules or vacuoles may be present in some cases. Very large, blastlike prolymphocytes may cause confusion with peripheral blood involvement by large-cell lymphoma or monoblastic leukemia. Some cases of PLL will show a greater spectrum of cells, ranging from mature lymphocytes with marked clumping of chromatin to more blastlike cells. These cases with variability in cell type suggest the possibility that the process has evolved from a CLL through prolymphocytic transformation, and cases with morphologic features intermediate between CLL and PLL may be very difficult to clearly diagnose as either PLL or CLL/PL because CLL/PL may show immunophenotypic features intermediate between CLL and PLL (13). Cases of PLL with a moderate amount of nuclear contour irregularity will suggest the possibility of a leukemic transformation of a lymphoma. Rare cases of PLL (as well as CLL and mantle-cell lymphoma) will have cytoplasmic projections, suggesting the possibility of a variant of hairy-cell leukemia or SLVL. Finally, TRAP stain may be positive in PLL and thus should not be used alone to distinguish PLL from hairy-cell leukemia.

Other Special Studies

Serum protein studies performed on PLL have shown occasional cases with monoclonal gammopathy (43). Clonal chromosomal abnormalities are frequently seen in PLL cells using B-cell mitogens. A 14q+ was the most frequently observed abnormality in B-cell PLL in one study (46). Catovsky's group has reported a high incidence of p53 mutation in PLL relative to CLL and other B-cell malignancies perhaps explaining its aggressive biology (47). Despite the partially transformed appearance of the cells in PLL and the chromosomal abnormalities, the cellular DNA content is normal in most cases, and there is a low percentage of cases in S-phase (48).

Change With Progression and Transformation

Because PLL is uncommon and also is a tumor of relatively transformed lymphoid cells, transformation to a large-cell process is very rare.

Staging, Therapy, and Prognostic Features

A widely accepted staging system similar to CLL does not exist for PLL; however, there are occasional cases of PLL with a low WBC that appear to be early-stage disease (45). Patients with PLL do not respond very well to conventional CLL management. Combination chemotherapy will induce remission in many patients; however, remission usually is short-lived.

Hairy-Cell Leukemia

Hairy-cell leukemia is an uncommon disease characterized by a proliferation of medium-sized lymphoid cells with hairy cytoplasmic projections (49). The disease predominantly involves the bone marrow and spleen, and the patient may or may not have a “leukemic” blood picture. There is a male predominance with a median age of onset in the sixth decade. The patient frequently presents with nonspecific constitutional symptoms, pancytopenia,

and splenomegaly, and the disease usually follows a very indolent course. Virtually all cases are of B-cell origin.

Diagnostic Features

The most characteristic diagnostic feature of hairy-cell leukemia is the appearance of hairy cells on blood and marrow smears (Fig. 44.8). Hairy cells range in size from that of a medium-sized lymphocyte to a monocyte. The cytoplasmic border has a frayed or shaggy appearance, and the cytoplasm is blue-gray with a variegated dark and light consistency. Cytoplasmic granules are uncommon. The nucleus generally is round to oval or slightly indented with a partly clumped, partly open or reticular chromatin, and a single indistinct nucleolus. The WBC usually is decreased, but hairy cells are found easily in the blood in most cases. The TRAP stain is classically positive in hairy-cell leukemia. Typical flow cytometry features are shown in Table 44.2 (1, 6, 7 and 8, 20). Typically, hairy cells are CD5 and CD23 negative and are positive for CD25 and CD11c. CD103 also may be used in difficult cases.

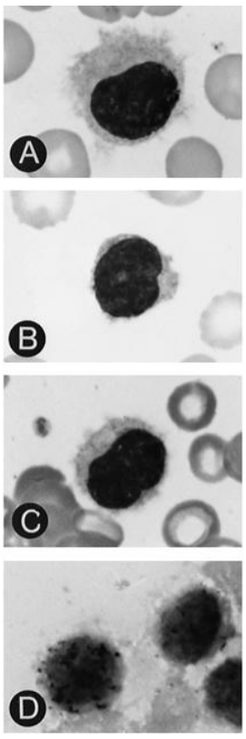


FIGURE 44.8. A, B, C: Typical hairy cells from the blood or marrow. **D:** Tartrate-resistant acid phosphatase (TRAP) stain. Note dispersion of positive granules throughout the cytoplasm. Wright's stain 1200X (**A**,**B**,**C**). TRAP stain 1200X (**D**).

In rare cases, it may be impossible to identify hairy cells on the blood and marrow smears, whereas other cases may have a markedly elevated WBC count with numerous hairy cells. A monocytopenia is characteristically seen in the peripheral blood and is a good differential diagnostic feature. The bone marrow in hairy-cell leukemia usually does not aspirate freely, and there may be a “dry tap.” Thus, in some cases with a dry tap and very few hairy cells in the peripheral blood, cells showing the characteristic hairy-cell morphology may be very difficult to find, and sufficient material may not be available for flow cytometry. Also, unfortunately, characteristic hairy-cell morphology is not as readily identified on touch preparations of bone marrow as on aspirates. In these cases, the diagnosis must rest more heavily on the appearance of cells on the sections and a TRAP stain on rare cells that aspirated or come from touch preps. The use of monoclonal antibodies on section material may be helpful if there are no cells on the blood or aspirate and touch preps (50, 51).

On bone marrow sections, the marrow usually is moderately to extensively replaced by hairy cells (Fig. 44.9). When the marrow is less extensively replaced, there is an interstitial pattern with varying numbers of normal marrow elements admixed. Characteristic hairy cells on sections have a bland-appearing lymphoid nucleus with mild to moderate nuclear contour irregularity and abundant pale cytoplasm between widely separated nuclei. Occasionally, the marrow will show some crush artifact resembling fibrosis (Fig. 44.9D). The spindle-shaped cells in this situation are hairy cells that have been strung out by the encasing reticulin fibers after being crushed. Although overt fibrosis is rarely seen on sections in hairy-cell leukemia, background reticulin usually is moderately increased on reticulin stain (Fig. 44.9E).

Morphologic Variability and Differential Diagnosis

Most cases of hairy-cell leukemia fit the description given above. However, some cases are difficult to diagnose as hairy-cell leukemia because of deviation from the norm (52, 53, 54, 55, 56, 57 and 58). Some of these appearances are listed in Table 44.8 and are illustrated in Fig. 44.10.

Hairy-cell leukemia can be mistaken for a number of other processes. Cases with elevated WBC and small hairy cells with round nuclear contour (Fig. 44.10A) can be mistaken for CLL (58). Rare cases of hairy-cell leukemia have abundant, clear nonhairy cytoplasm, giving them an exaggerated “fried egg” appearance of a Downey II/atypical lymphocyte (Fig. 44.10B, Fig. 44.10C) (53). Cases with large cells and more prominent nucleoli can be mistaken for PLL or even acute lymphoblastic leukemia (Fig. 44.10D) (52, 54). Cases with a monocytoid appearance can be mistaken for monocytic leukemia (Fig. 44.10E) (52). Cases with multilobulated nuclei can be mistaken for lymphoma (Fig. 44.10F) (55, 56), and cases with minimal marrow involvement and hairy cells in the 5% to 20% range and no obvious infiltrate in sections may be mistaken for normal.

Patients with splenomegaly, pseudofibrosis on sections, and minimal numbers of hairy cells on blood and marrow smears can be misdiagnosed as having agnogenic myeloid metaplasia (52). Patients with hypoplastic marrow and minimal splenomegaly can be misdiagnosed as having aplastic anemia (57). A focal or paratrabecular infiltrative pattern can be misinterpreted as evidence of lymphoma.

Of special note is the existence of a small percentage of cases that are midway morphologically and immunologically between PLL and hairy-cell leukemia and do not respond to conventional therapy for hairy-cell leukemia (54). There is also a set of cases with some morphologic and immunologic overlap between CLL and hairy-cell leukemia (58).

Cases of PLL and non-Hodgkin's lymphoma with hairy cytoplasm (SLVL) should be distinguished from hairy-cell leukemia (59, 60). In addition to marker studies, a helpful differential diagnostic feature is that most other lymphoproliferative disorders do not have the abundant pale cytoplasm in combination with the bland appearance of the nucleus seen in hairy-cell leukemia on sections. In addition, the infiltrate in hairy-cell leukemia is more interstitial or diffuse in the bone marrow and localized in the red pulp of the spleen. Lymphomas are distributed in a more focal pattern in the bone marrow and are more prone to have white pulp distribution in the spleen.

TRAP and Other Special Studies

Hairy cells are characterized by TRAP positivity (Fig. 44.8D) (61, 62). Many normal hematopoietic cells (including lymphoid cells) are acid-phosphatase positive; however, the reaction in most normal and neoplastic cells is inhibited by tartaric acid. Thus, a positive TRAP stain has become a cytochemical marker for hairy-cell leukemia. Unfortunately, a positive TRAP stain is not diagnostic, because a number of other processes may be TRAP positive. Also, the cells of some cases of hairy-cell

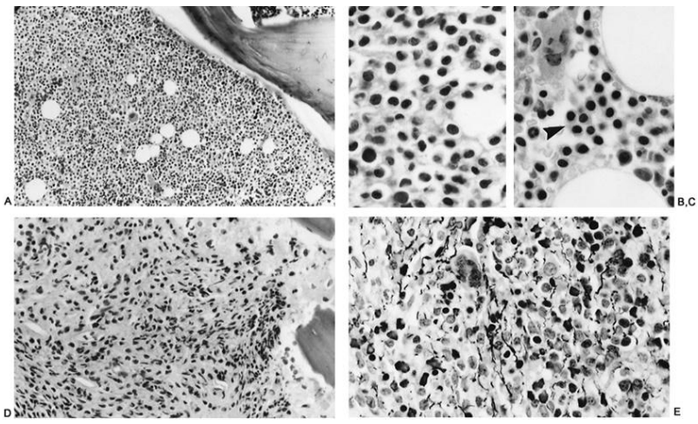


FIGURE 44.9. Trephine biopsy sections in hairy-cell leukemia. **A:** Note the diffuse infiltrate of hairy cells with scattered normal marrow elements. **B:** Note the mild-to-moderate nuclear contour irregularity. The halo around the hairy cells is from artifactual retraction in formaldehyde. **C:** Note the small collection of hairy cells in the center of field (arrow) from a marrow with minimal infiltration. These cells, which were fixed in B-5 fixative, do not show as much retraction artifact as formaldehyde fixation. **D:** Crush artifact, a frequent occurrence in hairy-cell leukemia, may resemble marrow fibrosis. **E:** Reticulin is frequently extensively increased even though overt fibrosis is not seen. Hematoxylin and eosin, 120X (**A**), 600X (**B,C**), 240X (**D**), Reticulin stain 600X (**E**).

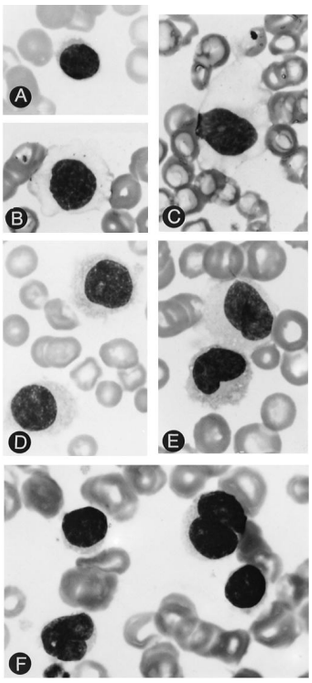


FIGURE 44.10. Hairy-cell morphologic variants. **A:** Lymphoid appearance. **B, C:** Hairy cells with non-hairy/Downey II-like cytoplasm. **D:** Prolymphocyte appearance. Note the prominent nucleolus in the cell in the upper right field. **E:** Monocytoid appearance. Note nuclear folds. **F:** Multilobular variant. Note multilobular cell in the upper right field and a deeply clefted cell in the lower left field. Wright's stain 1200X.

TABLE 44.8. MORPHOLOGIC VARIABILITY IN HAIRY-CELL LEUKEMIA

Smears
Lymphoid appearance
Nonhairy cytoplasm
Downey II appearance
Prolymphocytic appearance
Monocytoid appearance
Multilobular variant
Minimal involvement
Sections
Pseudofibrosis (crush artifact)
Hypoplastic bone marrow
Focal or paratrabecular infiltrate
No obvious infiltrate

leukemia react only weakly, or only rare strongly positive cells may be found. However, the TRAP stain is a good screening test when the diagnosis of hairy-cell leukemia is in question.

Patients with hairy-cell leukemia usually have a polyclonal gammopathy; however, rare cases of monoclonal gammopathy occur. This finding, in addition to the association of PCA-1 with hairy-cell leukemia, suggests to some investigators that hairy-cell leukemia may be more closely related to the plasma cell than other chronic lymphoproliferative disorders (63).

Many other sophisticated studies have been performed on hairy-cell leukemia, including cytogenetics (64), DNA content (27), electron microscopy (65, 66) and gene rearrangement (29); however, the studies of greatest diagnostic importance are the TRAP stain and immunologic markers.

Change with Progression and Transformation

Transformation and secondary hematologic malignancies in hairy-cell leukemia are rare (67, 68 and 69). Rarely, skeletal lesions, including osteosclerosis and lytic bone lesions or massive lymphadenopathy, may develop (70, 71 and 72).

Staging, Therapy, and Prognostic Features

Hairy-cell leukemia is an indolent disease and, if diagnosed in an asymptomatic patient, can be followed for many years without therapy. There are several therapeutic modalities available to treat symptomatic hairy-cell leukemia, which are capable of inducing long-term remission of the symptoms of disease. Splenectomy may help for chemotherapy failure. Treatment of hairy-cell leukemia by the types of chemotherapy used in other B-cell malignancies can result in a poor outcome; thus, it is very important to avoid erroneously diagnosing hairy-cell leukemia as some other B-cell process.

Chronic Lymphoproliferative Disorders of T-Cell Type

The chronic T-cell processes listed in Table 44.1 are a heterogeneous group of rare disorders with varied clinical presentation and prognosis (1, 73, 74, 75, 76 and 77). Unless these processes are thought of and studied with markers to determine their T-cell nature, they may not be readily diagnosed. Features that should lead to the suspicion that a patient with lymphocytosis in the blood or marrow has a T-cell lymphoproliferative disorder rather than a more common form of chronic B-cell leukemia include: skin lesions, cytoplasmic granules, cytopenia or lymphocytosis out of proportion to the extent of marrow involvement, and nuclear contour irregularity. Nuclear contour irregularity, when present in a T-cell process, may show more complexity (convoluted, cerebriform, or lobulated) (Fig. 44.11 and Fig. 44.12) than the simple clefting or folding of the nucleus seen in the B-cell processes that have nuclear contour irregularity.

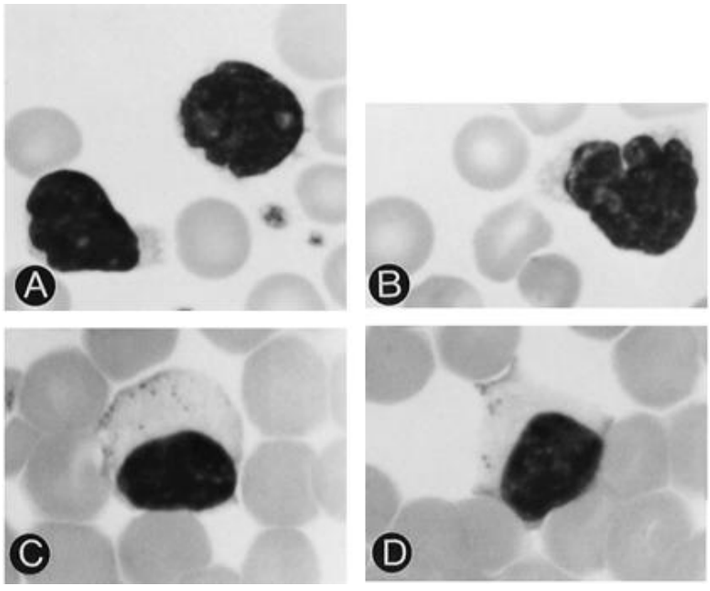


FIGURE 44.11. A, B: Cells of a chronic T-cell leukemia. Note the complex nuclear contour in the cells illustrated. **C, D:** Large granular lymphocytes. Note the abundant cytoplasm and multiple cytoplasmic granules. Wright's stain 1200X. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:325–357)

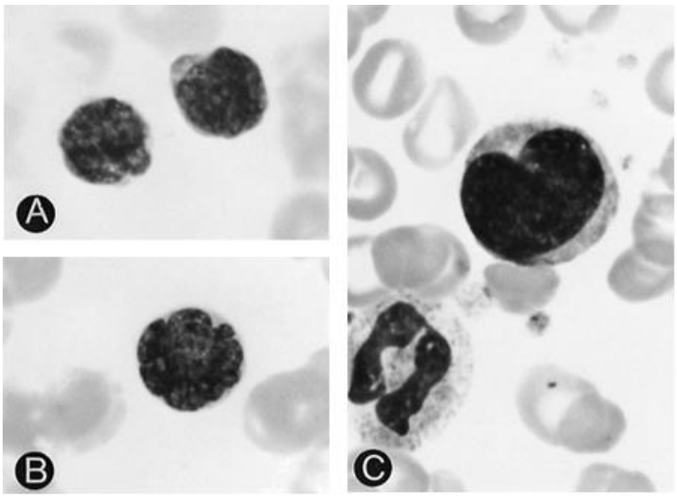


FIGURE 44.12. Cutaneous T-cell lymphoma cells in the peripheral blood. **A**, **B:** Sezary cells. Note complex internal nuclear convolutions. **C:** A large convoluted cell in the blood from a patient with disseminated mycosis fungoides. Wright's stain 1200X.

T-Cell Chronic Lymphocytic Leukemia

T-cell CLL is a very rare disease (76, 78, 79 and 80). It is seen in older individuals and frequently is associated with splenomegaly and

skin involvement. Cases of T-cell CLL may have cells with nuclear contour irregularity similar to the cells in Figure 44.11A and Figure 44.11B. However, they also may have a lymphocytosis morphologically indistinguishable from typical B-cell CLL. Earlier literature included large granular lymphocyte (LGL) leukemia as T-cell CLL; however LGL now is recognized as an entity distinct from T-cell CLL. Most cases of T-cell CLL are of helper T-cell phenotype (CD4); however, rare cases of cytotoxic/suppressor (CD8) CLL have been described (78, 79). A clonal immuno-marker such as the restricted light chain seen in CLL is not available to confirm clonality in T-cell lymphoproliferative disorders. Thus, in the differential between T-cell CLL (or other T-cell lymphoproliferative disorders) and a reactive T-cell lymphocytosis, molecular studies for T-cell clonality may be helpful (75). In addition, an abnormality involving chromosome 14 (q11;q32) is a frequent finding (76, 78).

T-cell CLL is less responsive than B-cell CLL to therapy. Because of the biologic similarity of T-cell CLL to T-cell PLL, some authors have suggested that T-cell CLL is merely a small-cell variant of T-cell PLL and recommend that they be lumped together as one entity, namely T-cell PLL (80).

T-Cell Prolymphocytic Leukemia

A T-cell phenotype is seen in approximately 20% of PLLs. T-cell PLL is similar clinically and morphologically to B-cell PLL, except that T-cell PLL is more likely to have adenopathy and skin infiltrates, and the prolymphocytes may have a higher nuclear cytoplasmic ratio, a greater degree of nuclear contour irregularity, and less prominent nucleoli than B-cell PLL (73, 74 and 75, 81). In addition, a small-cell variant of T-cell PLL has been described, which may overlap with T-cell CLL as described above (80). Similar to B-cell PLL, a small subset may pursue a more indolent course (82). A majority of cases of T-cell PLL are of helper phenotype (CD4); however, cases of cytotoxic/suppressor (CD8) phenotype also are reported (75). Similar to T-cell CLL, a clonal chromosomal abnormality involving chromosome 14 (q11;q32) is seen in a majority of cases (83).

Large Granular Lymphocyte Leukemia

LGL leukemia usually is an indolent disease characterized by a proliferation of mature-appearing lymphocytes with abundant cytoplasm and cytoplasmic granules (Fig. 44.11C and Fig. 44.11D) (1, 73, 74, 75 and 76, 84, 85). LGLs normally comprise about 15% of peripheral blood lymphocytes and they may be elevated above this level in reactive processes as well as in LGL leukemia. There are two types of LGL leukemia: the T-cell type, which is the most common, and the NK type (8, 75). The lymphocytes of the T-cell type express CD2, CD3, CD8, CD16, and are +/- for CD57. LGL leukemia cases with the NK-cell phenotype are distinguished from the T-cell type in that they are +/- for CD56 and CD57 and don't express CD3. The usual case of LGL leukemia presents with neutropenia and/or red-cell aplasia, mild splenomegaly, a mild peripheral blood lymphocytosis, and relative marrow sparing by the lymphoid infiltrate, which is focal, nonparatrabecular, or interstitial. Occasional cases are associated with rheumatoid arthritis or a positive rheumatoid factor. Although this disorder may not have a markedly elevated WBC count or an infiltrative character, most cases with a T-cell phenotype have been shown to be clonal, “neoplastic” processes by molecular techniques (75), a feature that may be useful in distinguishing this disorder from transient proliferations of LGLs (86). Cytogenetic studies have shown evidence of clonality in some cases; however, cytogenetic evaluation has been hampered because of the poor response of granular lymphocytes to mitogenic stimulation (87).

LGL also may be the cell of origin for rare cases of acute

leukemia and lymphoma (87, 88 and 89). These cases show more cytologic immaturity than typical LGL leukemia.

Sézary Syndrome

Sézary syndrome and mycosis fungoides are closely related, uncommon disorders of older individuals, and together they constitute the category called cutaneous T-cell lymphomas (73, 90). Sézary syndrome is primarily a disorder of the skin with secondary blood involvement. It is characterized by diffuse erythroderma resulting from infiltration of mature-appearing helper T cells with markedly convoluted nuclear contours. The disease shows varying degrees of lymphocytosis in the peripheral blood with Sézary cells, as illustrated in Fig. 44.12A and Fig. 44.12B. The bone marrow is relatively spared (91, 92). The diagnosis usually rests on a combination of clinical and morphologic features. A skin biopsy will show a bandlike infiltrate of small, convoluted lymphoid cells immediately beneath the epidermis. When the skin biopsy is not diagnostic, the identification of convoluted lymphocytes (Sézary cells) in the blood may be used to help substantiate the diagnosis. A small percentage of convoluted cells indistinguishable from the neoplastic Sézary cells (similar to those shown in Fig. 44.12A) can be seen in benign skin disorders; however, the presence of more than 15% to 20% of these cells or the presence of larger convoluted lymphoid cells similar to those shown in Fig. 44.12C and Fig. 44.12D are supportive of the diagnosis of Sézary syndrome (93, 94). The cells of Sézary syndrome as well as mycosis fungoides are T-helper (CD4+) cells and usually are CD7 negative (95). Immunophenotyping may be of some assistance in the differential diagnosis of Sézary syndrome with other T-cell malignancies involving the skin. When the differential diagnosis rests between Sézary syndrome and a reactive process, and skin biopsy and blood morphology are equivocal, molecular techniques for clonal T-cell gene rearrangement may provide help (96).

Mycosis fungoides characteristically lacks morphologic evidence of blood involvement and shows multifocal skin lesions rather than the diffuse involvement of Sézary syndrome. However, late in the course of mycosis fungoides with dissemination, large, atypical, convoluted lymphoid cells similar to those shown in Fig. 44.12 may be seen in the blood.

Adult T-Cell Leukemia/Lymphoma

Adult T-cell leukemia/lymphoma is an aggressive T-cell disorder caused by human T-lymphocytotrophic virus 1 (HTLV-1) (8, 73, 90, 97). It is seen predominantly in southwestern Japan, but also has been described in the Caribbean, the southeastern United States, and elsewhere. Morphologically, it is characterized by a proliferation of small to medium-sized lymphoid cells with marked nuclear lobulation similar to the cells illustrated in Fig. 44.11A and Fig. 44.11B. These cells infiltrate lymph nodes, bone marrow, blood, and skin. There may be relative marrow sparing compared with the involvement in the peripheral blood. While a majority of patients demonstrate a “leukemic pattern,” a subset of patients present with more of a “lymphoma-type” disease (97). Lytic bone lesions and hypercalcemia are common. The T-cells in adult T-cell leukemia/lymphoma are of helper (CD4) phenotype and usually are CD7 negative (8, 75, 97). The diagnosis usually is made by a combination of clinical, morphologic, and immunologic studies and is confirmed by antiviral antibodies or identification of the virus. This disorder usually has a very rapidly progressive course and responds poorly to aggressive combination chemotherapy, with a median survival of 1 year or less; however, some cases may follow a more indolent course.

IMMUNOPROLIFERATIVE DISORDERS

Part of "44 - Chronic Lymphoproliferative Disorders, Immunoproliferative Disorders, and Malignant Lymphoma"

Immunoproliferative disorders are neoplastic proliferations of plasma cells and B lymphocytes at the terminal stage of functional differentiation, and are characterized by the production and secretion of a monoclonal immunoprotein called a monoclonal gammopathy. “Disease” in immunoproliferative disorders may be caused by proliferation of the neoplastic cell and/or by secretion of an abnormal protein that has a pathologic effect. These disorders are almost exclusively in older individuals and their clinical course may be indolent or rapidly progressive. The immunoproliferative disorders include: plasma-cell myeloma, plasmacytoma, Waldenström's macroglobulinemia, heavy-chain disease, benign monoclonal gammopathy, and primary amyloidosis.

Plasma-Cell Myeloma

Plasma-cell myeloma is a relatively common hematologic disease characterized by proliferation of clusters and sheets of neoplastic plasma cells primarily in the bone marrow, resulting in focal bone lesions, diffuse osteoporosis, bone pain, fractures, and cytopenia (98, 99). The neoplastic proliferation usually produces a monoclonal gammopathy, which may result in hyperviscosity syndrome, coagulation abnormalities, immune deficiency, renal failure, or development of amyloidosis. Plasma-cell myeloma is seen primarily in older individuals (in the sixth to seventh decade) with a male predominance. It generally is a progressive malignant disorder; however, its course is variable, and prognosis is based to a large degree on the tumor burden present at diagnosis.

Diagnostic Features

The diagnosis of plasma-cell myeloma is dependent on a combination of clinical and laboratory features. From a laboratory standpoint, the diagnosis of plasma-cell myeloma should be suspected whenever there is a monoclonal gammopathy, and the bone marrow contains increased atypical plasma cells with a prominent single nucleolus and partially dispersed chromatin, sometimes referred to as “myeloma cells” (Fig. 44.13 and Fig. 44.14). The following morphologic changes are consistent with a diagnosis of plasma-cell myeloma:

1. The presence on marrow smears of 10% or more atypical plasma cells (Fig. 44.13A and Fig. 44.14A). These cells may infiltrate in an interstitial pattern, in multiple small clusters, or in sheets. The median percentage of atypical plasma cells on smears in patients with plasma-cell myeloma is about 30%. It is important to perform an actual differential (100 or more cells on three different

slides) to quantitate the number of plasma cells; plasma cells can easily be overestimated! Also, binucleation of plasma cells should not be considered an atypical feature; reactive plasma cells also are frequently binucleated.

2. Less than 10% atypical plasma cells on marrow smears in association with large sheets of atypical plasma cells on crush preparation or sections, with plasma cells representing nearly 100% of the cell population in the sheets (Fig. 44.13B) (100, 101).

Although these morphologic features are consistent with a diagnosis of plasma-cell myeloma, additional clinical and laboratory evidence of disease also may be necessary to make a definitive diagnosis. Additional evidence may be a monoclonal gammopathy in serum or urine, characteristic bone lesions, or, rarely, suppression of other immunoglobulins in the absence of a monoclonal gammopathy.

Caution should be taken in making an unequivocal diagnosis of plasma-cell myeloma in a patient with atypical plasma cells and a monoclonal gammopathy in the absence of any other evidence of disease. A small subset of patients who have 10% to 20% atypical plasma cells on marrow smears and a monoclonal gammopathy, but who also are asymptomatic, may not pursue an aggressive course. These cases have been diagnosed as indolent or smoldering myeloma (102, 103, 104 and 105).

A significant number of patients have up to 10% atypical plasma cells in the marrow with a monoclonal gammopathy and do not have disease manifestations of plasma-cell myeloma. These cases usually pursue a benign course and are called benign monoclonal gammopathy, or alternatively, monoclonal gammopathy of undetermined significance (MGUS). Benign monoclonal gammopathy is discussed in greater detail later in this chapter.

Less than 10% atypical plasma cells or more than 10% normal-appearing plasma cells can be seen in patients with progressive myeloma; however, these findings are more consistent with a benign monoclonal gammopathy or a reactive process, respectively. When these morphologies are present, the diagnosis of plasma-cell myeloma should rest even more heavily on clinical criteria.

Because patients can have progressive plasma-cell myeloma with fewer initial morphologic changes than those described earlier,

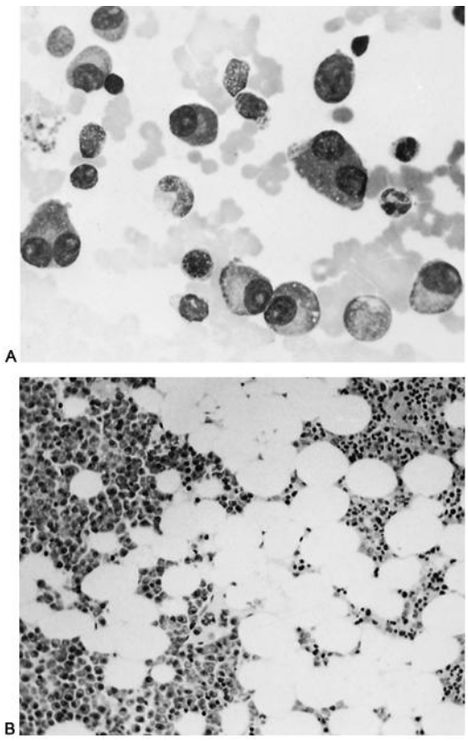


FIGURE 44.13. Plasma-cell myeloma. **A:** Marrow smear. Note immature plasma cells with a single prominent nucleolus. **B:** Bone marrow section. Note sheets of plasma cells in the upper and lower left field, and normal marrow elements to the right. Wright's stain 600X (**A**). Hematoxylin and eosin 120X (**B**). (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

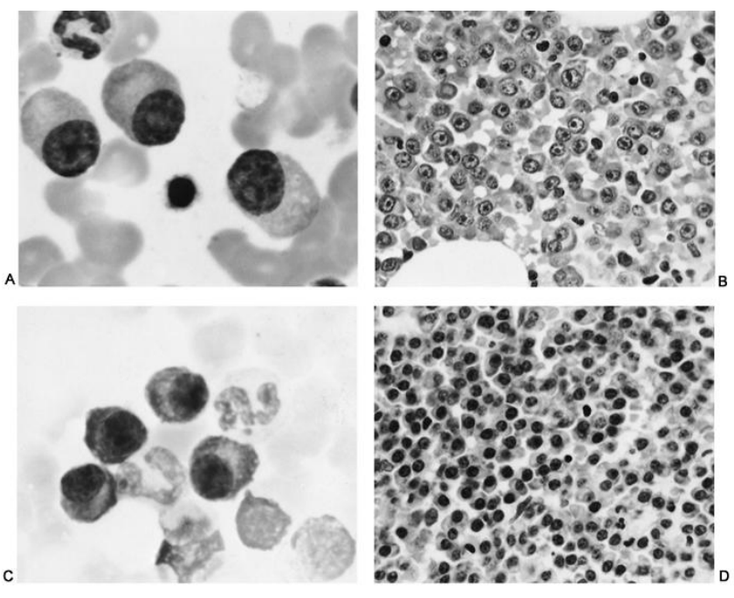


FIGURE 44.14. Plasma-cell myeloma. Marrow smears and sections. **A, B:** Myeloma with immature plasma cells having abundant eccentric basophilic cytoplasma and a prominent single nucleolus. **C, D:** Myeloma with smaller, more mature plasma cells. Chromatin is still partly dispersed but prominent nucleoli are not evident. Wright's stain 1200X (**A, C**). Hematoxylin and eosin, 600X (**B, D**). (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

or more rarely, because progressive disease may not develop when the morphologic features described above are found, several groups have developed a set of clinical pathologic criteria that serve to ensure that patients entered on protocols are diagnosed uniformly (105). A modification of one of these is illustrated in Table 44.9. Such criteria also are useful for routine diagnosis outside of protocol studies, because they emphasize the importance of using a combination of radiologic, clinical, morphologic, and laboratory features in arriving at the diagnosis of plasma-cell myeloma. These clinical-pathological features are helpful in distinguishing plasma-cell myeloma from an indolent or smoldering myeloma, and from solitary plasmacytoma and benign monoclonal gammopathy. Although these criteria may be very useful in some patients, sound judgment should be used when making the diagnosis of plasma-cell myeloma based only on these criteria. For example, if the criteria of Durie and Salmon are too rigidly adhered to, some patients with progressive plasma-cell myeloma may be excluded, or rare patients with plasmacytoma or benign monoclonal gammopathy and clinical findings caused by an unrelated disease may be included.

TABLE 44.9. CLINICAL PATHOLOGIC CRITERIA FOR INCLUSION IN STUDIES OF PLASMA-CELL MYELOMA

Major Criteria
I Plasmacytomas by biopsy
II > 30% marrow plasmacytosis
III Monoclonal gammopathy
> 3.5 g/dL of IgG
> 2.0 g/dL of IgA or
>1 gm/day κ or λ chains in urine without other significant proteinuria
Minor Criteria
A. 10%–30% marrow plasmacytosis
B. Monoclonal gammopathy with values less than in III
C. Lytic bone lesions
D. Suppressed normal immunoglobulins
< 50 mg/dL IgM,
< 100 mg/dL IgA or
< 600 mg/dL IgG
Symptomatic patient and
I plus B, C, or D
II plus B, C, or D
III
A, B, and C
A, B, and D

[From Dick FR. Chronic Lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984; 1:445–481. Modified from the criteria of Durie and Salmon (105).]

Occasional cases of reactive plasmacytosis with polyclonal gammopathy will show a significant infiltrate in the marrow with plasma cells, sometimes in sheets (101). The author of this chapter has observed two cases of reactive plasmacytosis in which there were over 50% reactive plasma cells with atypical plasma cells in a cellular bone marrow (Fig. 44.15). This emphasizes the need to identify laboratory and clinical criteria other than plasmacytosis greater than 10% or “sheets” of plasma cells in the bone marrow to establish the diagnosis of plasma-cell myeloma.

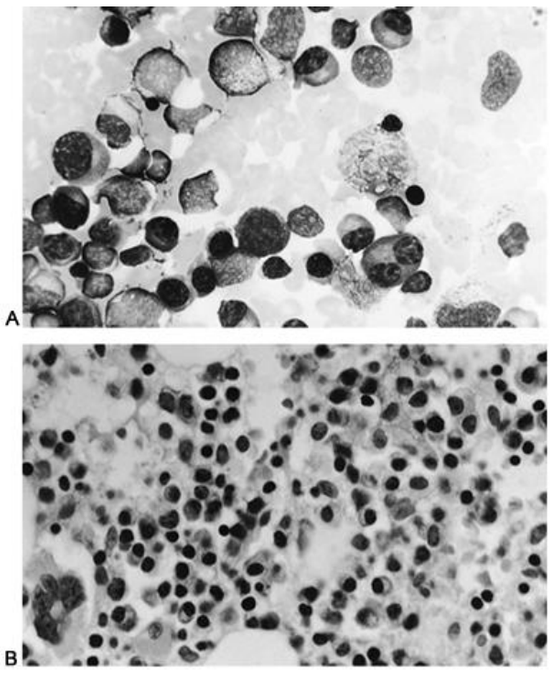


FIGURE 44.15. Bone marrow aspirate **A:** and biopsy **B:** showing a reactive plasmacytosis of greater than 50%. Note large and binucleated plasma cells. Single prominent nucleoli are not a prominent feature, however. Wright's stain 600X (**A**). Hematoxylin and eosin, 600X (**B**). (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

Morphologic Variability and Differential Diagnosis

The marrow in a classic case of plasma-cell myeloma shows a predominance of immature myeloma cells, as illustrated in Fig. 44.13 and Fig. 44.14A and Fig. 44.14B. Less commonly, smaller plasma cells with more mature nuclear features predominate (Fig. 44.14C and Fig. 44.14D). Other cases may show a spectrum of plasma cells ranging from nearly normal plasma cells to more classic myeloma cells (106).

Cytologic variability from the pattern described above occasionally occurs and should be recognized as part of the spectrum of plasma-cell myeloma (Table 44.10).

Flame cells, thesaurocytes, and inclusions can be seen in some of the plasma cells in a benign plasmacytosis; however, when they are present in plasma-cell myeloma, they usually are present in a major percentage of the plasma cells (Fig. 44.16). Flame cells and thesaurocytes (Fig. 44.16A and Fig. 44.16B) are frequently but not exclusively associated with an IgA monoclonal gammopathy (106).

A wide variety of cytoplasmic and nuclear inclusions may be seen in the cells of plasma-cell myeloma, some of which are illustrated in Figure 44.16C and Figure 44.16D (100, 107, 108, 109 and 110).

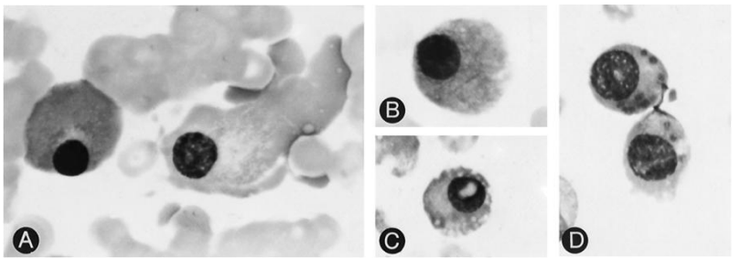


FIGURE 44.16. Variants of myeloma cells, bone marrow aspirate. **A:** Flame cells. The dark flowing cytoplasmic borders on these plasma cells are red or "flaming" on Wright's stain. **B:** Thesaurocytes are plasma cells with abundant, pale blue, foamy or reticulated cytoplasm. **C:** Dutcher bodies on smears are pale intranuclear inclusions. **D:** Russell bodies are reddish, globular, cytoplasmic inclusions. Wright's stain 1200X. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

Occasional cases may show a pleomorphic, a plasmablastic, or anaplastic appearance, as illustrated in Fig. 44.17. A myeloma with pleomorphic morphology has marked variability in cell size with large multinucleated and bizarre forms (106). Plasmablastic, anaplastic (or immunoblastic) morphology overlap (111, 112, 113 and 114). Cases with these morphologies show poor differentiation toward plasma cells, may have less abundant cytoplasm, and may have absence of eccentric cytoplasm with paranuclear pallor (hof). They are morphologically similar to, and may be difficult to distinguish from, a large-cell lymphoma with immunoblastic or plasmacytoid features, especially if the patient also has extramarrow masses (114).

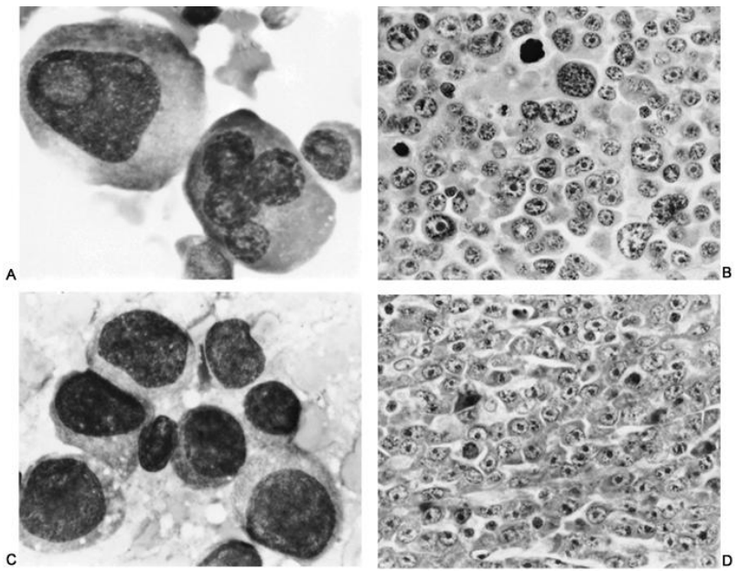


FIGURE 44.17. Variants of myeloma cells. Bone marrow aspirate and sections. **A, B:** Pleomorphic morphology in plasma-cell myeloma. Note multinucleation and large bizarre forms. **C, D:** Plasmablastic or anaplastic morphology. These plasma cells are difficult to recognize as plasma cells, and are morphologically similar to a large cell lymphoma. Wright's stain 1200X (**A, C**). Hematoxylin and eosin, 600X (**B, D**). (**A, B**, and **C**, From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

In rare cases of plasma-cell myeloma, the plasma cells in the bone marrow may have a lymphoid morphology, as illustrated in Fig. 44.18A (106, 115). These cases may be difficult to separate from lymphoma or Waldenström's macroglobulinemia if lytic lesions characteristic of plasma-cell myeloma are not present or if the patient has extra-marrow disease. Also, plasma cells, when present in the blood in typical plasma-cell myeloma, frequently have a more lymphoid appearance than in the marrow of the same patient.

Rare cases of myeloma will show considerable nuclear irregularity (multilobulated or monocytoid), as illustrated in Fig. 44.18B (116). On sections as well as on smears, these cells may be difficult to recognize as plasma cells. Rare cases of plasma-cell myeloma will slow phagocytic activity by the myeloma cells (117).

The marrow in plasma-cell myeloma usually aspirates quite freely; however, occasional cases may aspirate poorly because of increased reticulin or fibrosis; and rare cases may show osteosclerosis on radiograph and biopsy (118, 119 and 120).

Rarely, amyloidosis is present in the bone marrow in plasma-cell myeloma (121).

Plasma cells in small numbers can be seen in the peripheral blood at diagnosis in approximately 15% of cases of plasma-cell myeloma, and can be shown to be present in a much larger proportion of patients using special techniques to identify circulating plasma cells (122, 123). In about 2% of cases of plasma-cell myeloma, large numbers of plasma cells will be present in the blood (Fig. 44.19). When this occurs, the patient is said to have plasma-cell leukemia, which has been defined as having more than 2,000 plasma cells per cubic millimeter and more than 30% plasma cells in the peripheral blood (122). Plasma-cell leukemia has a younger age distribution, a greater incidence of hepatosplenomegaly, less clinical bone involvement, and a poorer prognosis than the usual case of plasma-cell myeloma. Again, plasma-cell leukemia also is more likely to show a lymphoid appearance than classic plasma-cell myeloma (106, 122).

Immunologic and Other Special Studies

A search for a monoclonal gammopathy is the single most important test that can be done in a patient suspected of having plasma-cell myeloma (Table 44.11 and Table 44.12) (124). The best procedure is to perform immunofixation electrophoresis on serum and urine as a screening test. If no monoclonal gammopathy is detected by these two procedures, it is highly unlikely that the patient has myeloma, because 1% or fewer cases of plasma-cell myeloma fail to produce at least a small monoclonal protein. Serum protein electrophoresis alone is not a good screening test for a monoclonal gammopathy because it is not as sensitive as immunofixation electrophoresis. Also, in 20% of patients

TABLE 44.10. MORPHOLOGIC VARIABILITY IN PLASMA CELL MYELOMA

Smears
Flame cells
Thesaurocytes
Inclusions
Pleomorphic cells
Plasmablastic or anaplastic cells
Lymphoid cells
Monocytoid or multilobated
Phagocytosis
Sections
Reticulin
Fibrosis
Osteosclerosis
Amyloid
Plasma-cell leukemia

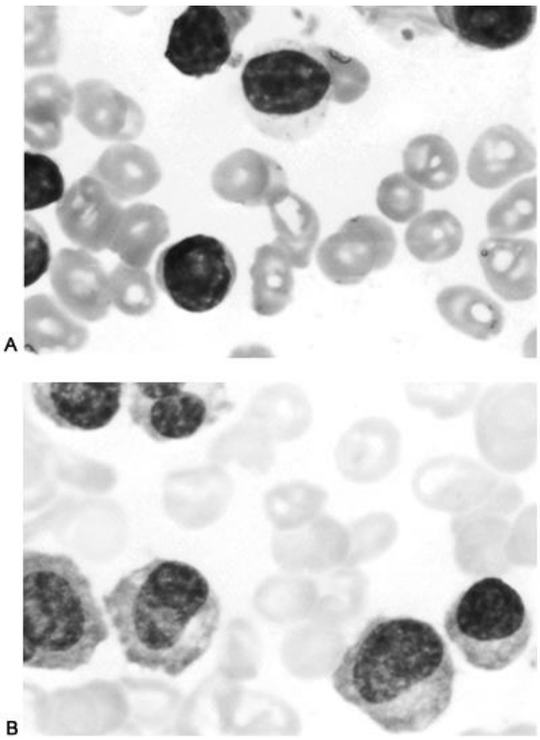


FIGURE 44.18. Variants of myeloma cells. **A:** Lymphoid morphology. **B:** Monocytoid or multilobated morphology. The marrow aspirates in both of these illustrations are from individuals with lytic bone lesions and IgG monoclonal gammopathy. Wright's stain 1200X. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

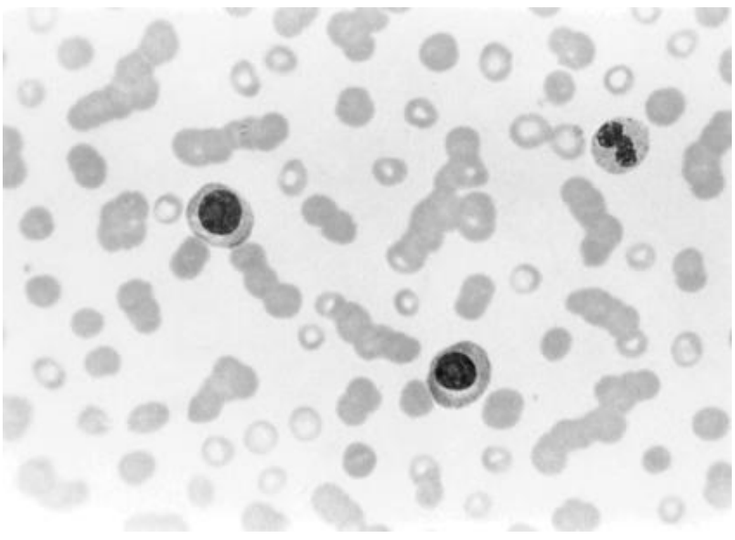


FIGURE 44.19. Plasma-cell leukemia (peripheral blood). Note the marked rouleaux and immature plasma cells. Wright's stain 600X. (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

with myeloma, the monoclonal gammopathy will be detected by urinalysis and not by serum protein electrophoresis. A serum-protein electrophoresis or quantitative immunoglobulins will, however, help detect those rare patients with myeloma who have hypogammaglobulinemia and no monoclonal gammopathy on serum immunofixation electrophoresis. As pointed out in the previous section, many patients with a monoclonal gammopathy (especially a small monoclonal gammopathy) will not have progressive plasma-cell myeloma. Thus, other features in addition to a monoclonal gammopathy are necessary to make the diagnosis of plasma-cell myeloma.

TABLE 44.11. FREQUENCY OF TYPE OF MONOCLONAL GAMMOPATHY IN PLASMA CELL MYELOMA

Monoclonal Gammopathy	Approximate Frequency of Occurrence (%)
IgG	57
IgA	20
κ or λ only	20
IgD	1
Nonsecretory	1
Biclonal	1
IgE	Rare
IgM	Rare
	100

(From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984; 1:445–481.)

TABLE 44.12. OTHER PROTEIN FINDINGS IN PLASMA-CELL MYELOMA (102)

Finding	Approximate Frequency of Occurrence in Plasma Cell Myeloma (%)
Size of monoclonal peak	
< 2 gm/dL	17
> 5 gm/dL	20
No obvious monoclonal peak on serum protein electrophoresis	20
Free light chains in urine	80
Hypoglobulinemia and no M-peak on serum electrophoresis	9
“Normal” serum protein electrophoresis	10–15
Location of monoclonal peak on protein electrophoresis	
Gamma	70
Beta	30
Alpha 2	Rare
Cryoglobulin	5
Pyroglobulin	1

(From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984; 1:445–481.)

Immunologic evaluation of the cells in myeloma is rarely needed because the morphology of the cells in myeloma usually is that of classical immature “myeloma cells,” and clonality is supported by the finding of a monoclonal gammopathy in the serum or urine. However if evaluated, the plasma cells in plasma-cell myeloma will show cytoplasmic immunoglobulin with clonally restricted light chain (125, 126 and 127). The plasma cells in plasma-cell myeloma generally are negative or SIg and B cell-associated antigens but do demonstrate plasma-cell–associated antigens.

Immunostaining of plasma cells in the bone marrow may be useful in distinguishing benign plasma cells from nonsecretory

TABLE 44.13. TRANSFORMATION AND SECONDARY TUMORS IN PLASMA-CELL MYELOMA

Plasma-cell leukemia
Pleomorphic, plasmablastic, or anaplastic myeloma
Immunoblastic sarcoma
Chronic lymphocytic leukemia
Extraosseous lesions
Myelodysplastic syndrome
Acute nonlymphocytic leukemia

myeloma, but may not clearly distinguish benign monoclonal gammopathy and amyloidosis with increased plasma cells from early myeloma. This is because plasma cells in the marrow of patients with benign monoclonal gammopathy and amyloidosis may also show a restricted light chain (125, 126, 127 and 128). However, in benign monoclonal gammopathy, the κ:λ ratio usually is less altered from normal than in plasma-cell myeloma (125).

Numerous other special studies have been performed in plasma-cell myeloma, including serum β_2-microglobulin (129, 130), plasma-cell labeling indices (131, 132), chromosome analysis (133, 134), ploidy analysis (135), and growth fraction with Ki-67 (136). Some of these may have potential prognostic importance.

Change with Progression and Transformation

A small percentage of patients with plasma-cell myeloma may show some morphologic change with time. These changes are listed in Table 44.13.

About one third of patients who qualify for the diagnosis of plasma-cell leukemia have transformed from a more typical presentation of plasma-cell myeloma (122). The remainder is *de novo*, as described earlier. Transformation to plasma-cell leukemia shows many clinical features in common with *de novo* plasma-cell leukemia.

Pleomorphic, plasmablastic, or anaplastic morphologies of myeloma can be seen at diagnosis; however, they also may develop in association with aggressive disease progression and extraosseous spread (32). A variety of other terms have been applied to this aggressive transformation, including dysplastic myeloma and immunoblastic sarcoma. These forms of transformation are probably derived from the same clone as the original tumor.

CLL associated with plasma-cell myeloma usually is diagnosed simultaneously; however, rare cases of CLL developing after a diagnosis of plasma-cell myeloma have been reported (137).

The increased incidence of myelodysplastic syndrome and acute nonlymphocytic leukemia in myeloma usually is attributed to prior alkalating agent therapy; however, there are isolated reports of nonlymphocytic leukemia developing in association with plasma-cell myeloma without prior therapy (138, 139). Therapy-related myelodysplastic syndrome and acute nonlymphocytic leukemia have much poorer prognoses than *de novo* disease.

Staging, Therapy, and Prognostic Features

Prognosis in plasma-cell myeloma is based primarily on the clinical staging scheme outlined in Table 44.14(140). The features listed in the table, which determine the stage of the patient, in general correlate with the tumor burden present. About 15% of patients present as stage I disease, and most of these have a creatinine level of 2 mg/dL or less (status A). Stage II comprises 30% of patients, and only 5% to 10% of these have renal status B. More than half of the patients present with stage III disease, and half of these have a creatinine level above 2 mg/dL (status B).

TABLE 44.14. CLINICAL STAGING OF PLASMA CELL MYELOMA (119)

Stage I	All of the following:
	Hg > 10 gm/dL
	Calcium > 12 mg/dL
	Normal bone on radiograph or solitary lesion only
	Monoclonal protein
	IgG < 5 gm/dL
	IgA < 3 gm/dL
	Urine light chains < 4 g/24 h
Stage II	Fitting neither stage I nor III
Stage III	One or more of the following:
	Hg < 8.5 gm/dL
	Calcium > 12 mg/dL
	Advanced lytic bone lesions
	Monoclonal protein
	IgG > 7 gm/dL
	IgA > 5 gm/dL
	Urine light chains > 12 g/24 h
Renal Status A = creatinine < 2 mg/dL	
Renal Status B = creatinine > 2 mg/dL	

(From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:445–481.)

In addition to the staging system, the following features of plasma-cell myeloma have been shown to be associated with a higher stage and worsened prognosis: elevated β_2-microglobulin (129, 130) and labeling indices (130), abnormal karyotype or DNA content (133, 134 and 135), immature or plasmablastic plasma cell morphology (111, 112, 113 and 114, 141), and pattern of marrow infiltration (142). β_2-Microglobulin appears to be the most widely used clinically.

Plasma-cell myeloma is a progressive disease that, similar to CLL, is difficult to cure. The median survival of patients with stage I disease is 6 years; with stage II, 4 years; and with stage III, 1 year. Patients with renal status A live a median of 3 years, while those with renal status B live less than 6 months (140).

Plasmacytoma

Plasmacytoma is a rare disorder characterized by a mass lesion of neoplastic plasma cells (143, 144). Although plasmacytomas usually are seen in the surgical pathology laboratory rather than the hematology laboratory, bone-marrow examination and serum and urine protein studies frequently are done subsequent to the diagnosis of plasmacytoma to rule out spread to the bone marrow. Also, as mentioned in the previous section, plasma-cell myeloma may develop extramedullary plasmacytomas late in the course of disease. Plasmacytomas can be solitary or multiple and may present in the bone (spine, pelvis, femur, etc.) or as an extramedullary lesion (upper respiratory tract, lymph nodes, spleen, skin, gastrointestinal tract, etc.). As with plasma-cell

myeloma, plasmacytomas are seen predominantly in older individuals, with a male predominance.

Immunologic Studies

When plasmacytomas are localized at presentation (solitary plasmacytoma), the presence of an associated monoclonal gammopathy is uncommon and, if present, should disappear as the solitary plasmacytoma is treated. With dissemination, the incidence of a monoclonal gammopathy increases, with IgG and κ being the predominant monoclonal gammopathies seen.

Therapy and Change with Progression

Eventual dissemination of plasmacytomas is seen in about one third to one half of patients with solitary plasmacytoma of bone. Eventual dissemination usually results in the classic picture of plasma-cell myeloma. Some cases may take up to 15 years before dissemination occurs. Extramedullary plasmacytomas have a lower incidence of eventual dissemination, and dissemination of extramedullary plasmacytoma is to lymph nodes, other extramedullary sites, and bone. Treatment for solitary plasmacytoma is radiation therapy.

Waldenström's Macroglobulinemia

Waldenström's macroglobulinemia is an uncommon disease characterized by an IgM monoclonal gammopathy and proliferation of lymphocytes in the bone marrow, which show varying degrees of differentiation into plasmacytoid lymphocytes and plasma cells (145). Waldenström's macroglobulinemia is derived from a cell that is functionally intermediate between the lymphocyte of CLL and the plasma cell of plasma-cell myeloma. The disease primarily affects the bone marrow, with frequent involvement of lymph nodes, spleen, and liver. Unlike plasma-cell myeloma, lytic bone lesions are rare. The morphology of the proliferation in Waldenström's macroglobulinemia is consistent with the diagnosis of lymphoplasmacytoid lymphoma in the REAL non-Hodgkin's lymphoma (discussed later). Thus, Waldenström's macroglobulinemia can be considered the bone marrow counterpart of lymphoplasmacytoid malignant lymphoma. In Waldenström's macroglobulinemia, as in plasma-cell myeloma, disease may be caused by the infiltrative nature of the neoplastic cells or by deleterious effects of the protein abnormality. Similar to CLL and plasma-cell myeloma, Waldenström's macroglobulinemia is a disease of older individuals and has a male predominance. Its prognosis is intermediate between CLL and plasma-cell myeloma.

Diagnostic Features

As stated above, the diagnosis of Waldenström's macroglobulinemia is made from a combination of bone-marrow findings and the presence of an IgM monoclonal gammopathy. The bone marrow frequently aspirates poorly because of increased reticulin in the marrow infiltrate. On marrow smears, there is a mild to marked lymphocytosis composed of mature lymphocytes with varying numbers of plasma cells, as illustrated in Fig. 44.20. Plasmacytoid

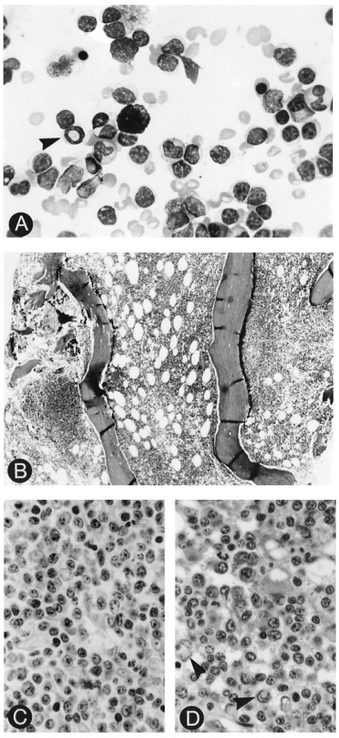


FIGURE 44.20. Waldenström's macroglobulinemia. **A:** Marrow smear. Note the numerous lymphocytes, several plasma cells, a Dutcher body (arrow), and a tissue mast cell. **B:** Trephine biopsy. In the left field, the infiltrate is focal. In the center between the two long trabeculae, it is primarily interstitial, and in the right field, it is paratrabecular. **C, D:** Note numerous lymphocytes and occasional plasma cells. In (**D**) Dutcher bodies (pale, intranuclear inclusions) are indicated with arrows. Hematoxylin and eosin, 600X (**A, C, D**), 50X (**B**). (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984;1:445–481.)

lymphocytes, which are morphologically intermediate between plasma cells and lymphocytes, may or may not be present to form a morphologic bridge between the lymphocytes and plasma cells. Tissue mast cells are increased frequently, and occasional Dutcher bodies may be observed. The peripheral blood in Waldenström's macroglobulinemia usually does not show lymphocytosis; however, there may occasionally be an elevated WBC count of up to 15,000/mm^3, with cells similar to those seen in the bone-marrow aspirate.

On bone-marrow sections, there usually are focal paratrabecular and nonparatrabecular infiltrates of mature lymphocytes with associated smaller numbers of plasma cells and plasmacytoid lymphocytes (Fig. 44.20B and Fig. 44.20C) (146, 147). The degree of involvement may range from small focal aggregates to extensive diffuse replacement, or, more rarely, there may be a diffuse interstitial infiltrate. In early marrow involvement, the marrow smears may show no increase in lymphocytes and plasma cells, and the marrow sections may show only small aggregates, which are difficult to distinguish from benign aggregates.

Morphologic Variability and Differential Diagnosis

Cases of Waldenström's macroglobulinemia may show more or less plasmacytic differentiation. This is the basis for the distinctions “lymphoplasmacytoid” and “lymphoplasmacytic” Waldenström's macroglobulinemia (146, 147). Scattered transformed cells similar to prolymphocytes of CLL, or cells with increased nuclear irregularity also may be seen in Waldenström's macroglobulinemia. Cases with increased large transformed cells (immunoblasts) have been termed “polymorphous” (146).

IgM monoclonal gammopathy can be seen in a variety of processes other than classic Waldenström's macroglobulinemia (145, 146, 149). Table 44.15 summarizes the variety of morphologies associated with monoclonal IgM. Technically, only patients with a morphology of lymphoplasmacytoid lymphoma in the bone marrow have classic Waldenström's macroglobulinemia. Patients with an IgM monoclonal gammopathy and small lymphocytic lymphoma (SLL) or CLL who have no plasmacytoid differentiation usually are not considered to have Waldenström's macroglobulinemia. The 10% lymphomas with IgM monoclonal gammopathy shown in Table 44.15 are lymphomas with more poorly differentiated morphologies such as large-cell lymphomas with plasmacytoid differentiation. These cases also should not be diagnosed as Waldenström's macroglobulinemia. It is of interest that lymphomas also can rarely have IgG monoclonal gammopathies (149).

TABLE 44.15. MORPHOLOGIC ENTITIES ASSOCIATED WITH MACROGLOBULINEMIA

Morphology	Approximate Frequency in Patients with Macroglobulinemia (%)
Lymphoplasmacytoid lymphoma	50
Small lymphocytic lymphoma or CLL	10
Other lymphomas	
diffuse	10
follicular	Rare
Other leukemias	Rare
Plasma-cell myeloma	Rare
Lymphoid aggregates in bone marrow	15
No morphologic lesion	15
	100

CLL, chronic lymphocytic leukemia (From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:445–481.)

Thirty percent of patients with a monoclonal IgM have no identifiable accompanying neoplastic lesion, but instead have benign-appearing lymphoid aggregates or no morphologic lesion. Such cases may be referred to as essential macroglobulinemia. These patients may have symptoms related to the protein abnormality (high serum viscosity, peripheral neuropathy cold agglutinin disease, coagulation disorders, etc.), or they may have no symptoms at all, the protein abnormality being identified incidentally on serum protein electrophoresis. This latter asymptomatic set of patients, who usually have less than 3 g/dL IgM monoclonal gammopathy would fit the category of benign monoclonal gammopathy (discussed in a later section).

Immunologic and Other Studies

As for plasma-cell myeloma, serum and urine immunofixation electrophoresis is of major diagnostic usefulness in patients suspected of having Waldenström's macroglobulinemia. Interestingly, a small percentage of patients with the classic clinical and morphologic features of Waldenström's macroglobulinemia have, instead of an IgM monoclonal gammopathy, an IgG or an IgA monoclonal gammopathy (150). Whether these patients should be included in the diagnostic category of Waldenström's macroglobulinemia is debatable.

Immunologic study of the cells on sections or smears in Waldenström's macroglobulinemia usually is not necessary because the monoclonal gammopathy serves as a clonal marker for the disease. When studied, the cells in Waldenström's macroglobulinemia will show cytoplasmic immunoglobulin of a clonally restricted light chain as well as B-cell surface markers (1, 8, 20, 151). Evaluation of bone-marrow cells using flow cytometry, immunofluorescence, and immunohistochemistry has demonstrated a light-chain restricted clone of cells in a majority of patients who have IgM monoclonal gammopathy but morphologically nondiagnostic bone marrow (151). This study concludes that patients with more than 20% “clonal” mononuclear cells in a morphologically nondiagnostic marrow without symptoms should be diagnosed as having lymphoplasmacytoid lymphoma (Waldenström's macroglobulinemia) rather than essential macroglobulinemia or a benign monoclonal gammopathy. Further investigation is necessary to confirm this conclusion.

Changes with Progression and Transformation

Occasional cases of Waldenström's macroglobulinemia show transformation to a more aggressive cell type, such as large-cell immunoblastic lymphoma (32). Like plasma-cell myeloma, rare cases also may develop acute nonlymphocytic leukemia after long-term therapy with alkylating agents (32).

About 20% of patients who present initially with an apparently benign IgM monoclonal gammopathy and no definitive

evidence of a lymphoid malignancy will develop a morphologically diagnosable hematologic malignancy. The median time for this to occur is slightly more than 4 years after the initial diagnosis (145).

Staging, Therapy, and Prognostic Features

Like CLL and plasma-cell myeloma, Waldenström's macroglobulinemia is a progressive, noncurable disease. It has a prognosis intermediate between CLL and plasma-cell myeloma, and the therapeutic approach is fairly similar to that for CLL. The bone marrow pattern and the degree of infiltration may have some prognostic significance, with a focal nodular pattern having a better prognosis than an extensive diffuse infiltrate (146, 152). Of the three morphologic variants described earlier, the lymphoplasmacytoid type has the best prognosis and the polymorphous type has the poorest (146, 147), however, another study could find no morphologic correlation with survival (147).

Heavy-Chain Disease

These rare disorders are characterized by a cellular proliferation of plasma cells and/or lymphoid cells and production of a monoclonal gammopathy composed of an incomplete heavy chain without an associated light chain (153). Thus, on immunofixation electrophoresis, the clonal protein reacts with antisera to heavy chains but not to light chains. Heavy-chain disease involving γ, μ, and α chains have been described (154, 155 and 156). The age range for heavy-chain disease is wider than that of the immunoproliferative disorders described in the preceding sections. Of the three disorders, α chain disease has the youngest age distribution, being seen primarily in the second and third decades. The organ distribution of the cellular proliferation depends on the heavy chain involved.

γ Heavy-Chain Disease

Clinical findings at presentation for this disorder, in addition to organomegaly and cytopenias, include autoimmune disease, swelling of the uvula, and neuropathy (153, 154). Approximately 70% of patients with a finding of γ heavy-chain disease have a neoplastic lesion, which may involve bone marrow, Waldeyer's ring, lymph nodes, spleen, and liver. The morphology of the lesion is characteristically a mixture of plasma cells, lymphocytes, and immunoblasts; however, more rarely, the lesion in γ heavy-chain disease may be identical to that of CLL, SLL-plasmacytoid, or large-cell immunoblastic lymphoma. The 30% of patients without a morphologic lesion associated with γ heavy-chain disease will have either atypical infiltrates in the aforementioned sites or no morphologically recognizable abnormal proliferation.

μ Heavy-Chain Disease

CLL and SLL are the most common morphologic entities associated with production of μ heavy chains (153, 155). Rare cases of large-cell lymphoma also may be seen. Plasma cells in the bone marrow in this process characteristically have cytoplasmic vacuoles (157). Most patients also have free light chains in the urine. This is the rarest of the heavy-chain diseases.

α Chain Disease

Patients with this disorder usually present with gastrointestinal symptoms (153, 156). This is the most common form of heavy-chain disease and is characteristically seen in the Mediterranean area. A morphologic lesion in α chain disease is most commonly observed in the gastrointestinal tract and abdominal lymph nodes. The cellular proliferation in the lesion can range from a dense infiltrate of plasma cells in the *lamina propria* of the bowel and nodes to a large-cell lymphoma. The bone marrow rarely may be involved (158).

"Benign" Monoclonal Gammopathy

Occasionally, a monoclonal gammopathy will be identified by serum protein electrophoresis during routine screening or workup of an elderly patient for diseases other than immunoproliferative disorders (159). A patient who has a small monoclonal gammopathy with no symptoms referable to the protein abnormality and no B-cell malignancy or amyloidosis is considered to have a benign monoclonal gammopathy, also called monoclonal gammopathy of undetermined significance (MGUS). The majority of these patients will follow a benign course, without developing plasma-cell myeloma, lymphoproliferative disorder, or complications from protein production.

Some patients with benign monoclonal gammopathy appear to have a true benign neoplastic disorder. Evidence for this includes persistence of the monoclonal gammopathy over time in an asymptomatic patient, and the presence of a mild (less than 10%) increase in atypical clonal plasma cells. Other patients appear to have a premalignant disorder. Evidence for this is the eventual development of plasma-cell myeloma or a related malignant immunoproliferative or lymphoproliferative disorder in up to 10% to 20% of cases. Currently, it is difficult to determine which patients with benign monoclonal gammopathy will experience progression to a malignant disorder. In other patients, the benign monoclonal gammopathy may represent a restricted response to some unidentified antigen. This is based on these findings: rare monoclonal gammopathies have known specificity; some benign monoclonal gammopathies are transient; and in a hospitalized setting, about 25% of benign monoclonal gammopathies are related to inflammatory, connective tissue, or neoplastic disease (159). The term essential monoclonal gammopathy has been applied to the first two sets of patients, and secondary monoclonal gammopathy to the set of patients in whom the monoclonal gammopathy may represent a restricted response (160). Although this distinction is conceptually attractive, it is not always possible to distinguish between an essential and a secondary disorder in a given patient.

Diagnostic Features

Bone-marrow examination, if performed in benign monoclonal gammopathy, usually shows less than 5% nearly normal-appearing plasma cells. However, up to 10% or more atypical plasma

cells also may be consistent with a benign course. The term smoldering myeloma, as discussed previously, has been used to describe the asymptomatic patient who has more than 10% atypical plasma cells,

TABLE 44.16. CLASSES OF SERUM IMMUNOGLOBULIN IN BENIGN MONOCLONAL GAMMOPATHY (140)

Ig Class	Frequency (%)
IgG	75
IgA	15
IgM	10

(From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:445–481.)

TABLE 44.17. SIZE OF M-COMPONENT IN BENIGN MONOCLONAL GAMMOPATHY (140)

Size (g/dL)	Approximate Percentage of Patients
< 1.5	30
1.6–2.5	65
> 2.5	5

(From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone 1984;1:445–481.)

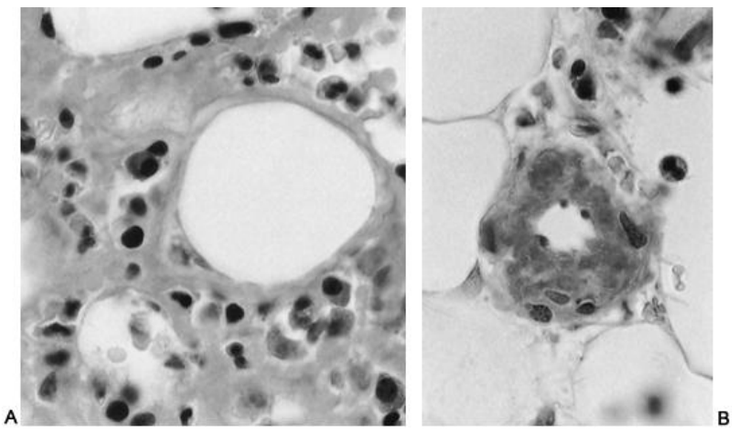


FIGURE 44.21. Primary amyloidosis, bone marrow section. **A:** The homogeneous interstitial deposits are amyloid. Note scattered plasma cells embedded in the amyloid. **B:** Deposition of amyloid in a vessel wall. Hematoxylin and eosin, 600X.

TABLE 44.18. HODGKIN'S DISEASE: CLASSIFICATION AND INCIDENCE OF BONE MARROW INVOLVEMENT

Rye Classification	Relative Frequency	Incidence of Bone Marrow Involvement
Lymphocyte predominant	14%	0%
Nodular sclerosis	41%	2–10%
Mixed cellularity	31%	5–20%
Lymphocyte depleted	14%	45–70%

has a monoclonal gammopathy above 3 g/dL, and follows a benign course (102, 103). Atypical plasma cells similar to those illustrated in Fig. 44.14 can be seen in patients with benign monoclonal gammopathy. However, markedly atypical cells should raise the possibility that the patient has significant disease at another site and will have progressive disease.

The incidence of benign monoclonal gammopathy increases with age. Screening of healthy populations using serum protein electrophoresis has shown that 1% of patients more than 25 years of age and 3% of patients more than 70 years of age will have a benign monoclonal gammopathy (161). Another study using more sensitive serum protein electrophoresis showed a 5% incidence in a healthy population less than 65 years old (162). The incidence may be even higher when using immunofixation electrophoresis.

Immunologic and Other Studies

The class of serum immunoglobulin seen in benign monoclonal gammopathy is shown in Table 44.16, and the size of the monoclonal gammopathy is shown in Table 44.17 (163). Approximately 15% of patients with benign monoclonal gammopathy will show some decrease in other immunoglobulins, and at least 10% of patients with a serum benign monoclonal gammopathy will have free light chains (usually less than 1 g per 24 hours) in the urine (159, 163). Patients with benign monoclonal gammopathy composed of free urinary light chains and no monoclonal serum protein also have been observed (159).

Immunohistochemical studies on the bone marrow in cases of benign monoclonal gammopathy have demonstrated an excess of plasma cells with a clonally restricted light chain in a majority of patients (125, 126 and 127). Thus, identification of a clonal population in this setting does not imply a malignant diagnosis.

Evaluation of $ß_2$-microglobulinemia, labeling index of bone marrow plasma cells and peripheral blood lymphocytes, cytoplasmic 5-nucleotidase, and short-term bone marrow culture have been reported in benign monoclonal gammopathy (131, 132, 159, 164, 165 and 166).

Therapy and Prognosis

Approximately 10% to 20% of patients with benign monoclonal gammopathy will eventually develop plasma-cell myeloma, Waldenström's macroglobulinemia, lymphoma, or amyloidosis (159). Although a progressive increase in the monoclonal protein level is a good indicator of the eventual development of progressive disease, a small percentage of patients who show an increase of 1 to 2 g in the monoclonal gammopathy are found not to have developed symptomatic neoplastic disease on long-term followup (159). A bone-marrow examination at the time of the discovery of benign monoclonal gammopathy may or may not be indicated, depending on other clinical and laboratory findings. There is some evidence to suggest that cases of benign monoclonal gammopathy that will progress to plasma-cell myeloma can be identified using peripheral blood labeling indices (132). $ß_2$-Microglobulin is not clearly useful in this regard (159). Benign monoclonal gammopathies should not be overinterpreted as evidence for plasma-cell myeloma; they can be seen in patients with diseases masquerading as plasma-cell myeloma, such as osteoporosis and metastatic carcinoma.

Primary Amyloidosis

Primary amyloidosis (AL amyloidosis) as opposed to secondary amyloidosis (AA amyloidosis) is associated with a clonal production of light chains with secretion in the urine (Bence-Jones protein), even when not associated with an overt malignant-cell proliferation, thus it is included under the category of immunoproliferative disorders. In addition, amyloidosis may eventually develop in about 15% of patients with plasma-cell myeloma, with an especially high propensity in patients with λ light-chain monoclonal gammopathy (121).

Amyloidosis observed on bone marrow sections usually involves vessels; however, an interstitial infiltrative pattern also may be seen, as illustrated in Fig. 44.21. The presence of amyloidosis can be confirmed on marrow sections with thioflavin-T and Congo red stains. Because amyloid can be detected in the bone marrow in only one fourth of patients with primary systemic amyloidosis; the abdominal fat aspirate is the screening method of choice in confirming the diagnosis rather than bone-marrow biopsy.

In keeping with the concept that primary amyloidosis is an immunoproliferative disorder, it has been demonstrated that, as in benign monoclonal gammopathy, there is an excess of plasma cells of restricted light chain type in almost three fourths of the bone marrows in primary amyloidosis (128). A mild atypical plasmacytosis, as seen in benign monoclonal gammopathies or smoldering myeloma, also is seen in primary amyloidosis, sometimes with plasma cells embedded in interstitial amyloid deposits. When a patient with laboratory and clinical features of primary amyloidosis has more than 10% atypical plasma cells but no other features of plasma-cell myeloma, it becomes problematic as to whether the patient also should be diagnosed as having plasma-cell myeloma.

MALIGNANT LYMPHOMA

Part of "44 - Chronic Lymphoproliferative Disorders, Immunoproliferative Disorders, and Malignant Lymphoma"

Malignant lymphomas include a diverse group of disorders ranging from morphologically mature processes with an indolent course to highly aggressive processes with morphologic immaturity. Although these processes are diagnosed primarily by tissue biopsy, they frequently have manifestations in the blood and bone marrow. Blood and bone marrow manifestations are emphasized in this chapter; more specific details of morphologic diagnosis in extramarrow sites can be found in textbooks and articles devoted to lymph-node morphology (167, 168 and 169).

Malignant lymphomas are of two major types: Hodgkin's disease and non-Hodgkin's lymphoma. There is a major difference between these two types of lymphomas. In Hodgkin's disease, the infiltrate has a minority population of a large, pleomorphic neoplastic cell (the Reed-Sternberg cell) and a major population of reactive cells including predominantly lymphocytes, with variable numbers of admixed histiocytes, eosinophils, and plasma cells. In non-Hodgkin's lymphomas, the infiltrate is composed predominantly of the neoplastic lymphoid cell.

There have been multiple classification schemes for subclassification of non-Hodgkin's lymphomas over the past several decades. In the past, non-Hodgkin's lymphomas were subclassified on the basis of infiltrative pattern and on the basis of the cell size and nuclear characteristic. Pattern and cytology are still important features in modern-day classification; however, now with the availability of monoclonal antibodies for use on paraffin section material, it is possible to subclassify the non-Hodgkin's lymphomas more precisely based on their functional cell of origin. Functional cell of origin is the major basis of the revised European-American classification of lymphoid neoplasms (REAL), which will be used in this chapter (8).

The classification of Hodgkin's disease, on the other hand, has changed only to a minor degree in the last three decades (Table 44.18). Classical Hodgkin's disease is subclassified on the basis of whether or not a nodular sclerosing pattern is present and on the basis of cellular composition, i.e., lymphocyte predominance or depletion, and number of Reed-Sternberg cells (1, 168). More recently, lymphocyte-predominant Hodgkin's lymphoma has been divided into the classical variant, which is immunophenotypically similar to other classical subtypes, and nodular lymphocyte-predominant Hodgkin's lymphoma in which the Reed-Sternberg cell has an easily identifiable B-cell phenotype, and unlike the Reed-Sternberg cell of classical Hodgkin's disease, is CD30 and CD15 negative (8, 168).

Malignant lymphomas have variable involvement of the bone marrow and peripheral blood. The relative frequency of lymphomas and their incidence in the bone marrow are shown in Table 44.18 and Table 44.19 (169, 170, 171 and 172).

Non-Hodgkin's Lymphoma

Diagnostic Features

Non-Hodgkin's lymphomas usually are initially diagnosed by examination of a tissue biopsy of a lymph node or other extra-marrow site rather than bone marrow sections. Tissue diagnosis is based on the effacement of architecture by a neoplastic lymphoid infiltrate distributed in a follicular, mantle zone, marginal zone, paracortical T-cell, or diffuse pattern. Once a diagnosis of a neoplastic lymphoid infiltrate and its pattern are identified, the cytologic subtype is assessed. Immunologic marker studies are frequently used if the specific subtype is still in question.

The details of morphologic features of the more commonly encountered types of non-Hodgkin's lymphomas are outlined in Fig. 44.22, Fig. 44.23, Fig. 44.24, Fig. 44.25 and Fig. 44.26. Small lymphocytic and lymphoplasmacytoid lymphoma are morphologically indistinguishable from CLL and Waldenström's macroglobulinemia, respectively described in Fig. 44.1, Fig. 44.2, Fig. 44.5, and Fig. 44.20.

Many lymphomas are diagnosed initially on tissue biopsy, and the marrow, if involved, is diagnosed secondarily. However, there are situations in which a primary diagnosis can be made by examination of bone-marrow sections. General principles that apply to the primary as well as secondary diagnosis of non-Hodgkin's lymphoma in the bone marrow follow.

1. Bone marrow involvement can take a variety of patterns. In the early stages of involvement, the bone marrow usually is focally involved by aggregates of lymphoma cells. These aggregates may be either nonparatrabecular or paratrabecular (171). A lymphoma with a paratrabecular pattern is illustrated in Fig. 44.27. Multiple paratrabecular aggregates may be diagnosed as lymphoma solely on the basis of their pattern, whereas nonparatrabecular aggregates, which mimic benign lymphoid aggregates,

TABLE 44.19. NON-HODGKIN'S MALIGNANT LYMPHOMAS: CLASSIFICATION AND INCIDENCE OF BONE MARROW INVOLVEMENT

REAL Classification Diagnoses	Approximate Relative Frequency	Incidence of Bone Marrow Involvement
B-cell Lymphomas		
Precursor B-lymphoblastic lymphoma	*	*
Peripheral B-cell lymphomas		
Small lymphocytic lymphoma	5%	70%
Lymphoplasmacytoid lymphoma	*	75%
Mantle-cell lymphoma	10%	65%
Follicle-center lymphoma		
Grade I (small cell)	15%	50%
Grade II (mixed small and large cell)	15%	30%
Grade III (large cell)	5%	15%
Diffuse (predominantly small cell)	*	*
Marginal zone B-cell lymphoma	10%	*
Diffuse large B-cell lymphoma	30%	10%
Burkitt's lymphoma	5%	15%
High-grade Burkitt like lymphoma	*	*
T-cell and NK-cell lymphomas		
Precursor T-lymphoblastic lymphoma	*	*
Peripheral T-cell and NK-cell lymphomas	5%	*

* Data is not readily available because of variability in the distinction between lymphoma vs. leukemia, because the diagnosis is poorly defined, or because it contains multiple subtypes.

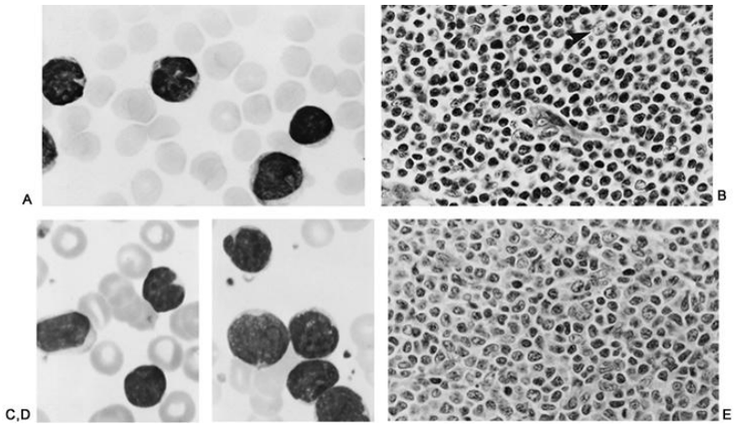


FIGURE 44.22. Lymphomas with a majority of cells the size of lymphocytes. **A, B:** Mantle-cell lymphoma. Many nearly normal appearing lymphocytes are present; however, some cells are slightly larger and there is moderate nuclear contour irregularity compared to small lymphocytic lymphoma. **C, D, E:** Small cleaved cells from a follicle-center lymphoma. On smears and sections, note the size range from small to large abnormal lymphocytes with scant cytoplasm and many sharply clefted nuclei. On sections in E, small cells predominate but scattered large cells the size of large-cell lymphoma are also seen. A normal histiocyte nucleus is present in the upper right center field of **B** (arrow) for size comparison. Wright's stain 1200X (**A, C, D**). Hematoxylin and eosin, 600X (**B, E**). (**A, C**, and **D**, From Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's lymphoma. In: Koepke JA, ed. Laboratory Hematology. New York: Churchill Livingstone, 1984;1:325–357.)

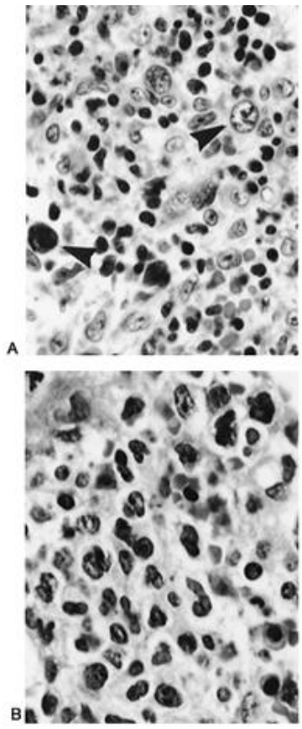


FIGURE 44.23. Large-cell lymphomas. Atypical large cells are numerous; however a mixture of cells the size of lymphocytes are also present. Note in (**A**) the hyperchromatic and convoluted large cells (arrows) and in (**B**) the numerous large clefted cells. The lymphoma in **A** immunophenotyped as a T-cell lymphoma and the lymphoma in **B** as a B-cell lymphoma. Hematoxylin and eosin, 600X.

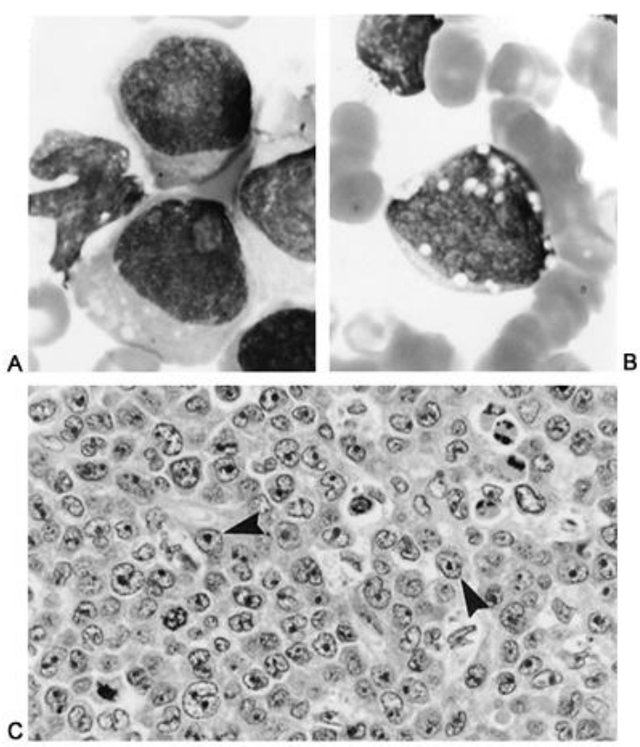


FIGURE 44.24. Large-cell lymphomas. These lymphomas have a majority of nuclei the size of a histiocyte nucleus or larger. On smears (**A, B**), there is marked variability in appearance of cells from one lymphoma to the next, (**A**) marked as a T-cell lymphoma and (**B**) as a B-cell lymphoma. **C:** This large-cell lymphoma shows many large non-clefted cells (arrows) with dispersed chromatin and nucleoli, as well as cells with nuclear pleomorphism. It immunophenotyped as a B-cell lymphoma. Wright's stain 1200X (**A, B**). Hematoxylin and eosin, 600X (**C**).

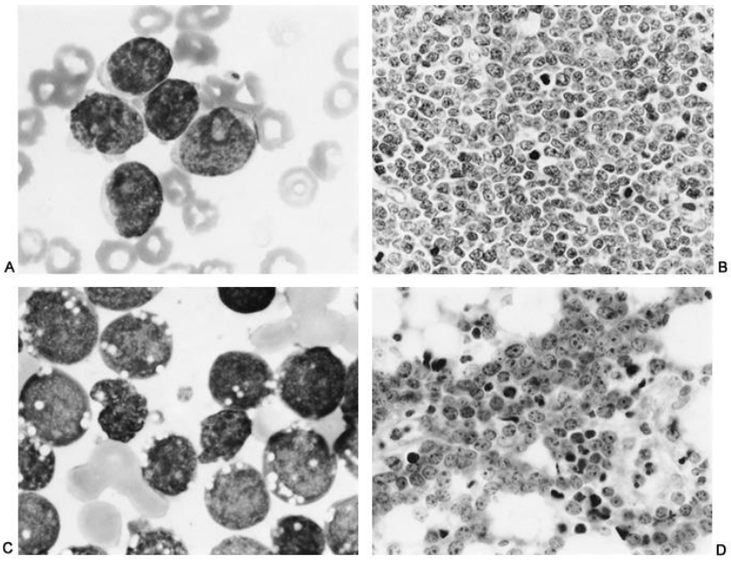


FIGURE 44.25. High-grade lymphomas with a blastlike appearance and very high mitotic activity. **A, B:** Lymphoblastic lymphoma. Lymphoma-cell nuclei are smaller than a histiocyte nucleus on sections, and on smears they are morphologically similar to blasts of acute lymphocytic leukemia of L1 and L2 type. Note the relative uniformity of cellular composition, stippled chromatin pattern, and mitotic activity on sections (**B**). Cells on smears and sections in lymphoblastic lymphomas (as in acute lymphocytic leukemia) may show a moderate degree of nuclear contour irregularity (convolution). **C, D:** Burkitt's lymphoma. Note the marked uniformity of cellular composition, hyperbasophilic cytoplasm, and crisp cytoplasmic and nuclear vacuoles. Nuclear size is slightly smaller than or about the size of a histiocyte nucleus on sections, and on smears the cells are morphologically similar to acute lymphocytic leukemia, L3 type. Note that some large-cell lymphoma cells (illustrated in Fig. 6A and Fig. 26B) may have, except for their size, some of the characteristics of Burkitt's lymphoma cells. Wright's stain 1200X (**A, C**). Hematoxylin and eosin, 600X (**B, D**).

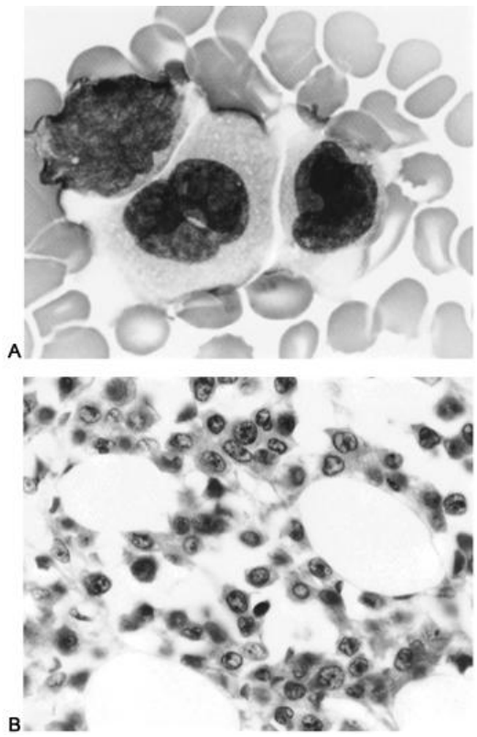


FIGURE 44.26. True histiocytic lymphoma/malignant histiocytosis, peripheral blood (A) and bone marrow section (B). Although this tumor resembles a large-cell lymphoma on sections, it is a malignancy of histiocytes. Note the resemblance of the cells on smears to large, highly pleomorphic monoblasts. The histiomonocytic nature of this process can be confirmed with a nonspecific esterase stain on smears or immunologic markers on sections or cell suspension. Wright's stain 1200X (A). Hematoxylin and eosin, 600X (B).

must be diagnosed using other features such as cytologic atypia or marker studies. As involvement increases, the marrow becomes more diffusely infiltrated and the focal pattern is obscured. Malignant lymphomas also may rarely have minimal involvement in an interstitial pattern in which lymphoma cells are diffusely admixed with normal marrow elements and don't form aggregates. When this occurs, it may be difficult to make a diagnosis from the examination of marrow sections, and the diagnosis may rest more heavily on identifying neoplastic lymphoid cells in aspirates or touch preparations, or on the use of immunologic markers.

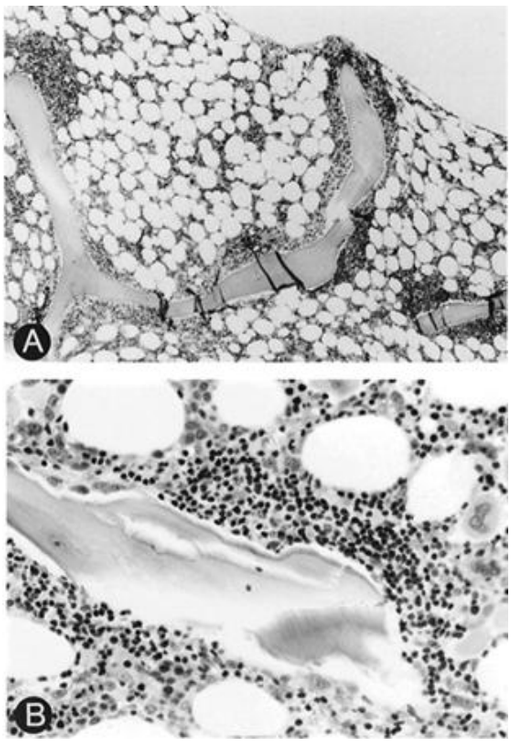


FIGURE 44.27. Paratrabecular pattern in a follicle center lymphoma. Note the “plastered on” or “piled up” appearance to most of the aggregates. The aggregate in the upper left field of (**A**) is not truly paratrabecular because its center appears to be away from the bone, and it could be touching bone by virtue of growth towards it. Note in (**B**) how the paratrabecular aggregate appears to be growing on the bone. Also note in (**B**) that the lymphocytes in this lymphoma are small and show only mild nuclear clefting, even though the lymph node biopsy showed considerable nuclear clefting. Hematoxylin and eosin, 50X (**A**), 240X (**B**).

2. Identification of the pattern of non-Hodgkin's lymphoma is rarely possible on bone-marrow examination. Focal aggregates of lymphoma can be seen in lymphomas that have a diffuse or follicular pattern on lymph-node biopsy. A focal pattern in the marrow is not equivalent to follicular lymphoma.
3. When assessing the cytology in non-Hodgkin's lymphoma, it is useful to compare the size of the nucleus of the lymphoma cells with the nucleus of a normal histiocyte on sections (or the nucleus of a myeloblast on smears). Small lymphocytic, lymphoplasmacytoid, mantle-cell, small-follicle-center cells, lymphoblastic, and small-cell variants of peripheral T-cell lymphoma have a nucleus that is definitely smaller than that of a histiocyte. Burkitt's lymphoma has a nucleus slightly smaller than or almost the size of a histiocyte nucleus, and the nuclei in large-cell lymphomas are the size of and larger than a histiocyte nucleus.
4. The bone marrow may show a lower grade subtype of lymphoma than an initial tissue biopsy (171, 172, 173, 174 and 175). For example, when a lymphoma of large-cell type is initially diagnosed in an extra-marrow site, the marrow may show a predominantly small cell type. This is particularly true of the follicle center and diffuse large-B-cell lymphomas. Thus, non-Hodgkin's lymphoma should not be definitively subtyped for purposes of therapy when a small-cell lymphoma is seen initially in the marrow, especially if an additional extra marrow site is available for biopsy. If the marrow is the only site available, then the clinician should be made aware of the potential problem of discordance. On the other hand, a primary diagnosis of large B-cell, or Burkitt's lymphoma in the marrow can be used for classification for therapy without biopsying an extra marrow site even if disease is present because discordance is not a problem with these subtypes. Although a primary diagnosis of a peripheral T-cell lymphoma can be made on bone marrow alone further subclassification T-cell lymphoma should be made with extreme caution. T-cell-rich B-cell variant of large B-cell lymphoma and Hodgkin's disease are in the differential diagnosis of T-cell lymphoma in the marrow (176, 177).
5. When faced with a lymphoma composed of small lymphocytes, and the marrow is the only site available, immunologic marker studies will be helpful in distinguishing among the small-cell types of B-cell lymphoma: small lymphocytic, lymphoplasmacytoid, mantle-cell, marginal-zone, and follicle-cell lymphomas (1, 177). These subtypes will be discussed later. Search for a monoclonal serum immunoglobulin should be performed if lymphoplasmacytoid lymphoma is a consideration.
6. Malignant lymphomas primarily represent tissue infiltration of neoplastic lymphoid cells, whereas leukemias are bone marrow and peripheral blood manifestations of neoplastic lymphoid cells. As shown in Table 44.20, for some specific cell types, the “lymphomatous” (tissue phase) and its “leukemic” or “marrow” counterpart have separate diagnostic terminology. For example, a neoplastic proliferation of the mature CD5-positive B cell with weak SIg is diagnosed as CLL when there is a peripheral blood lymphocytosis. When it involves tissues with minimal blood and marrow involvement, it is called SLL. Neoplastic proliferations of cells that comprise other types of lymphoma such as mantle-cell lymphoma; the follicle-center lymphomas; and large-cell lymphoma have no specific terminology for their leukemic phase, and thus these are simply diagnosed as the leukemic phase of malignant lymphoma. The archaic term “lymphosarcoma-cell leukemia” is no longer used.
7. In general, a trephine biopsy is superior to aspirate and particle sections in diagnosing malignant lymphoma in the marrow. This is because early bone marrow involvement is composed of small aggregates that may have increased reticulin and do not aspirate well. In addition, extensive marrow involvement may be partly fibrotic and result in a dry tap. In this latter situation, touch preparations are helpful in seeing cytologic detail. Because bone marrow involvement is frequently focal and may be minimal, bilateral bone marrow trephine biopsies of the posterior crests with five-stepped sections on each side generally are recommended (178, 179 and 180). The volume of bone marrow examined is also important because there may be focal, minimal involvement; thus, it is recommended that between the right and left biopsies a minimum total of 3 cm of trephine biopsy specimen in aggregate length be acquired, and more is preferable (181). This may require more than one pass of the trephine needle on each side. Lymphoma may rarely be seen on particle clot sections and not on trephine specimens; thus, these also should be carefully examined (178).
8. Of critical importance is the distinction between focal aggregates of lymphoma and benign lymphoid aggregates (Fig. 44.28A and Fig. 44.28B) (182). Important differential diagnostic features are listed in Table 44.21. Immunohistochemistry also may be of help as discussed in a later section.
9. Mixed cellular aggregates that are variably composed of lymphocytes, plasma cells, histiocytes, fibroblasts, endothelial cells, and immunoblasts may be seen in marrow involvement by lymphoma, especially peripheral T-cell lymphomas, Hodgkin's disease, and plasmacytoid lymphomas (Fig. 44.28C and Fig. 44.28D) (183, 184).

TABLE 44.20. TERMINOLOGY FOR "LYMPHOMATOUS" VS. "LEUKEMIC OR MARROW" MANIFESTATIONS OF LYMPHOID NEOPLASMS

"Lymphomatous" Counterpart	"Leukemic or Marrow" Counterpart
Small lymphocytic lymphoma	Chronic lymphocytic leukemia
Lymphoplasmacytoid lymphoma	Waldenström's macroglobulinemia
Burkitt's lymphoma	L3 acute lymphoblastic leukemia
Lymphoblastic lymphoma	L1 and L2 acute lymphoblastic leukemia
Mycosis fungoides	Sezary syndrome
Adult T-cell leukemia/lymphoma	Adult T-cell leukemia/lymphoma
Plasmacytoma	Plasma-cell myeloma

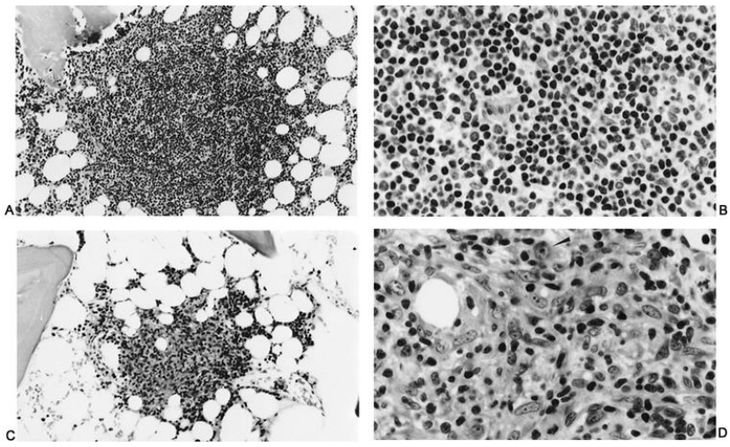


FIGURE 44.28. A, B: Large benign lymphoid aggregate. Note the mild to moderate degree of nuclear contour irregularity. **C, D:** Mixed cellular aggregate. Note the depletion of lymphocytes with admixed, histiocytes, fibroblasts, small vessels, and occasional large cells (immunoblasts) (arrow). This aggregate was seen in the marrow of a patient with a peripheral T-cell lymphoma. This aggregate is not diagnostic of marrow involvement because similar aggregates can be seen in a variety of other processes. Atypical large and small lymphoid cells similar to those seen in Fig. 44.23 or more extensive marrow infiltration would be necessary to make a definitive diagnosis of non-Hodgkin's lymphoma. Hematoxylin and eosin, 120X (**A, C**), 600X (**B, D**).

TABLE 44.21. DIFFERENTIAL DIAGNOSTIC FEATURES BETWEEN BENIGN LYMPHOID AGGREGATES AND FOCAL AGGREGATES OF LYMPHOMA

Feature of Aggregates	Benign	Lymphoma	Comment
Low Power Features			
Paratrabecular	Rare	Yes	
Nonparatrabecular	Usually	Yes	
Uniform rounded contour	Yes	Yes	
Irregular contour	Unusual	Yes	
Benign follicle center	Uncommon	Rare	Neoplastic follicles also are uncommon in bone marrow.
Poorly circumscribed spreading border	No	Yes	Especially in small lymphocytic lymphoma.
Circumscribed border	Yes	Yes	
Associated lipoid granulomas	Yes	Uncommon	
Large size	Uncommon	Yes	
Numerous	Yes	Yes	
Rare	Yes	Yes	
Small size	Yes	Yes	
Blood vessels	Yes	Uncommon	
Cytologic Features			
Minimal to mild nuclear contour irregularity	Yes	Yes	Seen in small lymphocytic, mantle-cell and follicle-center lymphoma.
Distinctly atypical cell type	No	Yes	Except for small lymphocytic, mantle-cell, and follicle-center lymphoma, which may show only mild nuclear contour irregularity.
Admixture of other elements (plasma cells, histiocytes, immunoblasts)	Yes	Sometimes	Lymphoplasmacytoid lymphoma, Hodgkin's disease, and other non-Hodgkin's lymphomas of peripheral T-cells may show a mixed cell population.

Because these mixed cellular aggregates also can be seen in infections and in autoimmune deficiency syndrome (AIDS) or angio-immunoblastic lymphadenopathy with dysproteinemia (AILD), which may be associated with or transform to lymphoma, the presence of mixed cellular aggregates in the marrow may result in very difficult differential diagnostic problems. Small mixed cellular aggregates should not be diagnosed as lymphoma unless they contain Reed-Sternberg cells morphologically and immunophenotypically consistent with a diagnosis of Hodgkin's disease, or contain atypical neoplastic lymphoid cells consistent with one of the subtypes of non-Hodgkin's lymphoma.

10. Criteria for marrow involvement vary depending on the lymphoma in question, and on whether the primary diagnosis has been made elsewhere. In general, less strict evidence is needed to secondarily diagnose marrow involvement than to make a primary diagnosis of lymphoma in the bone marrow. However, the presence of aggregates without neoplastic features should not be considered definitive evidence for secondary marrow involvement, even if the aggregates are numerous. Immunophenotyping may be of help in this situation if the aggregates contain a marked predominance of B-cells and the initial diagnosis was a B-cell lymphoma. This will be discussed in more detail later.
11. Lower-grade non-Hodgkin's lymphoma may undergo transformation to a higher-grade process, especially large-cell lymphoma. In this situation, the bone marrow may be the first site biopsied, resulting in a higher-grade marrow morphology than that seen in the original biopsy specimen (Table 44.22)(172, 173).

TABLE 44.22. MORPHOLOGIC VARIABILITY IN NHL AND HODGKIN'S DISEASE IN THE BONE MARROW, WHICH MAY LEAD TO DIAGNOSTIC DIFFICULTY

Non-Hodgkin's Lymphoma
Morphologic discordance with tissue biopsy
Interstitial infiltrative pattern
Leukemic presentation
Reed-Sternberg cells
Fibrosis
Epithelioid histiocytes
Necrosis
Posttherapy changes
Hodgkin's Disease
Fibrosis
Lack of Reed-Sternberg cells
Epithelioid histiocytes
Sheets of Reed-Sternberg cells
Posttherapy changes

NHL, non-Hodgkin's lymphoma

12. Reed-Sternberg–like cells can be seen in the marrow as well as in the lymph node in non-Hodgkin's lymphoma, especially in peripheral T-cell lymphoma, or in other lymphomas late in the course of the disease or with transformation (Fig. 44.29) (185, 186). The best way to distinguish non-Hodgkin's lymphoma from Hodgkin's disease is to identify an atypical lymphoid population in the infiltrate characteristic of a subtype of non-Hodgkin's lymphoma, in addition to the Reed-Sternberg–like cell. It is also helpful (but not always necessary) to have a tissue biopsy diagnosis of non-Hodgkin's lymphoma. Immunostaining also may be helpful if the Reed-Sternberg–like cells do not show the typical phenotype of Hodgkin's disease (187).

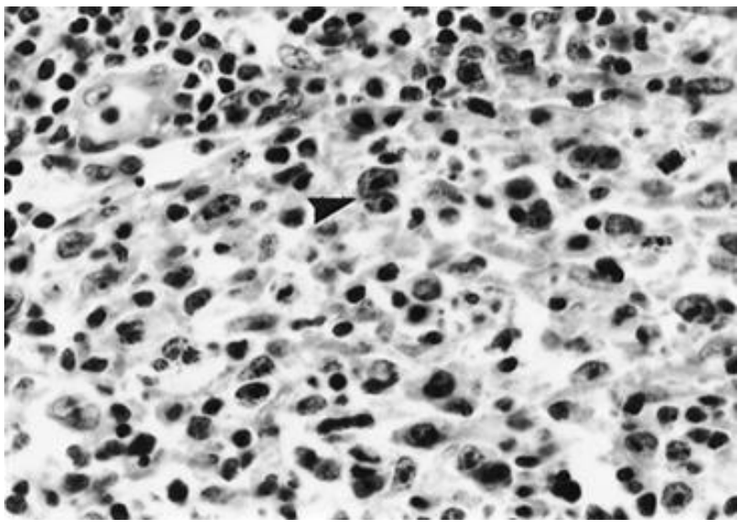


FIGURE 44.29. Follicle-center lymphoma with a Reed-Sternberg–like cell *(arrow)*. Hematoxylin and eosin, 600X.

Additional features of specific subtypes of non-Hodgkin's lymphoma and their differential diagnosis are described in the paragraphs that follow. Their characteristic marker studies are shown in Table 44.23.

Small Lymphocytic Lymphoma

SLL as diagnosed on tissue biopsy is essentially the tissue phase or the nonleukemic counterpart of CLL. Thus, to diagnose a patient as having SLL, the peripheral blood white count must be less than 5,000 lymphocytes/mm^3 (3). Otherwise, the lymph node or other tissue is diagnosed as SLL consistent with CLL, and the patient is diagnosed as having CLL. The pattern of marrow involvement

in SLL is similar to that of CLL, as illustrated in Fig. 44.1 and Fig. 44.2; however, a focal pattern is seen more commonly in SLL than is interstitial or diffuse extensive involvement (171). As in CLL, individual-cell morphology is not helpful in the diagnosis of SLL, because the tumor is made up of predominantly mature lymphocytes. The morphologic differential diagnosis of SLL includes primarily lymphoplasmacytoid lymphoma and mantle-cell lymphoma. Positivity for CD5, CD23, and CD43 with negative cyclin D1 and no IgM, monoclonal gammopathy is characteristic of SLL (16, 188, 189, 190, 191 and 192) (Table 44.23). If an aspirate is saved for flow cytometry, a clonal population of cells with weakly reacting surface immunoglobulin also will be helpful.

Lymphoplasmacytoid Lymphoma

As emphasized earlier, lymphoplasmacytoid lymphoma is morphologically the same disorder as Waldenström's macroglobulinemia, which was illustrated earlier in Fig. 44.20. The presence of a serum IgM monoclonal gammopathy and demonstration of plasmacytoid cells with clonal IgM cytoplasmic immunoglobulin if a significant monoclonal peak is not identified, are helpful in confirming the diagnosis.

Mantle-Cell Lymphoma

This subtype of lymphoma was not included in earlier classification schemes. It is now recognized as a distinct clinical pathologic entity (8, 15, 189, 191, 192, 193 and 194). It is characterized by a proliferation of small lymphocytes with a mild-to-moderate degree of nuclear irregularity (Fig. 44.22A and Fig. 44.22B) that in a lymph node may have a diffuse or a mantle-zone pattern that is characterized by nodules of lymphoma cells, some of which contain a benign-appearing germinal center. This lymphoma commonly has marrow involvement. The distribution in early bone marrow involvement usually is focal and nonparatrabecular; however, there also may be some element of a paratrabecular pattern. With extensive involvement, the pattern becomes more diffuse. On blood smears, many of the cells are fairly normal-appearing lymphocytes; however, a moderate number (10% to 20%) of cells with nuclear contour irregularity are seen (Fig. 44.22A). When the WBC is elevated, the process may be confused with CLL (15). The differential diagnosis of mantle-cell lymphoma includes primarily SLL and follicle-center lymphoma with a predominance of small cells. As indicated in Table 44.23 mantle-cell lymphoma can be distinguished on paraffin sections from these entities by its negativity for CD23 and CD10, and positivity with cyclinD-1 (16).

Marginal-Zone Lymphoma

Some cases of non-Hodgkin's lymphoma, especially the marginal-zone variant called splenic lymphoma with villous lymphocytes, may have hairy cytoplasmic borders and TRAP positivity and primarily involve the spleen (60, 177, 195). Marrow involvement is common but usually is more focal than hairy-cell leukemia. Distinction from hairy-cell leukemia may be very difficult, even with marker studies, however marker studies may be helpful if CD11c, CD25, and TRAP are negative. Because the infiltrative pattern in the spleen is different in non-Hodgkin's lymphoma (white pulp) than in hairy-cell leukemia (red pulp), splenectomy also may be useful in the differential diagnosis if the spleen is removed. As opposed to the splenic variant, nodal and extranodal variants of marginal-zone lymphoma rarely have marrow involvement.

Follicle Center Lymphomas

As indicated in Table 44.19, cells of follicle-center lymphomas, as initially described by Lukes and Collins, are composed of small and large (cleaved) cells and large (noncleaved or transformed) cells (196). On bone marrow sections, the early infiltrate of follicle-center lymphomas is characteristically focal with paratrabecular and occasional nonparatrabecular aggregates (171, 197). With more extensive involvement, the pattern becomes diffuse. To make a diagnosis of marrow involvement by follicle-center lymphoma when a diagnosis of lymphoma has previously been made on a marrow biopsy, one should see several distinctly paratrabecular lymphoid aggregates or focal infiltrates with definitely atypical cleaved lymphoid cells. The cytology on sections of follicle-center lymphoma involving the marrow may range from lymphoid cells with only mild nuclear contour irregularities (similar to mantle-cell lymphoma), to markedly clefted cells, to a mixture of large cells and small clefted cells, as illustrated in Fig. 44.22 and Fig. 44.23B. When the cytology of the cells is fairly bland and aggregates are small, the diagnosis of marrow involvement may have to be made on the basis of paratrabecular distribution of aggregates rather than on atypicality of the lymphoid cells. In follicle-center lymphomas, the cytologic atypicality and number of large cells in the bone marrow are very frequently less than in the lymph-node biopsy (171, 172).

When the marrow is extensively replaced, small clefted cells similar to those shown in Fig. 44.22C and Fig. 44.22D are seen on the aspirate. These cells are morphologically similar in appearance to the hematogones seen in the bone marrow of pediatric patients (171, 198).

The peripheral blood, similar to mantle-cell lymphoma, may be involved in follicle-center lymphomas. When the WBC is considerably increased, the peripheral blood also contains numerous small clefted cells. If the small clefted cells are not readily identified, and a lymph node biopsy has not been done, the case may be mistakenly diagnosed as CLL (198). If a previous or concurrent lymph node has not been biopsied, negativity of a SLL in the marrow to CD5, CD23, CD43, and Cyclin D1 and a positive CD10 suggests a follicle-center lymphoma.

As stated earlier, a primary diagnosis of small-cell lymphomas (SLL, mantle-cell, lymphoplasmacytoid, follicle-center, marginal-zone) should not be made on the basis of bone-marrow biopsy alone because a tissue biopsy may show a higher-grade cell type a follicular versus diffuse pattern cannot usually be determined and nuclear contour irregularity may be less pronounced in the marrow.

With minimal marrow involvement by the small cells malignant lymphomas described, it may be difficult to recognize lymphoma on bone-marrow aspirates if diagnostic aggregates are not present. Flow cytometry looking for a small clonal population may be indicated to identify minimal involvement in these cases.

Large-Cell Lymphomas

This morphologic subtype includes a heterogeneous group of lymphomas. Many cases represent the diffuse counterpart of follicle-center

lymphomas; however, others represent lymphomas derived from nonfollicular B cells or peripheral T cells (8, 167, 196). Bone-marrow involvement by large-cell lymphomas is illustrated in Fig. 44.23 and Fig. 44.24. The pattern of involvement by large cell lymphoma ranges from large focal areas of marrow replacement to interstitial infiltration by single cells that are difficult to identify on section material and may need immunohistochemistry to identify (Fig. 44.30). The differential diagnosis of large-cell lymphoma initially diagnosed in the bone marrow includes metastatic undifferentiated carcinoma, and if the marrow is completely replaced, acute leukemia. Occasionally, SLL may have numerous prolymphocytes and blastlike cells on lymph-node biopsy similar to Fig. 44.5B, and mimic a diffuse mixed lymphoma (17, 18, 199). When the bone marrow is examined in these cases, the marrow may show an infiltrate similar to SLL or CLL. This finding in the marrow should help the pathologist recognize the lymph node as a SLL with transformation as opposed to a more aggressive diffuse mixed lymphoma (199). Reed-Sternberg cells also may be seen in large-cell lymphoma in the bone marrow, and should be distinguished from Hodgkin's disease.

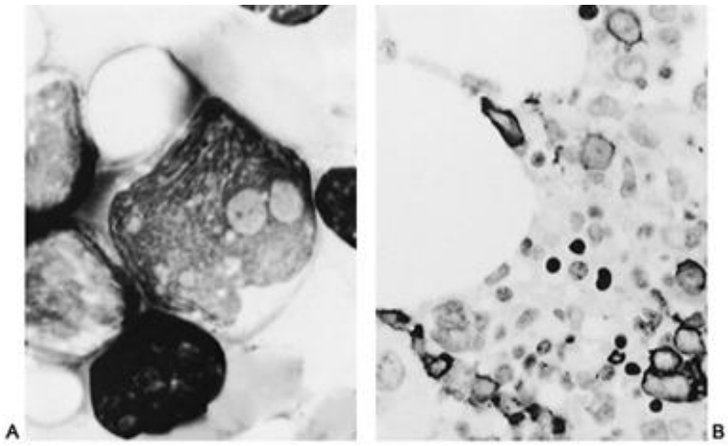


FIGURE 44.30. Large-cell lymphoma in bone marrow, diagnosed with the assistance of CD20 monoclonal antibody on paraffin section. **A:** On marrow smears in this case only, rare cells suspicious for large-cell lymphoma were seen. Sections (not shown) showed small clusters of large cells that could not be distinguished with certainty from reactive marrow elements. **B:** Immunoperoxidase staining with CD20 demonstrated small clusters and isolated positive cells (note dark staining of cytoplasmic outline), which were interpreted as B-cell, large-cell lymphoma. Wright's stain 1200X (**A**). Immunoperoxidase 600X (**B**).

Lymphoblastic Lymphoma and Burkitt's Lymphoma

As discussed previously, lymphoblastic lymphoma and acute lymphocytic leukemia (ALL) of types L1 and L2 are morphologically similar, and Burkitt's lymphoma and ALL L3 are morphologically similar. When a patient has a tissue mass diagnosed as lymphoblastic lymphoma or Burkitt's lymphoma, and the marrow shows only a small percentage of lymphoblasts with considerable marrow sparing, it is reasonable to diagnose the patient as having lymphoma. However, if the marrow is more extensively involved, then it becomes debatable whether the patient's condition should be called a malignant lymphoma with extensive marrow involvement or a leukemia with extramarrow mass(es) (see Table 44.20). In this situation, it may be advisable to diagnose the patient as having “leukemia/lymphoma” rather than struggle with deciding which came first. On the other hand, a patient presenting with a mass and blasts in the peripheral blood usually will have a bone-marrow biopsy done before a biopsy of the mass is done. At that point, a decision must be made whether to biopsy the mass, which will probably show morphology similar to the marrow.

Specific criteria for distinguishing leukemia from lymphoma for these cell types have been proposed; however, the criteria vary depending on the study group involved and thus before diagnosing an equivocal case as leukemia or lymphoma it would be prudent to consult with the hematologist/oncologist caring for the patient. Therapy for lymphoblastic lymphoma is frequently similar to aggressive therapy for ALL; thus, a firm commitment to a distinction between leukemia and lymphoma in this situation may not be necessary. Marker analysis of cells to distinguish leukemia from lymphoma is of little help. Marker analysis for ALL L3 and small, noncleaved-cell lymphoma is very similar. Lymphoblastic lymphoma is more often a precursor T-cell type (especially late thymocyte stage) than a precursor B-cell type; however, ALL also can be of precursor T-cell type. The morphologic features of the cells in ALL L1 and L2 and ALL L3 subtypes also are described in Chapter 43.

Morphologic Variability and Differential Diagnosis

As discussed in the preceding section, a characteristic of non-Hodgkin's lymphoma in the bone marrow is morphologic variability, in terms of both distribution and cytologic appearance. Some of the variability that may lead to confusion and difficulties in differential diagnosis are listed in Table 44.22. The first four of these already have been discussed in the previous section.

Fibrosis is an uncommon occurrence in the bone marrow in non-Hodgkin's lymphoma. When it does occur, however, diagnosis may be difficult unless abnormal lymphoid cells are identified in the fibrosis.

Epithelioid histiocytes, epithelioid-cell clusters, and granulomas may be admixed with non-Hodgkin's lymphoma in the bone marrow, but they also may be seen as an epiphenomenon without marrow involvement by lymphoma (200). Again, the most critical observation is to find abnormal lymphoma cells admixed with the histiocytes. Unless the histiocytes are a characteristic of the lymphoma at the previous site and lymphoma cells are admixed with histiocytes in the marrow, special stains for organisms should be performed.

Necrosis is uncommon but is very bothersome if present. This is because marrow necrosis in a patient with lymphoma is most likely from marrow involvement with lymphoma. However, a definite diagnosis cannot be made without viable tissue evident. Taking a biopsy at site may be helpful if it is crucial to know if the marrow is involved.

Following therapy for malignant lymphoma, the marrow may become altered by fibrosis or necrosis, making determination of residual disease difficult. In addition, aggregates in low-grade lymphoma may be reduced in size, so they are difficult to distinguish from benign lymphoid aggregates. When this occurs, it is best to not diagnose the lesion as lymphoma unless definite evidence of lymphoma is present. Instead, a comment should be made that lymphoma cannot be excluded.

Immunologic and Other Special Studies

A detailed description of the immunologic findings in non-Hodgkin's lymphoma is beyond the scope of this chapter. An overview of the evaluation of REAL classification of non-Hodgkin's lymphoma by marker studies can be found in the literature (8). A brief outline of the immunologic findings in non-Hodgkin's lymphoma is shown in Table 44.23. The more commonly used markers are represented. Except for lymphoblastic lymphomas, all of the non-Hodgkin's lymphomas show phenotypes similar to functionally mature/peripheral B or T-cells. Lymphoplasmacytoid lymphoma also may show cytoplasmic immunoglobulin.

Usually, marker analysis is not necessary to evaluate malignant lymphoma in the bone marrow because the diagnosis has been made on another site that may already have been immunophenotyped. Although most B-cell and T-cell lineage-specific markers can be done on paraffin section material, surface immunoglobulin to prove clonality and some lineage-specific markers such as CD11c can only reproducibly be done by flow cytometry; thus, it is suggested that a fresh aspirate for flow cytometry be saved on any patient with a suspected lymphoproliferative disorder in the bone marrow until the need for flow cytometry is determined. A fresh aspirate may be held as long as 24 hours, i.e., until after the section material has been processed and examined.

Specific situations in which immunophenotyping is useful in the evaluation of non-Hodgkin's lymphoma in the bone marrow or blood are described below.

1. The differential diagnosis of a lymphoma seen initially in the bone marrow, rests between two morphologically similar processes that have different immunophenotypes; for example, a large-cell lymphoma that is SIg negative or T-cell type can be differentiated from Burkitt's lymphoma or SLL can be distinguished from a mantle-cell lymphoma.
2. The morphology in the marrow has changed or is different from that of the tissue biopsy, and the question is whether a new tumor has developed or a transformation from the original tumor has occurred.
3. The differential diagnosis is a reactive lymphocytosis versus involvement with lymphoma. In this situation, identification of a clonally restricted light chain in B-cell process is particularly useful in confirming a lymphoid process as neoplastic. Some neoplastic T-cell processes may be identified if an abnormal T-cell phenotype characteristic of the disease in question is identified. However, with the exception of clonally restricted B-cells immunophenotyping may not help, and molecular techniques may be necessary to prove clonality.
4. Benign aggregates need to be distinguished from lymphoma. Immunohistochemistry may have some usefulness in distinguishing benign from neoplastic aggregates of low-grade lymphoma on paraffin sections if a marked predominance of B-cells possessing the immunologic characteristics of one of the low-grade lymphomas is identified (201, 202 and 203). Also lymphoplasmacytoid lymphoma may be identified by the presence of clonally restructed cytoplasmic immunoglobulin (151).
5. Identification of minimal disease. Current therapy is based on morphologic interpretation of involvement in the blood and bone marrow. New techniques, such as immunophenotypic or genotypic and clonal PCR analysis, have demonstrated that lymphoma cells may be present even when they are not seen morphologically (204, 205 and 206), especially in low-grade lymphomas. The importance of recognizing a minimal amount of marrow involvement is not yet clear. However, if autologous transplantation is a consideration, identification of residual disease in transplanted marrow or peripheral blood stem cells may assume greater importance.

A variety of other special studies have been performed on non-Hodgkin's malignant lymphomas, including chromosomal analysis (207), DNA content (27, 28), nuclear proliferation antigen Ki-67 (208, 209), and antigen receptor gene rearrangement (204, 205, 210, 211 and 212). Some of these may have applicability to lymphoma in the bone marrow and blood (185, 186).

Changes with Progression and Transformation

SLL and follicle-center lymphoma occasionally undergo morphologic transformation to lymphoma with the morphologic appearance of a more aggressive, large-cell lymphoma (32) (Table 44.24). This transformation may involve the bone marrow as well as other tissue sites; therefore, when a transformation is suspected, a bone-marrow examination may be performed before a tissue biopsy. Mantle-cell lymphoma also can undergo a blastic transformation that has a morphologic appearance of ALL in the bone marrow or blood with very blastlike cells (189). This can be distinguished from ALL with marker studies, because the blasts in transformation still retain the phenotype of the mature B-cell lymphoma. SLL may meet the criteria of CLL if the peripheral blood lymphocyte count exceeds 5,000/mm^3. Likewise, lymphoblastic lymphoma and small, noncleaved-cell lymphoma may develop bone marrow and blood involvement late in the course of the disease and become morphologically indistinguishable from acute leukemia L1, L2, or L3 subtypes. Although the marrow may become extensively involved late in the course of large-cell lymphoma, a leukemic blood picture with elevated white blood count is uncommon. Development of therapy-related myeloid leukemia similar to that seen in plasma-cell myeloma has been reported in non-Hodgkin's lymphoma(32).

TABLE 44.24. TRANSFORMATION AND SECONDARY TUMORS IN LYMPHOMAS

Non-Hodgkin's Lymphoma
Small lymphocytic lymphoma to large-cell lymphoma
Follicle-center lymphoma to large-cell lymphoma
Mantle-cell lymphoma to “blastic” morphology
Lymphoma to leukemia (see Table 44.20)
Acute myeloid leukemia
Myelodysplastic syndrome
Hodgkin's Disease
Progression towards lymphocyte depleted
Non-Hodgkin's lymphoma
Acute lymphoid leukemia
Acute myeloid leukemia
Myelodysplastic syndrome

Clinical Features

Prognosis in non-Hodgkin's lymphomas is highly dependent on the subtype of lymphoma and the extent of disease (stage). Depending

on the lymphoma type, the tumor can be seen in younger or older individuals. Low-grade lymphomas, which include peripheral B-cell lymphomas from small lymphocytic down through marginal zone on the classification list in Table 44.19, are seen almost exclusively in older individuals, whereas lymphoblastic lymphomas are seen predominantly in younger persons.

Therapy for non-Hodgkin's lymphoma is very dependent on the stage and grade of the lymphoma. Bone-marrow involvement automatically defines the patient as having widespread or stage IV disease. A few generalizations about therapy of non-Hodgkin's lymphomas follow.

Lymphomas composed of small lymphocytes usually are lumped together as low grade. Exceptions are that grade 2 follicle-center lymphoma are also considered low grade, and mantle-cell lymphoma acts more aggressively than the other low-grade lymphomas. Low-grade lymphomas are generally widespread and not curable; thus, therapy is directed at palliation as for CLL. The remaining lymphomas are frequently lumped together by clinicians as “high-grade” lymphoma. These lymphomas act aggressively if not treated but even if disseminated, long-term, disease-free survival may be achieved with intensive combination chemotherapy. Localized or bulky disease also may be treated with irradiation. Burkitt's lymphoma has the poorest prognosis of all tumors on the list. At present, autologous bone-marrow transplantation is a viable option for younger patients with a higher-grade lymphoma, widespread disease, and no marrow involvement. Autologous transplantation with purging or autologous peripheral stem-cell harvesting may have wider application in the future.

Hodgkin's Disease

Diagnostic Features

Hodgkin's disease usually is diagnosed on the basis of lymph node or other extramarrow tissue biopsy (8, 167, 168, 213). On tissue biopsy, there is effacement of architecture by an infiltrate that has as its basic component numerous mature lymphocytes and less frequent Reed-Sternberg cells. In addition, there may be varying numbers of benign histiocytes, eosinophils, and plasma cells. The subclassification of Hodgkin's disease is shown in Table 44.18.

When the bone marrow is involved by Hodgkin's disease, it is diagnosed by trephine biopsy and is almost never identified on clot sections, on smears, or in the peripheral blood (214, 215, 216, 217, 218 and 219). The marrow sections usually show focal infiltrates that are paratrabecular or nonparatrabecular. The infiltrates may be as small as lymphoid aggregates or may focally obliterate the marrow space, as in Fig. 44.31. The cellular infiltrates usually have a partly cellular and partly fibrotic appearance and are somewhat lymphocyte depleted, containing varying numbers of histiocytes and eosinophils and varying degrees of fibrosis. Reed-Sternberg cells usually are difficult to find, and numerous step sections may be necessary to identify them. With more extensive involvement, the marrow space can become completely obliterated by the infiltrate. The lymphocyte-depleted appearance of the bone marrow in Hodgkin's disease usually is present whether the lymph node shows a nodular sclerosing, mixed, or lymphocyte-depleted morphology. Thus, a subtype diagnosis of Hodgkin's disease should not be made on the basis of the bone marrow.

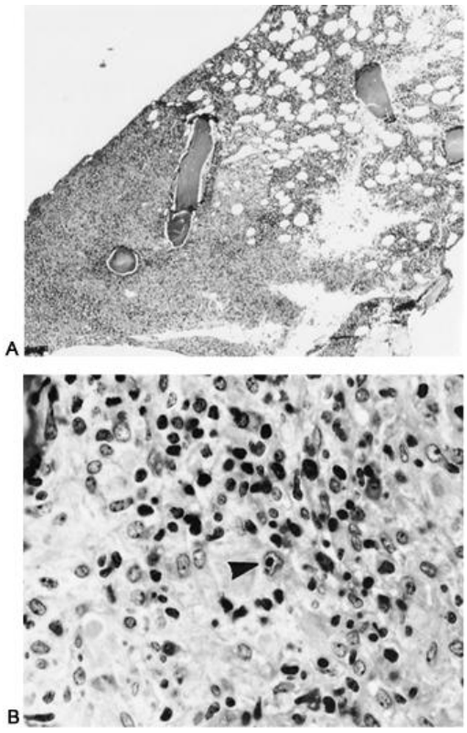


FIGURE 44.31. Hodgkin's disease, bone marrow. Note the extensive focal infiltrate in (**A**) with uninvolved normal marrow in the upper right field. Higher power (**B**) shows a mixed cellular infiltrate that is relatively lymphocyte depleted. A mononuclear Reed-Sternberg variant (half of a typical Reed-Sternberg cell) is indicated with an arrow. Hematoxylin and eosin, 25X (**A**), 600X (**B**).

Criteria for the diagnosis of Hodgkin's disease in the bone marrow are different depending on whether primary or secondary involvement is being diagnosed. When a diagnosis of Hodgkin's disease has been substantiated at another site, the requirement in the bone marrow is to see a cellular infiltrate consistent with Hodgkin's disease and at least one mononuclear variant of a Reed-Sternberg cell (Fig. 44.31B) (214, 215). To diagnose Hodgkin's disease primarily in the bone marrow, which may be necessary if the only other site of disease is in the retroperitoneum, several excellent typical Reed-Sternberg cells in addition to the typical infiltrate of Hodgkin's disease must be seen. Even then, one should be cautious, because non-Hodgkin's lymphoma can have Reed-Sternberg--like cells.

Lymphocyte-depleted Hodgkin's disease may present with marrow and retroperitoneal involvement and no easily accessible peripheral nodes on which a biopsy could be performed (218). Marker studies should be performed to confirm a primary diagnosis of Hodgkin's disease in the marrow. The background lymphocytes should predominantly express T-cell markers. The

Reed-Sternberg cells should be negative with CD45 and B- and T-cell markers, and be positive for CD30 and CD15 (8, 168, 213). Large cells associated with anaplastic large-cell non-Hodgkin's lymphomas that might mimic Reed-Sternberg cells are leukocyte common-antigen positive and should be T-cell or B-cell positive. They may also, however, be CD15 and CD30 positive, suggesting overlap between Hodgkin's disease and some anaplastic large-cell lymphomas (220).

Morphologic Variability and Differential Diagnosis

The morphologic variability that can be seen in Hodgkin's disease is listed in Table 44.22. In some cases of Hodgkin's disease, the marrow may be extensively fibrotic and mimic agnogenic myeloid metaplasia (Fig. 44.32) (219). Scattered lymphocytes within focal or diffuse areas of fibrosis should alert one to the possibility of Hodgkin's disease. When a fibrotic bone marrow infiltrate is seen in a patient with known Hodgkin's disease at another site, and an exhaustive search of adequate, stepped sectioned, bilateral trephine biopsy reveals no Reed-Sternberg cells or variants, the marrow should be considered suspicious for Hodgkin's disease. If the lesion contains an extensive, mixed cellular infiltrate that is otherwise typical of Hodgkin's disease, but no Reed-Sternberg variants, then the marrow may in consultation with the clinician be considered “consistent with, but not diagnostic of involvement” for purposes of staging.

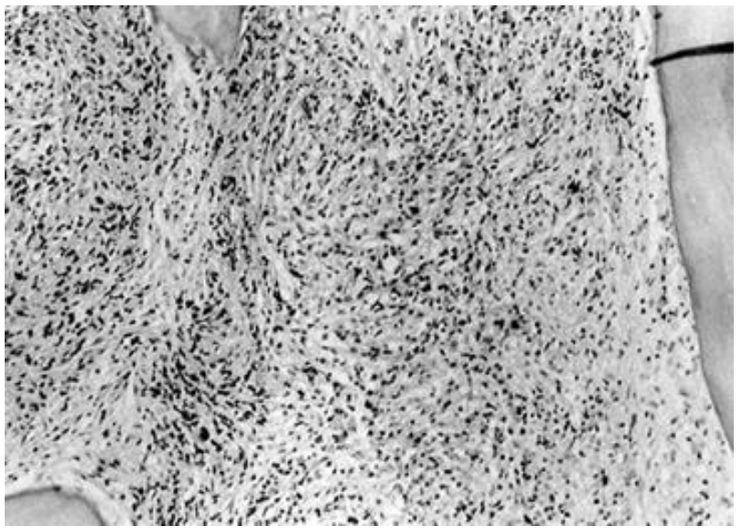


FIGURE 44.32. Hodgkin's disease showing extensive diffuse replacement of the marrow space by fibrosis. Lymphocytes admixed with the fibrosis suggested that this was Hodgkin's disease rather than agnogenic myeloid metaplasia, and Reed-Sternberg cells were identified after a careful search on multiple step sections. Hematoxylin and eosin, 120X.

Because of the mixed cellular infiltrate and the frequent presence of epithelioid histiocytes, the differential diagnosis of Hodgkin's disease in the bone marrow includes infectious disease, AIDS, and AILD. In some cases of Hodgkin's disease, granulomas may accompany Hodgkin's disease in the bone marrow; however, in other cases, nonspecific granulomas may be present in the bone marrow without marrow involvement by Hodgkin's disease (200).

Peripheral T-cell lymphomas are particularly prone to mimic Hodgkin's disease in the bone marrow (185). The critical morphologic feature to distinguish Hodgkin's disease from non-Hodgkin's lymphoma in the bone marrow, when a primary diagnosis is being considered on the marrow, is the appearance of the background lymphocytes. In Hodgkin's disease, as opposed to non-Hodgkin's lymphoma, these lymphocytes are nearly normal in appearance and do not show transitional forms to Reed-Sternberg cells. On the other hand, some variants of Hodgkin's disease may have numerous Reed-Sternberg cells and mimic an anaplastic large-cell lymphoma. This usually is more of a differential diagnostic problem with tissue biopsy than with bone-marrow biopsy.

Changes with Progression and Transformation

In general, the natural progression of Hodgkin's disease is toward the lymphocyte-depleted variant (See Table 44.24).

It is well documented that Hodgkin's disease is often associated with or develops non-Hodgkin's lymphoma (32). Cases of Hodgkin's disease that transform to large-cell lymphoma should be carefully studied to rule out the possibility of initial misdiagnosis of Hodgkin's disease or transformation of Hodgkin's disease to a syncytial or lymphocyte-depleted form of Hodgkin's disease with sheets of Reed-Sternberg cells (221, 222). Rare cases of development of ALLs also have been reported (32). Similar to other B-cell processes, myeloid leukemia can develop subsequent to therapy in Hodgkin's disease (32).

Staging, Therapy, and Prognostic Features

Survival in Hodgkin's disease is dependent on the stage of disease. Furthermore, morphologic subtypes of Hodgkin's disease correlate well with stage. Lymphocyte-predominant and nodular sclerosing forms generally correlate with a low incidence of marrow involvement, early stage of disease, and good survival, whereas mixed cellularity and lymphocyte-depleted Hodgkin's disease are correlated with a higher incidence of marrow involvement, a higher stage, and poorer survival. In general, therapy for early disease is local radiation with or without chemotherapy; for more advanced disease, therapy is combination chemotherapy with or without radiation. Transplantation is also used for younger patients with refractory disease.

Post-Transplant Lymphoproliferative Disorders

Post-transplant lymphoproliferative disorders usually develop months to years after transplantation and can be seen following solid-organ transplantation or bone-marrow transplantation (223, 224 and 225). The tumors arise secondary to severe immunocompromise, are usually of B-cell phenotype, and have been very highly associated with Epstein-Barr virus (EBV). They have many features in common with non-Hodgkin's lymphoma arising in severely immunocompromised hosts in general. They are particularly likely to develop in mismatched and T-cell–depleted bone-marrow transplant recipients.

Clinically, these disorders frequently develop as isolated extranodal disease but may present as widespread nodal disease, or they may rarely present with peripheral blood, bone marrow, or

CSF involvement (223). Morphologically, they show a spectrum, and three morphologic subsets have been described (225). One is that of a monomorphic-malignant process usually a large-cell lymphoma or less commonly plasma-cell myeloma. A second subset is described as a polymorphic lymphoproliferative disorder, containing a mixture of large cells (immunoblasts), immature plasma cells, mature plasma cells, and lymphocytes (Fig. 44.33A). Finally, a plasmacytic hyperplasia can be the tissue finding (225). When the former occurs on a tissue biopsy, and the diagnosis of malignant lymphoma is undisputed. However, when a polymorphous morphology is seen, distinction from a reactive process may be more difficult.

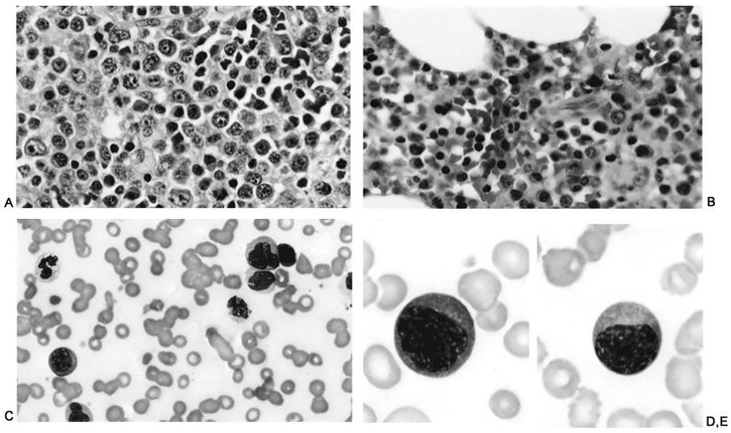


FIGURE 44.33. Post-transplant lymphoproliferative disorders. **A:** This polymorphous infiltrate is from a lymph node of a renal transplant patient. Note the spectrum of cells including lymphocytes, plasma cells, immature plasma cells, and large cells (immunoblasts). **B, C, D, E:** From a patient with a mismatched, T-cell depleted bone marrow transplant. Note the variety of large atypical plasmacytoid cells with deeply basophilic cytoplasm and dark, condensed nuclear chromatin. Hematoxylin and eosin, 600X (**A, B**). Wright's stain 600X (**C**), 1200X (**D, E**).

Post-transplantation lymphoproliferative disorders may occasionally involve the peripheral blood, bone marrow, or body fluid. The disorder should be strongly considered when large immature plasmacytoid cells are present in these sites in a patient at risk. These plasmacytoid cells are distinctly different in appearance from reactive (Downey) cells characteristic of infectious mononucleosis or cytomegolovirus (CMV) infection (Fig. 44.33B, Fig. 44.33C, Fig. 44.33D and Fig. 44.33E). This distinction is important because “Downey” cells may be associated with reactive processes in the post-transplant population. Immunophenotyping may be of some assistance in distinguishing reactive cells from cells suspicious for post-transplant lymphoproliferative disorder because Downey cells are T-cells, whereas the plasmacytoid cells of post-transplant lymphoproliferative disorders usually are B-cells, presumably transformed by EBV. Rare cases of T-cell post-transplant lymphoproliferative disorders have been described (225, 226).

The development of post-transplant lymphoproliferative disorders is thought to be a multistep process that starts in the immunocompromised host with EBV infection of B cells, followed by a polyclonal expansion of infected B cells (224). Out of this polyclonal expansion, oligoclonal B-cell proliferation can develop, presumably from lack of immune regulation. Subsequent to this, there may be eventual development of a monomorphic or dominant B-cell clone, which then behaves in a malignant fashion. If post-transplant lymphoproliferative disorders are studied for surface and cytoplasmic immunoglobulin at the polyclonal or oligoclonal stage, they are likely to express both κ and λ light chains, whereas the tumor at the monomorphic stage may show a clonally restricted light chain. An oligoclonal proliferation may be identified at an earlier stage by evaluating for clonal immunoglobulin gene rearrangement, clonal EBV incorporation, or clonal C-myc translocation (224, 225).

Therapeutic intervention in post-transplant lymphoproliferative disorders consists of removal of immunosuppression and treatment with acyclovir and interferon alpha. For more aggressive monomorphic disease combinations, chemotherapy may be

used. Prognosis for post-transplant lymphoproliferative disease is very poor; however, patients with less evidence of clonal transformation are more likely to respond to antiviral therapy and removal of immunosuppression than those with monomorphic and monoclonal proliferations (224).

Acknowledgment

The author would like to thank Barbara Ashbacher for her assistance in the preparation of this manuscript.

REFERENCES

1. Foucar K. Chronic lymphoid leukemias and lymphoproliferative disorders. *Mod Pathol* 1999;12:141–150

2. Gale RP, Foon KA. Biology of chronic lymphocytic leukemia. *Semin Hematol* 1987;24:209–229.

3. Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-Sponsored Working Group Guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. *Blood* 1996;12:4990–4997.

4. Spier CM, Kjeldsberg CR, Head DR, et al. Chronic lymphocytic leukemia in young adults. *Am J Clin Pathol* 1985;84:675–678.

5. Perreault C, Boileau J, Gyger M, et al. Chronic B-cell lymphocytosis. *Eur J Haematol* 1989;42:361–367.

6. Davis BH, Foucar K, Sczcarkowski W, et al. U.S.-Canadian recommendations on the immunophenotypic analysis of hematologic neoplasia by flow cytometry: medical indications. *Cytometry* 1997;30:249–263.

7. Jennings CD, Foon KA. Recent advances in flow cytometry: application to the diagnosis of hematologic malignancy. *Blood* 1997;90:2863–2892.

8. Harris NL, Jaffe ES, Stein H, et al. A Revised European-American Classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood* 1994;84:1361–1392.

9. Peterson LC, Bloomfield CD, Brunning RD. Relationship of clinical staging and lymphocyte morphology to survival in chronic lymphocytic leukaemia. *Br J Haematol* 1980;45:563–567.

10. Criel A, Verhoeff G, Vlietinck R, et al. Further characterization of morphologically defined typical and atypical CLL: a clinical, immunophenotypic, cytogenetic and prognostic study on 390 cases. *Br J Haematol* 1997;97:383–391.

11. Oscier DG, Matutes E, Copplestone A, et al. Atypical lymphocyte morphology: an adverse prognostic factor for disease progression in stage A CLL independent of trisomy 12. *Br J Haematol* 1997;98:934–939.

12. Dick FR. Chronic lymphocytic leukemia, prolymphocytic leukemia and leukemic non-Hodgkin's 1ymphoma. In: Koepke JA, ed. *Laboratory Hematology*. New York:Churchill Livingstone, 1984;1:325–357.

13. Melo JV, Catovsky D, Galton DAG. Chronic lymphocytic leukemia and prolymphocytic leukemia: a clinicopathologic reappraisal. *Blood Cells* 1987;12:339–353.

14. Ghani AM, Krause JR. Investigation of cell size and nuclear clefts as prognostic parameters in chronic lymphocytic leukemia. *Cancer* 1986;58:2233–2238.

15. Pombo De Oliveira MS, Jaffe ES, Catovsky D. Leukaemic phase of mantle zone (intermediate) lymphoma: its characterisation in 11 cases. *J Clin Pathol* 1989;42:962–972.

16. Vasef MA, Medeiros J, Koo C, et al. Cyclin D1 immunohistochemical staining is useful in distinguishing mantle cell lymphoma from other low grade B-cell neoplasms in bone marrow. *Am J Clin Pathol* 1997;108:302–307.

17. Dick FR, Maca R. The lymph node in chronic lymphocytic leukemia. *Cancer* 1978;41:283–292.

18. Ben-Ezra J, Burke JS, Swartz WG, et al. Small lymphocytic lymphoma: a clinicopathologic analysis of 268 cases. *Blood* 1989;73:579–587.

19. Kimura A, Hyodo H, Nakata Y, Kuramoto A. Chronic lymphocytic leukemia associated with bone marrow fibrosis: possible role of interleuken 1a in the pathogenesis. *Am J Hematol* 1993;43:47–50.

20. DiGiuseppe JA, Borowitz MJ. Clinical utility of flow cytometry in the chronic lymphoid leukemias. *Semin Oncol* 1998;25:6–10.

21. Moreau EJ, Matutes E, A'hern RP, et al. Improvement of the chronic lymphocytic leukemia scoring system with the monoclonal antibody SN8 (CD79b). *Am J Clin Pathol* 1997;108:378–382.

22. Huh YO, Pugh WC, Hagop M, et al. Detection of subgroups of chronic B-cell leukemias by FMC7 monoclonal antibody. *Am J Clin Pathol* 1994;101:283–289.

23. Tefferi A, Bartholmai BJ, Witzig TE, et al. Heterogeneity and clinical relevance of the intensity of CD20 and immunoglobulin light-chain expression in B-cell chronic lymphocytic leukemia. *Am J Clin Pathol* 1966;106:457–461.

24. Deegan MJ, Abraham JP, Sawdyk M, et al. High incidence of monoclonal proteins in the serum and urine of chronic lymphocytic leukemia patients. *Blood* 1984;64:1207–1211.

25. Finn WG, Maya T, Krishna KY, et al. Karyotype correlates with peripheral blood morphology and immunophenotype in chronic lymphocytic leukemia. *Am J Clin Pathol* 1966;105:458–467.

26. Matutes E, Oscier D, Garcia-Marco J, et al. Trisomy 12 defines a group of CLL with atypical morphology: correlation between cytogenetic, clinical and laboratory features in 54 patients. *Br J Haematol* 1966;92:382–388.

27. Braylan RC, Fowlkes BJ, Jaffe ES, et al. Cell volumes and DNA distributions of normal and neoplastic human lymphoid cells. *Cancer* 1978;41:201–209.

28. Diamond LW, Nathwani BN, Rappaport H. Flow cytometry in the diagnosis and classification of malignant lymphoma and leukemia. *Cancer* 1982;50:1122–1135.

29. Foroni L, Catovsky D, Luzzatto L. Immunoglobulin gene rearrangements in hairy cell leukemia and other chronic B cell lymphoproliferative disorders. *Leukemia* 1987;4:389–392.

30. Norton JD, Pattinson J, Hoffbrand AV, et al. Rearrangement and expression of T cell antigen receptor genes in B cell chronic lymphocytic leukemia. *Blood* 1988;71:178–185.

31. Kjeldsberg CR, Marty J. Prolymphocytic transformation of chronic lymphocytic leukemia. *Cancer* 1981;48:2447–2457.

32. York JC, Glick AD, Cousar JB, et al. Changes in the appearance of hematopoietic and lymphoid neoplasms: Clinical, pathologic, and biologic implications. *Hum Pathol* 1984;15:11–38.

33. Sun T, Susin M, Desner M, et al. The clonal origin of two cell populations in Richter's syndrome. *Hum Pathol* 1990;21:722–728.

34. Williams J, Schned A, Cotelingam JD, Jaffe ES. Chronic lymphocytic leukemia with coexistent Hodgkin's disease. Implications for the origin of the Reed Sternberg cell. *Am J Surg Pathol* 1991;15:33–42.

35. Brecher M, Banks PM. Hodgkin's disease variant of Richter's syndrome. *Am J Clin Pathol* 1990;93:333–339.

36. Herndier CJ, Glassberg AB, Hamill TR. A case of composite Hodgkin's Disease and chronic lymphocytic leukemia in bone marrow. Lack of Epstein-Barr virus. *Arch Pathol Lab Med* 1996;120:386–389.

37. Saltman DL, Ross JA, Banks RE, et al. Molecular evidence for a single clonal origin in biphenotypic concomitant chronic lymphocytic leukemia and multiple myeloma. *Blood* 1989;74:2062–2065.

38. Torelli UL, Torelli GM, Emilia G, et al. Simultaneously increased expression of the c-myc and µ-chain genes in the acute blastic transformation of a chronic lymphocytic leukaemia. *Br J Haematol* 1987;65:165–170.

39. Lai R, Arber DA, Brynes RK, et al. Untreated chronic lymphocytic leukemia concurrent with or followed by acute myelogenous leukemia or myelodysplastic syndrome. A report of five cases and review of the literature. *Am J Clin Pathol* 1999;111:373–378.

40. Geisler C, Ralfkiaer E, Hansen MM, et al. The bone marrow histological pattern has independent prognostic value in early stage chronic lymphocytic leukaemia. *Br J Haematol* 1986;62:47–54.

41. Stone, RM. Prolymphocytic leukemia. *Hematol Oncol Clin North Am* 1990;4:457–471.

42. Bearman RM, Pangalis GA, Rappaport H. Prolymphocytic leukemia: clinical, histopathological, and cytochemical observations. *Cancer* 1978;42:2360–2372.

43. Galton DAG, Goldman JM, Wiltshaw E, et al. Prolymphocytic leukaemia. *Br J Haematol* 1974;27:7–23.

44. Nieto LH, Lampert IA, Catovsky D. Bone marrow histological patterns in B-cell and T-cell prolymphocytic leukemia. *Hematol Pathol* 1989;3:79–84.

45. Lahuerta-Palacios JJ, Valdes MD, Navas-Palacios JN, et al. Six new cases of prolymphocytic leukemia with heterogeneous prognosis: clinical and immunologic features, light microscopy, and ultrastructural findings. *Cancer* 1985;55:2550–2557.

46. Brito-Babapulle V, Pittman S, Melo JV, et al. Cytogenetic studies on prolymphocytic leukemia: 1. B-cell prolymphocytic leukemia. *Hematol Pathol* 1987;1:27–33.

47. Lens D, Schouwer PJ, Hamoudi RA, et al. p53 abnormalities in B-cell prolymphocytic leukemia. *Blood* 1997;89:2015–2023.

48. Diamond LW, Bearman RM, Berry PK, et al. Prolymphocytic leukemia: flow microfluorometric, immunologic, and cytogenetic observations. *Am J Hematol* 1980;9:319–330.

49. Catovsky D, Golomb HM, Golde DW, guest eds. Special Issue: The 2nd International Workshop on Hairy Cell Leukemia. *Leukemia* 1987;1:284–408.

50. Strickler JG, Schmidt CM, Wick MR. Immunophenotype of hairy cell leukemia in paraffin sections. *Mod Pathol* 1990;3:518–523.

51. Hoyer JD, Li CY, Yam LT, et al. Immunohistochemical demonstration of acid phophatase isoenzyme 5 (tartrate-resistant) in paraffin sections of hairy cell leukemia and other hematologic disorders. *Am J Clin Pathol* 1997;108:308–315.

52. Yam LT, Phyliky RL, Li CY. Benign and neoplastic disorders simulating hairy cell leukemia. *Semin Oncol* 1984;11:353–361.

53. Woessner S, Lafuente R, Sans-Sabrafen J. Unusual cytological appearance of hairy cell leukemia. *Lancet* 1976;2:812.

54. Sainati L, Matutes E, Mulligan S, et al. A variant form of hairy cell leukemia resistant to α-interferon: clinical and phenotypic characteristics of 17 patients. *Blood* 1990;76:157–162.

55. Bartl R, Frisch B, Hill W, et al. Bone marrow histology in hairy cell leukemia: identification of subtypes and their prognostic significance. *Am J Clin Pathol* 1983;79:531–545.

56. Hanson CA, Ward PCJ, Schnitzer B. A multilobular variant of hairy cell leukemia with morphologic similarities to T-cell lymphoma. *Am J Surg Pathol* 1989;13:671–679.

57. Lee WM, Beckstead JH. Hairy cell leukemia with bone marrow hypoplasia. *Cancer* 1982;50:2207–2210.

58. Hanson CA, Gribbin TE, Schnitzer B, et al. CD11c (Leu-M5) expression characterizes a B-cell chronic lymphoproliferative disorder with features of both chronic lymphocytic leukemia and hairy cell leukemia. *Blood* 1990;76:2360–2367.

59. Agnarsson BA, Kadin ME. An unusual B-cell lymphoma simulating hairy cell leukemia. *Am J Clin Pathol* 1987;88:752–759.

60. Melo JV, Robinson DSF, Gregory C, et al. Splenic B cell lymphoma with “villous” lymphocytes in the peripheral blood: a disorder distinct from hairy cell leukemia. *Leukemia* 1987;1:294–299.

61. Katayama I, Yang JPS. Reassessment of a cytochemical test for differential diagnosis of leukemic reticuloendotheliosis. *Am J Clin Pathol* 1977;68:268–272.

62. Yam LT, Janckila AJ, Li CY, et al. Cytochemistry of tartrate resistant acid phosphatase: 15 years' experience. *Leukemia* 1987;1:285–288.

63. Anderson KC, Boyd AW, Fisher DC, et al. Hairy cell leukemia: a tumor of pre-plasma cells. *Blood* 1985;65:620–629.

64. Brito-Babapulle V, Pittman S, Melo JV, et al. The 14q+ marker in hairy cell 1eukaemia: a cytogenetic study of 15 cases. *Leuk Res* 1986;10:131–138.

65. Katayama I, Li CY, Yam LT. Ultrastructural characteristics of the “hairy cells” of leukemic reticuloendotheliosis. *Am J Pathol* 1972;67:361–370.

66. Golomb HM, Braylan R, Polliack A. “Hairy” cell leukaemia (leukaemic reticuloendotheliosis): a scanning electron microscopic study of eight cases. *Br J Haematol* 1975;29:455–460.

67. Jacobs RH, Vokes EE, Golomb HM. Second malignancies in hairy cell leukemia. *Cancer* 1985;56:1462–1467.

68. Davis KM, Spindel E, Franzine DA, et al. Anaplastic neoplasm in a patient with hairy cell leukemia. *Cancer* 1985;56:2470–2475.

69. Nazeer T, Burkart P, Dunn H, et al. Blastic transformation of hairy cell leukemia. *Arch Pathol Lab Med* 1997;121:707–713.

70. VanderMolen IA, Urba WJ, Longo DL, et al. Diffuse osteosclerosis in hairy cell leukemia. *Blood* 1989;74:2066–2069.

71. Demanes DJ, Lane N, Beckstead JH. Bone involvement in hairy-cell leukemia. *Cancer* 1982;49:1697–1701.

72. Mercieca J, Matutes E, Moskovic E, et al. Massive abdominal lymphadenopathy in hairy cell leukemia: a report of 12 cases. *Br J Haematol* 1992;82:547–554.

73. Knowles DM II. The human T-cell leukemias: clinical, cytomorphologic, immunophenotypic, and genotypic characteristics. *Hum Pathol* 1986;17:14–33.

74. Berliner N. T gamma lymphocytosis and T cell chronic leukemias. *Hematol Oncol Clin North Am* 1990;4:473–487.

75. Knowles DM. Immunophenotypic and antigen receptor gene rearrangement analysis in T cell neoplasia. *Am J Pathol* 1989;134:761–785.

76. Kingreen D, Seigert W. Chronic lymphatic leukemias of T and NK type. *Leukemia* 1997;11(suppl 2):46–49.

77. Ginaldi L, Matutes E, Farahat N, et al. Differential expression of CD3 and CD7 in T-cell malignancies: a quantitative study by flow cytometry. *Br J Haematol* 1996;93:921–927.

78. Hoyer JD, Ross CW, Witzig TE, et al. True T-cell chronic lymphocytic leukemia: a morphologic and immunophenotypic study of 25 cases. *Blood* 1995;86:1163–1169.

79. Hui PK, Feller AC, Pileri S, et al. New aggressive variant of suppressor/cytotoxic T-CLL. *Am J Clin Pathol* 1987;87:55–59.

80. Matutes E, Catovsky D. Similarities between T-cell chronic lymphocytic leukemia and the small-cell variant of T-prolymphocytic leukemia (letter). *Blood* 1996;87:3520–3521.

81. Matutes E, Talavera G, O'Brien M, et al. The morphological spectrum of T-prolymphocytic leukaemia. *Br J Haematol* 1986;64:111–124.

82. Garand R, Goasguen J, Brizard A, et al. Indolent course as a relatively frequent presentation in T-prolymphocytic leukemia. *Br J Haematol* 1998;103:488–494.

83. Brito-Babapulle V, Pomfret M, Matutes E. Cytogenetic studies on prolymphocytic leukemia: II. T cell prolymphocytic leukemia. *Blood* 1987;70:926–931.

84. Pandolfi F, Loughran TP, Starkebaum G, et al. Clinical course and prognosis of the lymphoproliferative disease of granular lymphocytes: a multicenter study. *Cancer* 1990;65:341–348.

85. Agnarsson BA, Loughran TP, Starkebaum G, et al. The pathology of large granular lymphocyte leukemia. *Hum Pathol* 1989;20:643–651.

86. Scott CS, Richards SJ, Sivakumaran M, et al. Transient and persistent expansions of large granular lymphocytes (LGL) and NK-associated (NKa) cells: the Yorkshire Leukemia Group study. *Br J Haematol* 1993;83:504–515.

87. Ohno Y, Amakawa R, Fukuhara S, et al. Acute transformation of chronic large granular lymphocyte leukemia associated with additional chromosome abnormality. *Cancer* 1989;64:62–67.

88. Imamura N, Kusunoki Y, Kawa-Ha K, et al. Aggressive natural killer cell leukaemia/lymphoma: sport of four cases and review of the literature: possible existence of a new clinical entity originating from the third lineage of lymphoid cells. *Br J Haematol* 1990;75:49–59.

89. Longacre TA, Listrom MB, Spigel JH, et al. Aggressive jejunal lymphoma of large granular lymphocytes: immunohistochemical, ultrastructural, molecular, and DNA content analysis. *Am J Clin Pathol* 1990;93:124–132.

90. Nagatani T, Matsuzaki T, Iemoto G, et al. Comparative study of cutaneous T-cell lymphoma and adult T-cell leukemia/lymphoma: clinical, histopathologic, and immunohistochemical analyses. *Cancer* 1990;66:2380–2386.

91. Salhany KE, Greer JP, Cousar JB, et al. Marrow involvement in cutaneous T-cell lymphoma. *Am J Clin Pathol* 1989;92:747–754.

92. Pinto A, Zagonts el V, Carbone A. Bone marrow involvement in cutaneous T-cell lymphoma. *Am J Clin Pathol* 1990;94:119–120.

93. Willemze R, Van Vloten WA, Hermans J, et al. Diagnostic criteria in Sézary's syndrome: a multiparameter study of peripheral blood lymphocytes in 32 patients with erythroderma. *J Invest Dermatol* 1983;81:392–397.

94. Vonderheid EC, Sobel EL, Nowell PC, et al. Diagnostic and prognostic significance of Sézary cells in peripheral blood smears from patients with cutaneous T cell 1ymphoma. *Blood* 1985;66:358–366.

95. Bogen SA, Pelley D, Charif M, et al. Immunophenotypic identification of Sézary cells in peripheral blood. *Am J Clin Pathol* 1996;106:739–748.

96. Muche JM, Lukowski A, Asadullah K, et al. Demonstration of frequent occurrence of clonal T-cells in the peripheral blood of patients with primary cutaneous T-cell lymphoma. *Blood* 1997;90:1636–1642.

97. Shimamoto Y, Suga K, Nishimura J, et al. Major prognostic factors of Japanese patients with lymphoma-type adult T-cell leukemia. *Am J Hematol* 1990;35:232–237.

98. Proceedings of the International Conference on Multiple Myeloma Biology, Pathophysiology, Prognosis and Treatment, June 19–22, 1989, Bologna (Italy). *Eur J Haematol* 1989;43(suppl 51):11–195.

99. Hallek M, Leif Bergsagel P, Anderson KC. Multiple Myeloma: increasing evidence for a multistep transformation process. *Blood* 91 1998;91:3–21.

100. Dick FR. Plasma cell myeloma and related disorders with monoclonal gammopathy. In: Koepke JA, ed. *Laboratory Hematology*. New York: Churchill Livingstone, 1984:445–481.

101. Buss DH, Prichard RW, Hartz JW, et al. Initial bone marrow findings in multiple myeloma: significance of plasma cell nodules. *Arch Pathol Lab Med* 1986;110:30–33.

102. Kyle RA, Greipp PR. Smoldering multiple myeloma. *N Engl J Med* 1980;302:1347–1349.

103. Crowley JP. Smoldering multiple myeloma. *N Engl J Med* 1980;303:941.

104. Durie BGM. Staging and kinetics of multiple myeloma. *Semin Oncol* 1986;13:300–309.

105. Ong F, Hermans J, Noordijk EM and Kluin-Nelemans JC. Is the Durie and Salmon diagnostic classification system for plasma cell dyscrasias still the best choice? Application of three classification systems to a large population-based registry of paraproteinemia and multiple myeloma. *Ann Hematol* 1975;70:19–24.

106. Reed M, McKenna RW, Bridges R, et al. Morphologic manifestations of monoclonal gammopathies. *Am J Clin Pathol* 1981;76:8–23.

107. Brunning RD, Parkin J. Intranuclear inclusions in plasma cells and lymphocytes from patients with monoclonal gammopathies. *Am J Clin Pathol* 1976;66:10–21.

108. Hsu SU, Hsu PL, McMillan PN, et al. Russell bodies: a light and electron microscopic immunoperoxidase study. *Am J Clin Pathol* 1982;77:26–31.

109. Gabriel L, Escribano L, Perales J, et al. Multiple myeloma with crystalline inclusions in most hemopoietic cells. *Am J Hematol* 1985;18:405–411.

110. Maldonado JE, Brown AL, Bayrd ED, et al. Cytoplasmic and intranuclear electron-dense bodies in the myeloma cell. *Arch Pathol* 1966;81:484–499.

111. Greipp PR, Raymond NM, Kyle RA, et al. Multiple myeloma: significance of plasmablastic subtype in morphological classification. *Blood* 1985;65:305–310.

112. Schambeck CM, Bartl R, Hochtlen-Vollmar W, et al. Characterization of myeloma cells by means of labeling index, bone marrow histology, and serum beta 2-microglobulin. *Am J Clin Pathol* 1996;106:64–68.

113. Murakami H, Nemoto K, Sawamura M, et al. Prognostic relevance of morphological classification in multiple myeloma. *Acta Haematol* 1992;87:113–117.

114. Strand WR, Banks PM, Kyle RA. Anaplastic plasma cell myeloma and immunoblastic lymphoma: clinical, pathologic, and immunologic comparison. *Am J Med* 1984;76:861–867.

115. Sun NCJ, Fishkin BG, Nies KM, et al. Lymphoplasmacytic myeloma: an immunological, immunohistochemical and electron microscopic study. *Cancer* 1979;43:2268–2278.

116. Zukerberg I-R, Ferry JA, Conlon M, et al. Plasma cell myeloma with cleaved, multilobated, and monocytoid nuclei. *Am J Clin Pathol* 1990;93:657–661.

117. Kanoh T, Fujii H. Phagocytic myeloma cells: report of a case and review of the literature. *Am J Clin Pathol* 1985;84:121–124.

118. Krzyzaniak RL, Buss DH, Cooper R, et al. Marrow fibrosis and multiple myeloma. *Am J Clin Pathol* 1988;89:63–68.

119. Raman S, Frame B, Saeed SM, et al. Diffuse nonsecretory osteosclerotic myeloma with extensive erythrophagocytosis. *Am J Clin Pathol* 1983;80:84–88.

120. Lacy MQ, Gertz MA, Hanson CA et al. Multiple myeloma associated with diffuse osteosclerotic bone lesions: a clinical entity distinct from osteosclerotic myeloma (POEMS syndrome). *Am J Hematol* 1997;56:288–293.

121. Stone, MJ. Amyloidosis: a final common pathway for protein deposition in tissues. *Blood* 1990;75:531–545.

122. Kyle RA, Maldonado JE, Bayrd ED. Plasma cell leukemia: report on 17 cases. *Arch Intern Med* 1974;133:813–818.

123. Witzig TE, Dhodapkar MV, Kyle RA, Greipp PR. Quantitation of circulating peripheral blood plasma cells and their relationship to disease activity in patients with multiple myeloma. *Cancer* 1993;72:108–113.

124. Kyle RA, Bayrd ED. *The Monoclonal Gammopathies: Multiple Myeloma and Related Plasma Cell Disorders*. Springfield, IL: Charles C Thomas, 1979.

125. Peterson LC, Brown BA, Crosson JT, et al. Application of the immunoperoxidase technic to bone marrow trephine biopsies in the classification of patients with monoclonal gammopathies. *Am J Clin Pathol* 1986;85:688–693.

126. Wolf BC, Brady K, O'Murchadha MT, et al. An evaluation of immunohistologic stains for immunoglobulin light chains in bone marrow biopsies in benign and malignant plasma cell proliferations. *Am J Clin Pathol* 1990;94:742–746.

127. Hitzman JL, Li CY, Kyle RA. Immunoperoxidase staining of bone marrow sections. *Cancer* 1981;48:2438–2446.

128. Wolf BC, Kumar A, Vera JC, et al. Bone marrow morphology and immunology in systemic amyloidosis. *Am J Clin Pathol* 1986;86:84–88.

129. Durie BGM, Stock-Novack D, Salmon SE, et al. Prognostic value of pretreatment serum β_2-microglobulin in myeloma: a southwest oncology group study. *Blood* 1990;75:823–830.

130. Greipp PR, Katzmann JA, O'Fallon WM, et al. Value of β_2-microglobulin level and plasma cell labeling indices as prognostic factors in patients with newly diagnosed myeloma. *Blood* 1988;72:219–223.

131. Greipp PR, Witzig TE, Gonchoroff NJ, et al. Immunofluorescence labeling indices in myeloma and related monoclonal gammopathies. *Mayo Clin Proc* 1987;62:969–977.

132. Witzig TE, Gonchoroff NJ, Katzmann JA, et al. Peripheral blood B cell labeling indices are a measure of disease activity in patients with monoclonal gammopathies. *J Clin Oncol* 1988;6:1041–1046.

133. Gould J, Alexanian R, Goodacre A, et al. Plasma cell karyotype in multiple myeloma. *Blood* 1988;71:453–456.

134. Zandecki M, Lai JL, Facon T. Multiple myeloma: almost all patients are cytogenetically abnormal. *Br J Haematol* 1996;94:217–227.

135. Barlogie B, Alexanian R, Dixon D, et al. Prognostic implications of tumor cell DNA and RNA content in multiple myeloma. *Blood* 1985;66:338–341.

136. Lokhorst HM, Boom SE, Terpstra W, et al. Determination of the growth fraction in monoclonal gammopathy with the monoclonal antibody Ki-67. *Br J Haematol* 1988;69:477–481.

137. Zalcberg JR, Cornell FN, Ireton JC, et al. Chronic lymphatic leukemia developing in a patient with multiple myeloma: immunologic demonstration of a clonally distinct second malignancy. *Cancer* 1982;50:594–597.

138. Foucar K, McKenna RW, Bloomfield CD, et al. Therapy related leukemia: a panmyelosis. *Cancer* 1979;43:1285–1296.

139. Mufti GJ, Hamblin TJ, Clein GP, et al. Coexistent myelodysplasia and plasma cell neoplasia. *Br J Haematol* 1983;54:91–96.

140. Durie BGM, Salmon SE. A clinical staging system for multiple myeloma. *Cancer* 1975;36:842–854.

141. Pasqualetti P, Casale R, Collacciani A, et al. Multiple myeloma: relationship between survival and cellular morphology. *Am J Hematol* 1990;33:145–147.

142. Carbone A, Volpe R, Manconi R, et al. Bone marrow pattern and clinical staging in multiple myeloma (Letter). *Br J Haematol* 1987;65:502–503.

143. Knowling MA, Harwood AR, Bergsagel DE. Comparison of extramedullary plasmacytomas with solitary and multiple plasma cell tumors of bone. *J Clin Oncol* 1983;1:255–262.

144. Meis JM, Butler JJ, Osborne BM, et al. Solitary plasmacytomas of bone and extramedullary plasmacytomas: a clinicopathologic and immunohistochemical study. *Cancer* 1987;59:1475–1485.

145. Kyle RA, Garton JP. The spectrum of IgM monoclonal gammopathy in 430 cases. *Mayo Clin Proc* 1987;62:719–731.

146. Bartl R, Frisch B, Mahl G, et al. Bone marrow histology in Waldenström's macroglobulinaemia: clinical relevance of subtype recognition. *Scand J Haematol* 1983;31:359–375.

147. Andriko JA, Aguilera NS, Chu WS et al. Waldenström's macroglobulinemia: a clinicopathologic study of 22 cases. *Cancer* 1997;80:1926–1935.

148. Rywlin AM, Civantos F, Ortega RS, et al. Bone marrow histology in monoclonal macroglobulinemia. *Am J Clin Pathol* 1975;63:769–778.

149. Moore DF, Migliore PJ, Shullenberger CC, et al. Monoclonal macroglobulinemia in malignant lymphoma. *Ann Intern Med* 1970;72:43–47.

150. Tursz T, Brouet JC, Flandrin G, et al. Clinical and pathologic features of Waldenström's macroglobulinemia in seven patients with serum monoclonal IgG or IgA. *Am J Med* 1977;63:499–502.

151. Feiner HD, Rizk CC, Finfer MD, et al. IgM monoclonal gammopathy/Waldenström's macroglobulinemia: a morphological and immunophenotypic study of the bone marrow. *Mod Pathol* 1990;3:348–356.

152. Chelazzi G, Bettini R, Pinotti G. Bone-marrow-patterns and survival in Waldenström's macroglobulinaemia. *Lancet* 1979;2:965–966.

153. Seligmann M, Mihaesco E, Preud'Homme JL, et al. Heavy chain diseases: current findings and concepts. *Immunol Rev* 1979;48:145–167.

154. Kyle RA, Greipp PR, Banks PM. The diverse picture of gamma heavy-chain disease: report of seven cases and review of literature. *Mayo Clin Proc* 1981;56:439–451.

155. Brouet JC, Seligmann M, Danon F, et al. μ-Chain disease: report of two new cases. *Arch Intern Med* 1979;139:672–674.

156. Galian A, Lececstre MJ, Scotto J, et al. Pathological study of alpha-chain disease, with special emphasis on evolution. *Cancer* 1977;39:2081–2101.

157. Ohno YI, Koya M, Yasuda N, et al. Ultrastructure of vacuolated plasma cells in macroglobulinemia associated with production of mu-chain fragment. *Cancer* 1982;49:2489–2492.

158. Reyes F, Piquet J, Gourdin MF, et al. Immunoblastic lymphoma involving the bone marrow in a patient with alpha chain disease: clinical and immunoelectron microscopic study. *Cancer* 1985;55:1007–1014.

159. Kyle RA, Lust JA. Monoclonal gammopathies of undetermined significance. *Semin Hematol* 1989;26:176–200.

160. Lichtman MA. Essential and secondary monoclonal gammopathies. In: Williams WJ, Bently E, Erslev AJ, et al. eds. *Hematology*. 4th ed. New York: McGraw-Hill, 1990:1109–1114.

161. Axelsson U, Bachmann R, Hallen J. Frequency of pathological proteins (M-components) in 6,995 sera from an adult population. *Acta Med Scand* 1966;179:235–247.

162. Papadopoulos NM, Elin RJ, Wilson DM. Incidence of γ-globulin banding in a healthy population by high-resolution electrophoresis. *Clin Chem* 1982;28:707–708.

163. Kyle RA. Monoclonal gammopathy of undetermined significance: natural history in 241 cases. *Am J Med* 1978;64:814–826.

164. Majumdar G, Heard SE, Singh AK. Use of cytoplasmic 5-nucleotidase for differentiating malignant from benign monoclonal gammopathies. *J Clin Pathol* 1990;43:891–892.

165. Bezwoda WR, Gordon V, Bagg A, et al. Light chain restriction analysis of bone marrow plasma cells in patients with MGUS or “solitary” plasmacytoma: diagnostic value and correlation with clinical course. *Br J Haematol* 1990;74:420–423.

166. Majumdar G, Singh AK. Excess of monoclonal plasma cells in MGUS and diagnosis of early myeloma (Letter). *Br J Haematol* 1990;74:441–442.

167. Jaffe ES, ed. *Surgical Pathology of the Lymph Nodes and Related Organs*. 2nd ed. Philadelphia: WB Saunders, 1991.

168. Harris NL. Hodgkin's disease: classification and differential diagnosis. *Mod Pathol* 1999;12:159–176.

169. The Non-Hodgkin's Lymphoma Classification Project. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. *Blood* 1997;89:3909–3918.

170. Foucar K, Dick FR. Histopathology of non-Hodgkin's lymphoma and Hodgkin's disease. In: Keopke JA, ed. *Laboratory Hematology*. New York: Churchill-Livingstone, 1984:359–415.

171. Dick F, Bloomfield CD, Brunning RD. Incidence, cytology, and histopathology of non-Hodgkin's lymphomas in the bone marrow. *Cancer* 1974;33:1382–1398.

172. Kluin PM, Van Krieken JH, Kleiverda K, et al. Discordant morphologic characteristics of B-cell lymphomas in bone marrow and lymph node biopsies. *Am J Clin Pathol* 1990;94:59–66.

173. Fisher DE, Jacobson JO, Ault KA, et al. Diffuse large cell lymphoma with discordant bone marrow histology: clinical features and biological implications. *Cancer* 1989;64:1879–1887.

174. Conlan MG, Bast M, Armitage JO, et al. Bone marrow involvement by non-Hodgkin's lymphoma: the clinical significance of morphologic discordance between the lymph node and bone marrow. *J Clin Oncol* 1990;8:1163–1173.

175. Crisan D, Mattson JC. Discordant morphologic features in bone marrow involvement by malignant lymphoma. *Am J Hematol* 1995;49:299–309.

176. Skinnider BF, et al. Bone marrow involvement in T-cell-rich B-cell lymphoma. *Am J Clin Pathol* 1997; 108:570–578.

177. Swerdlow SH. Small B-cell lymphomas of the lymph nodes and spleen: practical insights to diagnosis and pathogenesis. *Mod Pathol* 1999;12:125–140.

178. Brynes RK, McKenna RW, Sundberg RD. Bone marrow aspiration and trephine biopsy: an approach to a thorough study. *Am J Clin Pathol* 1978;70:753–759.

179. Ebie N, Lowe JM, Gregory SA. Bilateral trephine bone marrow biopsy for staging non-Hodgkin's lymphoma/a second look. *Hematol Pathol* 1989;3:29–33.

180. Barekman CL, et al. Comparative utility of diagnostic bone marrow components: a ten year study. *Am J Hematol* 1997;56:37–41.

181. Jacobs, P. Core length in bone-marrow biopsy. *Lancet* 1979;1:1405–1406.

182. Maeda K, Hyun BH, Rebuck JW. Lymphoid follicles in bone marrow aspirates. *Am J Clin Pathol* 1977;67:41–48.

183. Pangalis GA, Moran EM, Rappaport H. Blood and bone marrow findings in angioimmunoblastic lymphadenopathy. *Blood* 1978;51:71–83.

184. Sun NCJ, Shapshak P, Lachant NA, et al. Bone marrow examination in patients with AIDS and AIDS-related complex (ARC). *Am J Clin Pathol* 1989;92:589–594.

185. Hanson CA, Brunning RD, Gajl-Peczalska KJ, et al. Bone marrow manifestations of peripheral T-cell lymphoma. *Am J Clin Pathol* 1986;86:449–460.

186. McKenna RW, Brunning RD. Reed-Sternberg-like cells in nodular lymphoma involving the bone marrow. *Am J Clin Pathol* 1975;63:779–785.

187. Shin SS, Ben-Ezra J, Burke JS et al. Reed Sternberg-like cells in low-grade lymphomas are transformed neoplastic cells of B-cell lineage. *Am J Clin Pathol* 1993; 99:658–662.

188. Contos MJ, Kornstein MJ, Innes DJ and Ben-Ezra J. The utility of CD20 and CD42 in subclassification of low-grade B-cell lymphoma on paraffin sections. *Mod Pathol* 1992;5:631–633.

189. Dunphy CH, Wheaton SE and Perkins SL. CD23 expression in transformed small lymphocytic lymphomas/chronic lymphocytic leukemias and blastic transformation of mantle cell lymphoma. *Mod Pathol* 1997;10:818–822.

190. Fineberg S, Marsh E, Alfonso F, et al. Immunophenotypic evaluation of bone marrow in non-Hodgkin's lymphoma. *Hum Pathol* 1993;24:636–642.

191. Kumar S, Green GA, Teruya-Feldstein J, et al. Use of CD23 (Bu38) on paraffin sections in the diagnosis of small lymphocytic lymphoma and mantle cell lymphoma. *Mod Pathol* 1996;9:925–929.

192. Perry, DA, Bast MA, Armitage JO, et al. Diffuse intermediate lymphocytic lymphoma: a clinicopathologic study and comparison with small lymphocytic lymphoma and diffuse small cleaved cell lymphoma. *Cancer* 1990;66:1995–2000.

193. Lardelli P, Bookman MA, Sundeen J, et al. Lymphocytic lymphoma of intermediate differentiation: morphologic and immunophenotypic spectrum and clinical correlations. *Am J Surg Pathol* 1990;14:752–763.

194. Wasman J, Rosenthal NS, Farhi DC. Mantle cell lymphoma: morphologic findings in bone marrow involvement. *Am J Clin Pathol* 1966; 106:196–200.

195. Neiman RS, Sullivan AL, Jaffe R. Malignant lymphoma simulating leukemic reticuloendotheliosis: a clinicopathologic study of ten cases. *Cancer* 1979;43:329–342.

196. Lukes RJ, Collins RD. Immunological characterization of human malignant lymphomas. *Cancer* 1974;34:1488–1503.

197. McKenna RW, Bloomfield CD, Brunning RD. Nodular lymphoma: bone marrow and blood manifestations. *Cancer* 1975;36:428–440.

198. Spiro S, Galton D, Wiltshaw E, et al. Follicular lymphoma: a survey of 75 cases with special reference to the syndrome resembling chronic lymphocytic leukemia. *Br J Cancer* 1975;31:60–72.

199. Foucar K, Armitage JO, Dick FR. Malignant lymphoma mixed small and large cell: a clinicopathologic study of 47 cases. *Cancer* 1983;51:2090–2099.

200. Bhargava V, Farhi DC. Bone marrow granulomas: clinicopathologic findings in 72 cases and review of the literature. *Hematol Pathol* 1988;2:43–50.

201. Kubic VL, Brunning RD. Immunohistochemical evaluation of neoplasms in bone marrow biopsies using monoclonal antibodies reactive in paraffin-embedded tissue. *Mod Pathol* 1989;2:618–629.

202. Horny HP, Engst U, Walz RS, et al. In situ immunophenotyping of lymphocytes in human bone marrow: an immunohistochemical study. *Br J Haematol* 1989;71:313–321.

203. Bluth RF, Casey TT, McCurley TL. Differentiation of reactive from neoplastic small-cell lymphoid aggregates in paraffin-embedded marrow particle preparations using L-26 (CD20) and UCHL-1 (CD45RO) monoclonal antibodies. *Am J Clin Pathol* 1993;99:150–156.

204. Horning SJ, Galili N, Cleary M, et al. Detection of non-Hodgkin's lymphoma in the peripheral blood by analysis of antigen receptor gene rearrangements: results of a prospective study. *Blood* 1990;75:1139–1145.

205. Coad JE, Olson DJ, Christensen DR, et al. Correlation of PCR-detected clonal gene rearrangements with bone marrow morphology in patients with B-lineage lymphomas. *Am J Surg Pathol* 1997;21:1047–1056.

206. Naughton MJ, Hess JL, Zutter MM, Bartlett NL. Bone marrow staging in patients with non-Hodgkin's lymphoma: Is flow cytometry a useful test? *Cancer* 1998;82:1154–115.

207. Rowley JD. Recurring chromosome abnormalities in leukemia and lymphoma. *Semin Hematol* 1990;27:122–136.

208. Grogan TM, Lippman SM, Spier CM, et al. Independent prognostic significance of a nuclear proliferation antigen in diffuse large cell lymphomas as determined by the monoclonal antibody Ki-67. *Blood* 1988;71:1157–1160.

209. Weiss LM, Strickler JG, Medeiros LJ, et al. Proliferative rates of non-Hodgkin's lymphomas as assessed by Ki-67 antibody. *Hum Pathol* 1987;18:1155–1159.

210. Kamat D, Laszewski MJ, Kemp JD, et al. The diagnostic utility of immunophenotyping and immunogenotyping in the pathologic evaluation of lymphoid proliferations. *Mod Pathol* 1990;3:105–112.

211. Ellison DJ, Hu E, Zovich D, et al. Immunogenetic analysis of bone marrow aspirates in patients with non-Hodgkin lymphomas. *Am J Hematol* 1990;33:160–166.

212. Sandhaus LM, Voelkerding KV, Dougherty J, et al. Combined utility of gene rearrangement analysis and flow cytometry in the diagnosis of lymphoproliferative disease in the bone marrow. *Hematol Pathol* 1990;4:135–148.

213. Anastasi J, Bitter MA, Vardiman JW. The histopathologic diagnosis and subclassification of Hodgkin's disease. *Hematol Oncol Clin North Am* 1989;3:187–204.

214. Lukes RJ. Criteria for involvement of lymph node, bone marrow, spleen, and liver in Hodgkin's disease. *Cancer Res* 1971;31:1755–1767.

215. Weiss RB, Brunning RD, Kennedy BJ. Hodgkin's disease in the bone marrow. *Cancer* 1975;36:2077–2083.

216. Myers CE, Chabner BA, De Vita VT, et al. Bone marrow involvement in Hodgkin's disease: pathology and response to MOPP chemotherapy. *Blood* 1974;44:197–204.

217. Bartl R, Frisch B, Burkhards R, et al. Assessment of bone marrow histology in Hodgkin's disease: correlation with clinical factors. *Br J Haematol* 1982;51:345–360.

218. Kinney MC, Greer JP, Stein RS, et al. Lymphocyte-depletion Hodgkin's disease: histopathologic diagnosis of marrow involvement. *Am J Surg Pathol* 1986;10:219–226.

219. Meadows LM, Rosse WR, Moore JO, et al. Hodgkin's disease presenting as myelofibrosis. *Cancer* 1989;64:1720–1726.

220. Leoncini L, Del Vecchio MT, Kraft R, et al. Hodgkin's disease and CD30-positive anaplastic large cell lymphomas a continuous spectrum of malignant disorders: a quantitative morphometric and immunohistologic study. *Am J Pathol* 1990;137:1047–1057.

221. Casey TT, Cousar JB, Mangum M, et al. Monomorphic lymphomas arising in patients with Hodgkin's disease: correlation of morphologic, immunophenotypic, and molecular genetic findings in 12 cases. *Am J Pathol* 1990;136:81–94.

222. Hansmann ML, Stein H, Fellbaum C, et al. Nodular paragranuloma can transform into high-grade malignant lymphoma of B type. *Hum Pathol* 1989;20:1169–1175.

223. Davey DD, Kamat D, Laszewski M, et al. Epstein-Barr virus-related lymphoproliferative disorders following bone marrow transplantation: an immunologic and genotypic analysis. *Mod Pathol* 1989;2:27–34.

224. Locker J, Nalesnik M. Molecular genetic analysis of lymphoid tumors arising after organ transplantation. *Am J Pathol* 1989;135:977–987.

225. Knowles DM. Immunodeficiency-associated lymphoproliferative disorders. *Mod Pathol*; 1999;12:200–217.

226. Zutter MM, Durnam DM, Hackman RC, et al. Secondary T-cell lymphoproliferation after marrow transplantation. *Am J Clin Pathol* 1990;94:714–721.

45
Chronic Myeloproliferative Disorders

Jane C. Huang

Hesham M. Amin

Serhan Alkan

Chronic myeloproliferative disorders (CMPDs) refer to a group of related primary disorders of the bone marrow that morphologically manifests as proliferations of differentiated myeloid cells. However, they are actually neoplastic, stem cell disorders, as proven by cytogenetic and molecular studies. CMPDs share overlapping features, but can be differentiated most of the time from each other by clinical, pathologic, and genetic findings. The four major categories of CMPDs consist of chronic myelogenous leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET), and chronic idiopathic myelofibrosis (CIMF). A very simplistic way of distinguishing these disorders is to think of the predominant myeloid line that is manifested: granulocytes in CML, erythroid cells in PV, and megakaryocytes/platelets in ET. Other less common processes include hypereosinophilic syndrome/chronic eosinophilic leukemia and chronic neutrophilic leukemia. Chronic myelomonocytic leukemia, atypical chronic myeloid leukemia, and juvenile myelomonocytic leukemia were previously classified as CMPDs, but the new World Health Organization (WHO) classification of myeloid neoplasms has included them under a new category of myelodysplastic/myeloproliferative disorders (1).

It is important to remember the similar features that justify lumping these diseases together in the general category of CMPDs. CMPDs generally present with splenomegaly and peripheral blood cytosis. Significant dysplasia is usually not associated with CMPDs. The bone marrow frequently is fibrotic to various degrees, initially reticulin fibrosis and eventually progressing to collagen fibrosis. Megakaryocytes often are increased and characteristically clustered. The relationship between myelofibrosis and megakaryocytes has been extensively studied (see section on chronic idiopathic myelofibrosis).

These general features are useful in differentiating CMPDs from myelodysplastic syndromes. Like CMPDs, the bone marrow in myelodysplastic syndromes is hypercellular. However, myelodysplastic syndromes usually present with peripheral cytopenias and morphologic dysplasia without significant myelofibrosis or splenomegaly. Nevertheless, there remain some cases in which the distinction might be very difficult, as recognized by the WHO with the introduction of a new category, myelodysplastic/myeloproliferative disorders (1).

CHRONIC MYELOGENOUS LEUKEMIA

Part of "45 - Chronic Myeloproliferative Disorders"

CML is the best characterized and most common of all the CMPDs. This disease is defined by a molecular marker, the Philadelphia chromosome (Ph) with the breakpoint cluster region-Abelson murine leukemia virus (BCR-ABL) fusion gene that results from the translocation of chromosomes 9 and 22 [t(9;22) (q34;q11)]. Although CML is a stem cell disorder, it manifests as a predominant granulocytic proliferation. Therefore, some authorities prefer the term chronic granulocytic leukemia to refer specifically to this type of CMPD with the BCR-ABL rearrangement (2). It is important to distinguish CML from other CMPDs because the clinical course of CML is not indolent; CML patients usually progress to acute leukemia.

Etiology/Pathogenesis

The BCR-ABL hybrid protein, which is produced as a result of the characteristic translocation, is a constitutively active, cytoplasmic tyrosine kinase (3, 4). This upregulated tyrosine kinase activates multiple signal transduction pathways by phosphorylation (3, 4). Cytokines that regulate the growth of normal hematopoietic cells utilize the same signal pathways. As a result, CML cells escape normal cellular control of growth and differentiation (3).

Laboratory studies also have shown decreased adhesion of CML cells to bone marrow stromal cells (5, 6). This observation may account for the peripheral circulation of immature granulocytes and may indicate that the stromal cells have lost control of the growth of hematopoietic cells. The BCR-ABL hybrid protein has been demonstrated to transform hematopoietic cells to become independent of cytokines and growth factors (4, 5, 7). In addition, some protection from apoptosis or programmed cell death may be conferred on the CML cells by the BCR-ABL oncoprotein (5).

The underlying etiology for the development of CML is unknown. No genetic factors have been reported. Exposure to radiation has been implicated with an increased incidence of CML in atomic bomb survivors in Japan and ankylosing spondylitis patients who received spinal radiation (8, 9).

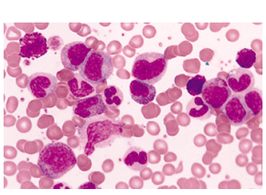


FIGURE 45.1. Characteristic peripheral blood smear of chronic myelogenous leukemia shows basophilia and granulocytosis with neutrophils and immature granulocytes.

Clinical Findings

The reported incidence of CML is about 10 to 20 cases per million people (10). The median age at diagnosis is 50 to 60 years, but CML can occur in much younger patients, including children (8). There may be a slight male predominance. Symptoms are variable and nonspecific, consisting predominantly of fatigue, anorexia, and weight loss. Up to 45% of the patients are asymptomatic at the time of diagnosis with the disease detected as a result of routine blood testing in many cases (4, 8). Some degree of splenomegaly (50% on physical examination) is usually present (4, 8). Hepatomegaly is present in about 10% to 40% of the patients (8). Lymphadenopathy is uncommon in the chronic phase of CML. In contrast to polycythemia vera and essential thrombocythemia hemorrhagic, and thrombotic complications are uncommon in the chronic phase of CML except in patients with marked thrombocytosis. Marked leukocytosis also appears to predispose patients to thromboses (8) and complications of hyperviscosity, including priapism, cerebrovascular accidents, and visual disturbances (8).

Chronic Phase

Peripheral Blood Findings

CML manifests as a primary proliferation of granulocytic elements. CML is included in the differential diagnosis whenever a granulocytosis is encountered in a complete blood count. Granulocytosis is defined as an absolute granulocyte count greater than 7.5×10^9/L. Granulocytes include not only segmented neutrophils, but also bands, metamyelocytes, myelocytes, promyelocytes and myeloblasts. Before considering the possibility of CML, reactive causes of neutrophilia must be excluded (Table 45.1). With the widespread use of colony stimulating factor (CSF) therapy, it is not uncommon to encounter a peripheral leukocytosis with a granulocytic left shift (11). A neutrophilia is expected in certain, “normal” physiologic states and responses such as infections that may be associated with a leukemoid reaction, pregnancy, neonates, stress/exercise, and acute hemorrhage/hemolysis.

TABLE 45.1. DIFFERENTIAL DIAGNOSIS OF REACTIVE NEUTROPHILIA

Infections/leukemoid reaction
Physiologic response (pregnancy, neonates, stress/exercise, acute hemorrhage/hemolysis)
Drugs/hormones (colony stimulating factor therapy, epinephrine, steroids)
Miscellaneous • Gout • Trauma • Autoimmune disorders

Peripheral blood findings in CML are rather characteristic and many times diagnostic. The white blood cell count is increased, but is variable and ranges between 20 and $1{,}000 \times 10^9$/L (median 150 to 200×10^9/L). The two most important findings are basophilia (greater than 0.2×10^9/L basophils) and granulocytosis at various stages of maturation (Fig. 45.1). Peripheral blood smear shows segmented neutrophils, bands, metamyelocytes, myelocytes, and promyelocytes without significant dysplasia. Often occasional blasts are seen, but usually no more than 1% to 2%. The presence of blasts in adults almost always indicates a hematologic malignancy, except in patients receiving CSF. Other features of peripheral blood smears in CSF therapy include very intense granulation in the granulocytes (11), but basophilia is often not present. In addition, occasional blasts in the peripheral blood smears of infants and children are not unusual, especially in tertiary medical settings.

Other peripheral blood findings in the chronic phase of CML include eosinophilia in 90% of the cases and some degree of monocytosis (9, 12). Thrombocytosis of variable degree usually is present, and levels as high as $1{,}000 \times 10^9$/L have been reached in some patients. Thrombocytopenia is extremely rare at diagnosis, and its presence would indicate an accelerated or blastic phase (8). Most patients are anemic at presentation with normal red cell morphology.

Bone Marrow Findings

Morphologic findings in the bone marrow biopsy do not add significant information for diagnostic purposes. However, a bone marrow biopsy is still recommended (13). Cytogenetic analysis for confirmation of the diagnosis of CML and for prognostic purposes is much better on bone marrow aspiration material than peripheral blood. A baseline picture of the bone marrow allows comparison with future changes, assessing initial degree of reticulin fibrosis, and looking specifically for foci of blasts in the core biopsy to indicate focal blastic transformation.

The cellularity is markedly increased, and a marked granulocytic hyperplasia is present with a myeloid to erythroid ratio of 10:1 or greater (Fig. 45.2). Megakaryocytes are increased and typically small and hypolobated (Fig. 45.3), in contrast to other CMPDs. Reticulin stain may show mildly increased reticulin fibrosis. Pseudo-Gaucher cells may be seen in 10% to 20% of the cases, representing the high release of sphingolipids from the leukemic granulocytes (9).

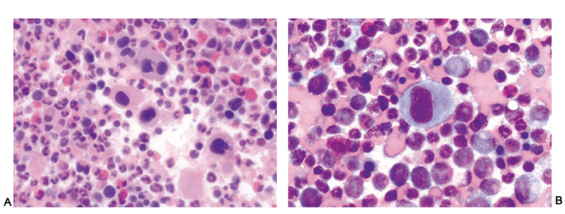


FIGURE 45.3. Typical megakaryocytes in chronic myelogenous leukemia are small and hypolobated. **A:** bone marrow clot section; **B:** bone marrow aspirate smear.

Genetic Studies

Evidence of the BCR-ABL translocation is now a diagnostic criterion. Standard cytogenetic evaluation is useful not only in confirming the diagnosis of CML, but has been useful in assessing

the efficacy of therapeutic regimens. Karyotypic analyses of bone marrow (preferably) or peripheral blood will show the t(9;22) or the Ph chromosome. The Ph chromosome is the abnormal, shortened chromosome 22 with fusion of the BCR gene in the long arm of chromosome 22 and ABL (human homologue of the Abelson murine leukemia virus) oncogene from long arm of chromosome 9 (Fig. 45.4). The breakpoints in the BCR gene of chromosome 22 involve two major areas: the major breakpoint cluster region (M-bcr) between exons 2 and 3 or between exons 3 and 4 and the minor breakpoint cluster region (m-bcr) within the first intron (upstream of M-bcr) (14). The M-bcr is associated with CML while m-bcr is seen in acute lymphoblastic leukemia cases.

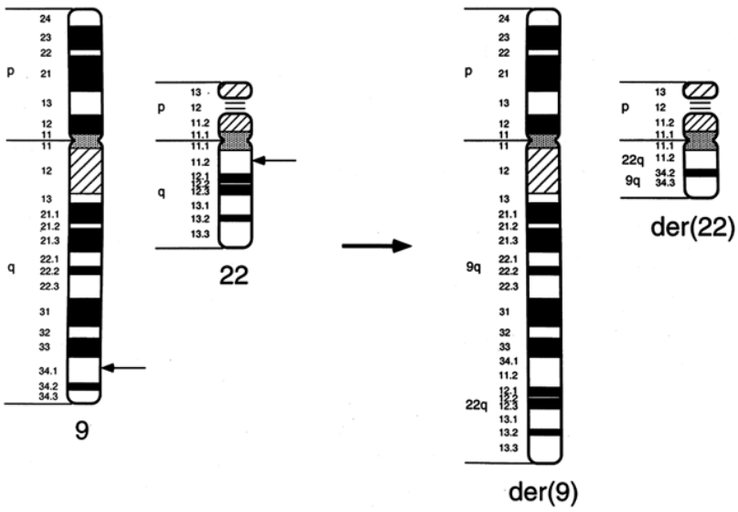


FIGURE 45.4. Schematic of the t(9;22) in chronic myelogenous leukemia. (Courtesy of V. Lindgren, PhD, Department of Pathology, Loyola University Medical Center, Maywood, IL.)

About 5% to 10% of CML cases will be negative for the t(9;22) by standard cytogenetics (15, 16). However, almost half of the Ph-negative CML cases will show evidence of the characteristic translocation by molecular studies. Because the translocation breakpoints in CML occur in a narrow 5.8 kb region of the BCR gene, Southern blot analysis following restriction enzyme digestion can be used for diagnosis (14). However, reverse transcriptase-polymerase chain reaction (RT-PCR) and fluorescence *in situ* hybridization (FISH) methods are now more commonly used in diagnostic laboratories.

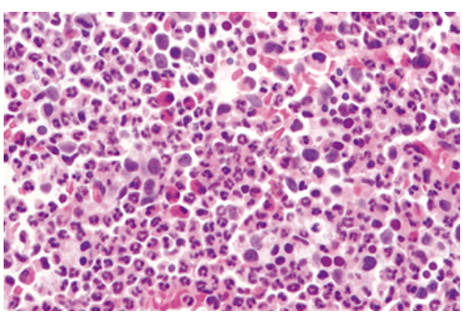


FIGURE 45.2. Bone marrow biopsy of chronic myelogenous leukemia showing marked hypercellularity and marked myeloid hyperplasia.

The breakpoint in the c-ABL gene of chromosome 9 can occur over a large region, more than 100 to 200 kb at the 5′ end upstream of exon 2 (14). As a result, PCR analysis of the BCR-ABL fusion gene is not possible, because PCR can only successfully amplify shorter DNA segments. However, RT-PCR using BCR-ABL mRNA as the template can be performed to detect the translocation. Breakpoint in the M-bcr region yields an 8.5-kb mRNA and 210-kd fusion protein that is typically found in chronic phase CML patients (14). Breakpoint in the m-bcr region produces a 7.0-kb mRNA and a 190-kd fusion protein that is seen in about half of the Ph-positive acute lymphoblastic leukemia (14). Although RT-PCR is more sensitive than standard cytogenetic analyses and southern blots, false positivity as a result of contamination is a significant concern. False negativity in RT-PCR also may occur because of different breakpoint region involvement.

Although RT-PCR analysis has been shown to be extremely sensitive in detecting the BCR-ABL rearrangement, the significance of PCR results in the monitoring of minimal residual disease after bone marrow transplantation or interferon therapy is not clear (17,18 and 19). A positive PCR analysis does not always predict impending relapse, and a negative result does not mean cure. In fact, there have been two studies demonstrating the detection of BCR-ABL transcripts by RT-PCR in normal, healthy individuals (20, 21). Therefore, using PCR results for CML patients in clinical therapeutic decision making are not recommended until more prospective studies are performed (22).

Cytogeneticists have been using FISH more frequently to detect the Ph chromosome. Directly conjugated fluorescent probes against the BCR and c-ABL genes are commercially available. Assuming the ABL probe is attached to a red/rhodamine fluorochrome, and the BCR probe has a green/FITC fluorochrome, normal cells will show two red and two green signals. In contrast, cells with the translocation will show one red signal, one green signal, and one yellow signal due to the fusion of the red and green signals, indicative of the BCR-ABL translocation (Fig. 45.5). Studies have shown the efficacy of FISH in the detection of the translocation at the time of diagnosis (23) and that FISH is a viable alternative to karyotyping and Southern blot analysis (24). FISH is fast and simple; results can be available in 24 hours. If

fresh material is not available for karyotyping or molecular studies, old peripheral blood or bone marrow smears may be used. Another advantage of FISH over standard cytogenetics is the capability to evaluate nuclei in interphase; therefore, the FISH method is not restricted by the quality and quantity of metaphases. However, there are limitations to FISH. No interlaboratory reproducibility and consistency have been established as specific regulations and guidelines are not available. The cutoff point for determining positive or negative results has to be determined by each individual laboratory. Therefore, the use of FISH in monitoring minimal residual disease would be difficult (23).

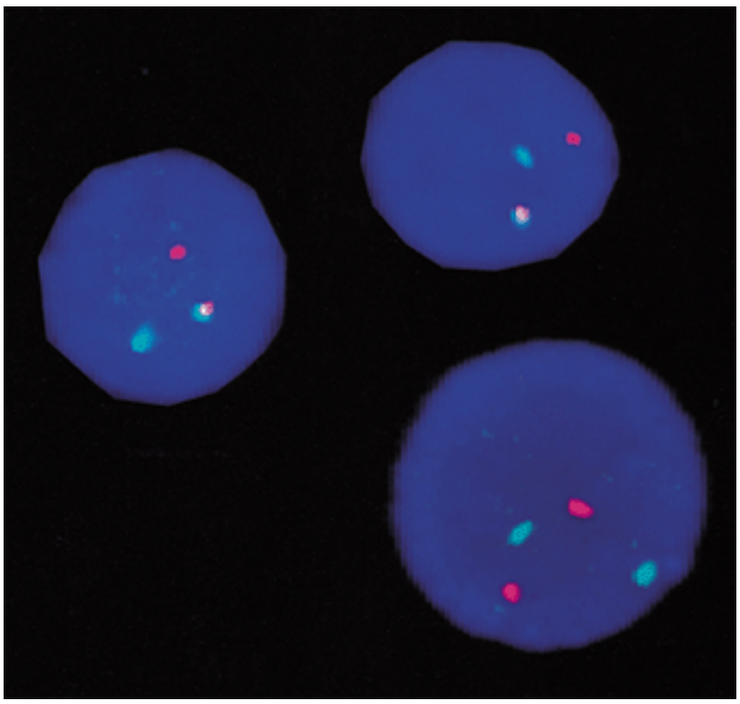


FIGURE 45.5. FISH (fluorescence *in situ* hybridization) analysis of the t(9;22) in chronic myelogenous leukemia. The normal cell shows two red signals of ABL gene and two green signals of the BCR gene. The two abnormal cells show one red signal, one green signal, and one yellow signal of the translocation due to the fusion of the red and green signals. (Courtesy of V. Lindgren, PhD, Department of Pathology, Loyola University Medical Center, Maywood, IL.)

There are other disorders with the t(9;22) besides CML (13). A small percentage of acute lymphoblastic leukemias also has a t(9;22), about 5% of pediatric and 20% to 25% of adult acute lymphoblastic leukemia. About half of acute lymphoblastic leukemia with the t(9;22) involve identical breakpoints in the BCR gene as CML (M-bcr), resulting in the 210-kD BCR-ABL protein. About half involve the breakpoint in the m-bcr (first intron) that produces the 190-kD fusion protein (14). The presence of t(9;22) in acute lymphoblastic leukemia is a poor prognostic indicator; patients in this particular group usually are older, present with a high white blood cell count, and are resistant to chemotherapy (25, 26). The Ph chromosome is also detected in a few cases of acute myelogenous leukemias (<1%). It has been reported in acute basophilic leukemia (27) and acute biphenotypic or mixed lineage leukemia (28).

Flow Cytometry

Flow cytometric analysis is not indicated for the chronic phase of CML. However, if blastic transformation occurs, immunophenotyping is critical in the characterization of the blasts to determine myeloid or lymphoid lineage.

Leukocyte Alkaline Phosphatase (LAP) Score

An additional study that may be useful in the workup of a possible CML case is the LAP score (29). The LAP score test is performed on fresh, heparinized peripheral blood. A cytochemical stain detecting the level of enzyme activity is analyzed by examining 100 neutrophils or bands for individual intensity of cytoplasmic reaction on a scale of 0 to 4. The individual scores of 100 cells are summed to obtain the total score. A wide normal range of the total score (30 to 150) has been reported. Because of the variability of observers and dyes among different laboratories, each laboratory should establish its own normal range. A low LAP score less than the established normal range is strong supportive evidence of CML. Paroxysmal nocturnal hemoglobinuria is an example of other rare conditions that may cause a low LAP score. However, a normal or high LAP score does not exclude the possibility of CML because superimposed infections/leukemoid reactions, pregnancy, drugs, and other inflammatory disorders can elevate the LAP score. Because of the analytical uncertainty of the LAP score, interpretation should be made with great caution (30).

Differential Diagnosis of CML

Leukemoid Reaction

The marked, polyclonal increase in neutrophils and bands secondary to inflammation, infection, cancer, and drugs is known as leukemoid reaction (31). The distinction between CML and leukemoid reaction is generally straightforward. While the total leukocyte count in leukemoid reaction is increased to more than 50×10^9/L, peripheral blood smear shows mostly mature neutrophils and bands with absence of basophilia. Furthermore, leukemoid reaction is associated with a normal cytogenetic analysis and high, or less commonly normal, LAP scores.

Chronic Myelomonocytic Leukemia and Atypical Chronic Myeloid Leukemia

As stated above, about 5% of CML lack cytogenetic (Ph-negative) and molecular evidence of the BCR-ABL fusion. The differential diagnosis for most of the remaining 5% of the cases includes chronic myelomonocytic leukemia (CMML) and atypical chronic myeloid leukemia (aCML) (32). Diagnostic criteria for CMML and aCML are not straightforward, but guidelines have been proposed (32,33 and 34). Most of the differentiating features are based on peripheral blood findings with emphasis on the degree of monocytosis, basophilia, immature granulocytosis, and dysplasia. The presence of peripheral basophilia has been found to be the best morphologic discriminator for Ph-positive CML (35). The presence of granulocytic dysplasia and monocytosis raises the possibility of CMML and aCML (34). The bone marrow findings in CMML and aCML are similar with hypercellular marrows except for the percentage of erythroid precursors (34).

CMML is a disease primarily seen in the elderly. Patients typically present with anemia and thrombocytopenia as well as splenomegaly and leukocytosis resulting from neutrophilia and monocytosis. The monocytosis is generally greater than 10% (33) and exceeds 1×10^9/L. However, basophilia usually is absent (less than 2% and less than 0.2×10^9/L), and immature granulocytes (metamyelocytes, myelocytes, and promyelocytes) are less than or equal to 10% (33). Granulocytic dysplasia is mild to absent. In summary, the peripheral blood smear of CMML shows basically a mixture of increased numbers of mature neutrophils and monocytes with only occasional myelocyte or metamyelocyte (32). The bone marrow contains more erythroid precursors than aCML or Ph-positive CML (usually greater than or equal to 15%) (34, 36).

CMML has been a difficult disorder to classify with some cases better classified as a myelodysplastic disorder and others better classified as a myeloproliferative disorder. The distinction between the two phases had been based primarily on the presenting leukocyte count and the presence or absence of splenomegaly. A recent study divided CMML into a myelodysplastic type and a myeloproliferative type based on the leukocyte count, using 13×10^9/L as the cutoff. The study found no significant difference in the 2-year survival or overall prognosis between the two types of CMML (37).

The new WHO classification of acute leukemia and myelodysplastic syndromes removes the CMML category from the myelodysplastic disorders and places it in a new “myelodysplastic/myeloproliferative disorders” category, recognizing that there is a group of disorders that have both myelodysplastic and myeloproliferative features. This mixed category also includes aCML and juvenile myelomonocytic leukemia (formerly known as juvenile chronic myeloid leukemia) (1).

aCML, similar to CML, shows a neutrophilia with shift to immaturity (immature granulocytes between 10% and 20%), but lacks basophilia and shows minimal to mild monocytosis (between 3% and 10%). Another distinguishing feature of aCML is the presence of trilineage dysplasia with hypogranular and hypo/hypersegmented granulocytes, dyserythropoiesis, and dysmegakaryopoiesis (32). Erythroid precursors in the bone marrow are rare.

Recognition of Ph-negative and BCR-ABL negative CML cases has prognostic significance. aCML appears to be clinically aggressive with a reported median survival of less than 2 years (34). Prognosis for CMML is variable because of the heterogeneity in cases reported as CMML. In general, CMML is thought to behave more poorly than the BCR-ABL positive CML (chronic granulocytic leukemia) (34).

Juvenile Chronic Myelogenous Leukemia

Chronic myeloproliferative disorders in children are uncommon and overlap with myelodysplastic syndromes of childhood, as recognized by the new WHO classification (1). Juvenile chronic myelogenous leukemia (JCML) represents the most common form of myeloproliferative/myelodysplastic disorders in children

under the age of four. JCML has also been described under different names such as chronic myelomonocytic leukemia in children, monosomy 7 syndrome, and juvenile myelomonocytic leukemia (JMML). The latter designation of JMML was proposed by the recently established International Juvenile Myelomonocytic Leukemia Working Group (38, 39).

This group of disorders is distinctly different from the Ph-positive adult CML. JCML or JMML lacks the Ph chromosome and molecular evidence of the BCR-ABL rearrangement. Additional criteria for this diagnosis, as proposed by the International Working Group (38), include absolute peripheral monocytosis of greater than 1×10^9/L [usually > 5×10^9/L (39)], and bone marrow blasts less than 20%. Two additional criteria from the following list must also be met: elevated hemoglobin F for age [greater than 10%, but seen in only one half to two thirds of the patients (39, 40)], granulocytic left shift in the peripheral blood, peripheral leukocytosis over 10×10^9/L, cytogenetic abnormality [about two thirds with normal karyotype, but monosomy 7 in one quarter of the patients analyzed (41)], or *in vitro* cell culture studies demonstrating hypersensitivity of myeloid progenitor cells to granulocyte/macrophage colony stimulating factor (42).

Most patients present before the age of four, and median age at presentation in one series was 1.8 years (41). Characteristic clinical features of this disorder consist of hepatosplenomegaly, lymphadenopathy, skin rashes, and bleeding in the form of purpura, ecchymoses, petechiae, and epistaxis secondary to thrombocytopenia (43, 44). Skin manifestations include erythematous maculopapular rashes of the face and lesions associated with neurofibromatosis (e.g., xanthomas, café-au-lait spots and neurofibromas) (43).

About 10% to 25% of the cases are associated with neurofibromatosis, type 1 (44). In fact, the risk for JMML in children with neurofibromatosis is increased by more than 300-fold (41, 45). The association with neurofibromatosis has shed some light on the pathogenesis of JMML/JCML (39). Loss of the normal NF1 allele results in decreased downregulation of Ras-GTP. The increased Ras-GTP level causes constitutive activation of its signal transduction pathway. In addition, 30% of the cases have mutations in the RAS gene that also result in increased levels of Ras bound to GTP.

The prognosis, in general, is poor, and some patients do develop blastic transformation (40, 43). Chemotherapy is of very limited usefulness, and bone marrow transplantation (BMT) provides the only chance of prolonging survival (39, 41). Poor prognostic factors include age at presentation over 2 years, elevated hemoglobin F greater than 10%, and low platelet count (40, 41). High blast and normoblast counts in peripheral blood were also found to be predictive of short survival (43).

Chronic Neutrophilic Leukemia

This disorder is very rare and remains a controversial entity. The pathologic findings closely mimic leukemoid reaction (9, 46), characterized by a persistent, unexplained, predominantly mature neutrophilia without basophilia. There is a slight left shift in the granulocytes, including occasional myeloblasts. Toxic granulation and Dohle bodies are often seen in the neutrophils. Despite the reactive changes seen in the neutrophils, these patients do not have any evidence of an underlying infection. Instead, these patients are usually elderly with prolonged splenomegaly accompanying the neutrophilia. Cytogenetics is helpful, as these patients do not have the Ph chromosome, but occasionally may show other clonal abnormalities (47). Blastic transformation has been reported (48).

Clinical Course and Progression of CML

The chronic phase of CML usually lasts 3 years and often progresses to an accelerated phase and ultimately, blastic transformation and death. About 20% to 25% of the patients progress directly to the blastic phase without evidence of the accelerated phase (8). Median survival after patient reaches the accelerated phase is 6 to 18 months, and survival is less than 6 months with blastic transformation. Common causes of death are leukemic transformation and bone marrow failure/fibrosis.

Accelerated phase

The diagnostic criteria for the accelerated phase are vague and include both morphologic and clinical findings (49). Clinically, the patient may be refractory to therapy with increasing splenomegaly or may require higher doses of drugs to control the disease. In the peripheral blood, the white cell count may increase or decrease with increasing blasts, but blasts constitute less than 30%. Peripheral blasts greater than or equal to 15% (but less than 30%) or peripheral blasts and promyelocytes greater than or equal to 30% (but blasts less than 30%) have been used as criteria for the accelerated phase (49). Increasing basophilia greater than or equal to 20% and morphologic evidence of dysplasia are indicative of disease progression (9). Hemoglobin levels and platelet counts may drop with platelet counts less than 100×10^9/L found to be a poor, independent prognostic factor (49). In the bone marrow, increasing numbers of blasts (but less than 30%) and increasing basophilia may be seen. A reticulin stain to evaluate for increasing myelofibrosis is useful. Clonal cytogenetic evolution with additional cytogenetic abnormalities is indicative of progression of CML. Common additional abnormalities include a second Ph chromosome, trisomy 8 (+8), and isochromosome 17q.

Blastic Transformation

A diagnosis of blastic transformation of CML can be made if any of the following three criteria are met (50). First, there are 30% or more blasts in the peripheral blood or bone marrow (Fig. 45.6). The development of an extramedullary myeloid tumor or lymphoblastic lymphoma also indicates a blast crisis. The extramedullary disease may be the first and only evidence of blastic transformation with no increased blasts in the bone marrow and peripheral blood (9, 51). Finally, blastic transformation has occurred if blasts in the peripheral blood and bone marrow are less than 30%, but the bone marrow core biopsy shows large intramedullary foci of blasts. This latter finding is usually referred to as focal blast transformation.

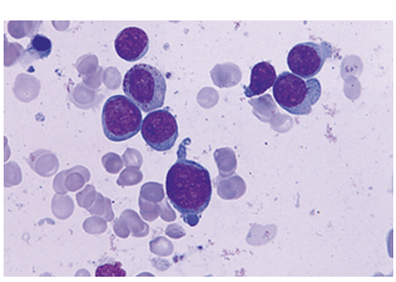


FIGURE 45.6. Numerous blasts in the bone marrow aspirate smear from a patient with chronic myelogenous leukemia in blastic transformation.

Analysis of 242 patients with CML in blast crisis was performed

by Kantarjian et al (52). Three independent, poor prognostic factors were identified: myeloid or undifferentiated blasts, evidence of clonal evolution, and thrombocytopenia less than 50 × 10^9/L (indicative of low marrow reserve). These factors affected overall survival and response to therapy. The morphologic findings of myeloid, lymphoid, or undifferentiated blast crises are similar to their respective counterparts in *de novo* acute leukemia. Appropriate workup includes immunophenotypic characterization of blasts by flow cytometry and fresh aspirate material sent for cytogenetic evaluation because many patients [up to 60% in one study (52)] demonstrate clonal evolution. Clonal evolution most frequently involves the development of a second Ph chromosome and trisomy 8 (Fig. 45.7). The presence of isochromosome 17q is also considered as evidence of clonal evolution (52, 53).

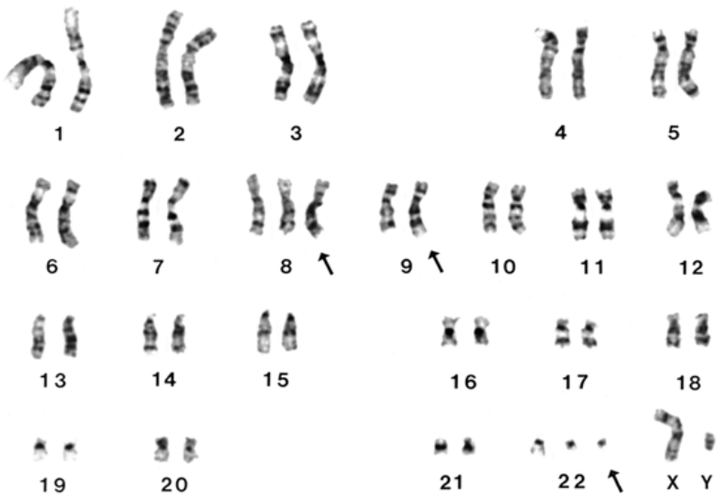


FIGURE 45.7. Karyotype of a man with chronic myelogenous leukemia in blastic transformation demonstrating several chromosomal abnormalities. The abnormalities include Philadelphia chromosome (abnormally shortened chromosome 22), a second Philadelphia chromosome, and trisomy 8. (Courtesy of V. Lindgren, PhD, Department of Pathology, Loyola University Medical Center, Maywood, IL.)

About two thirds of the blast crises are myeloid in origin and one third is lymphoid in origin. It is clinically and prognostically important to determine lineage of these blast crises. Although blast crises of CML in general are clinically very aggressive with survival usually less than 6 months, several studies have found that the lymphoid blast crisis in CML is associated with longer median survival and better response to therapy when compared with the myeloid blast crisis (54,55 and 56). The lymphoid blast crisis appears to respond to acute lymphoblastic leukemia therapeutic regimens containing vincristine and prednisone. Although the prognosis is better with the lymphoid blast crisis, the overall survival advantage is of short duration with median survival of 9 to 12 months for a lymphoid blast crisis versus 3 to 4 months for a nonlymphoid blast crisis. As a result, allogeneic BMT is recommended for patients with a lymphoid blast crisis after induction of a second chronic phase (54).

Other characteristics associated with the lymphoid blast crisis in CML include abrupt transformation without a preceding accelerated phase, less frequent organomegaly, lower white blood cell count with lower blast percentage, fewer peripheral basophils, lesser degree of anemia, more extensive involvement of bone marrow by blasts, and more frequent extramedullary involvement (54,55 and 56). The overwhelming majority of the lymphoid blast crisis show a B-lineage with only rare cases of T-cell blast crisis (57). Aberrant expression of myeloid antigens in these lymphoid blast crises appears to be relatively common (55, 56).

Extramedullary blast transformation can occur anywhere, but common locations include lymph nodes, skin, spleen, and soft tissues. A leukemic tumor in extramedullary sites usually consists of a proliferation of relatively uniform, medium to large mononuclear cells (Fig. 45.8). Differentiation from large-cell lymphoma and other large, poorly differentiated neoplasms is difficult based on morphology alone. Sometimes an eosinophilic infiltrate may accompany the tumor cells and provide a clue to the leukemic nature of the neoplasm.

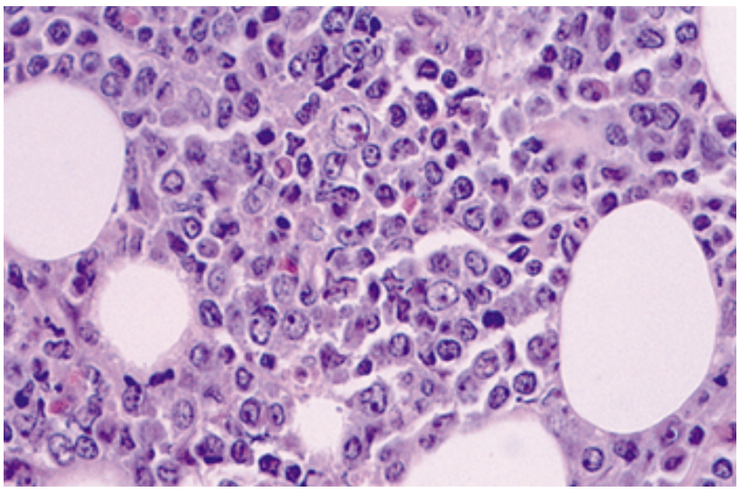


FIGURE 45.8. Extramedullary myeloid tumor (granulocytic sarcoma) in retroperitoneal tissue from a patient with chronic myelogenous leukemia. The tumor cells are relatively uniform, medium to large, mononuclear with high nuclear/cytoplasmic ratio.

Immunohistochemical analysis should be performed (58). For an extramedullary myeloid tumor (also known as granulocytic sarcoma), myeloperoxidase and lysozyme are good myeloid markers (Fig. 45.9). For lymphoblastic lymphoma, TdT is probably the most useful in confirming the diagnosis. For purposes of determining lineage, CD79a and CD10-immunopositivity indicate B-cell lineage. CD20 is not as useful, given that most

precursor B-acute lymphoblastic leukemias are negative for CD20. For T-cell lineage, antibody against CD3 can be used to detect both surface and cytoplasmic CD3 expression. Although CD43 is a nonspecific T-cell marker, it also marks myeloid cells, and many times it may be the only marker initially positive if an extramedullary myeloid tumor was not considered in the differential diagnosis. CD34 is a marker of progenitor, hematopoietic cells, and its expression is consistent with acute leukemia, but it does not indicate myeloid or lymphoid origin.

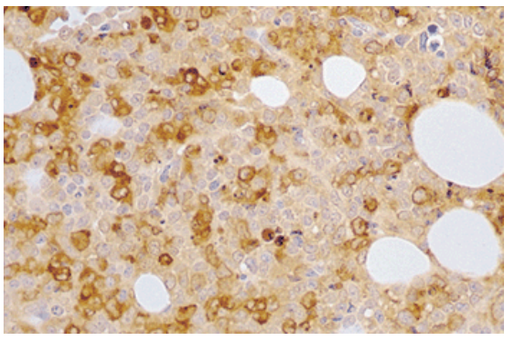


FIGURE 45.9. Myeloperoxidase positivity in extramedullary myeloid tumor (granulocytic sarcoma). This tumor is the same case presented in Fig. 8.

Therapy

Most patients with favorable clinical findings are treated with hydroxyurea, busulfan, or recombinant interferon-α. Allogeneic BMT is the preferred first-line therapy if the patient is young (preferably less than 50 years of age) and if an HLA-matched related donor is available. Transplantation in the early stages is advocated because better survival has been seen in patients transplanted in the chronic phase than in the accelerated or blastic phase. In addition, patients have better chances of survival if they are transplanted within the first year after diagnosis rather than later in the chronic phase (4). However, the benefits of BMT should be weighed against potential morbidity and mortality associated with transplantation (59).

For patients in whom BMT is not possible, interferon-α alone or in combination with other agents is the preferred optimal therapy (59). Interferon-α has been shown to induce hematologic and cytogenetic remissions that are not long-lasting. One of the important assays for clinical response to interferon-α is FISH detection of the Ph chromosome. Other experimental therapies include anti-BCR-ABL antisense oligonucleotides, BCR-ABL–specific tyrosine kinase inhibitors, and intensive chemotherapy with purged auto stem cell transplantation.

The role of autologous transplantation in CML as primary or salvage therapy has not been clearly defined and is still experimental. Investigators have attempted to improve methods for purging the graft by *ex vivo* and *in vivo* means (60). *Ex vivo* purging includes longterm liquid cultures, antisense oligonucleotides, tyrosine kinase inhibitors, and selection of CD34-positive, HLA-DR negative progenitor cells. *In vivo* purging primarily involves decreasing the tumor burden by intensive chemotherapy prior to harvesting of the bone marrow or peripheral blood.

POLYCYTHEMIA VERA

Part of "45 - Chronic Myeloproliferative Disorders"

PV is distinguishable from the other chronic myeloproliferative disorders by its predominant proliferation of the erythroid cell line. However, it is actually a stem cell disorder that causes proliferation of all the myeloid components. The increased red cell mass is the hallmark of this disorder, but the white blood cell count is often mildly increased, and quite significant thrombocytosis is generally present. In addition, the bone marrow often shows panhyperplasia, not just erythroid hyperplasia.

Etiology/Pathogenesis

PV is a clonal disorder, as confirmed by cytogenetic and clonality analyses. About 17% of untreated patients with PV have karyotypic abnormalities (61,62,63,64 and 65). The most common abnormalities found at the time of diagnosis include deletion of the long arm of chromosomes 20 and 13 (20q- and 13q-), trisomy 8 (+8), and trisomy 9 (+9). With progression of disease to myelofibrosis or acute leukemia, the frequency of chromosomal anomalies increases to 50% to 80% (65). Overall 65% of patients who were treated with myelosuppressive agents showed cytogenetic abnormalities, most commonly involving chromosomes 5 and 7 (61,62,63,64,65 and 66).

Clonality assays looking at glucose-6-phosphate dehydrogenase isoenzymes and allelic patterns of X-linked genes have demonstrated monoclonality with a monoallelic pattern (67,68 and 69). A skewed pattern of X-inactivation in granulocytes together with a balanced pattern in lymphocytes generally indicates clonality. However, similar findings have been found in 25% to 50% of normal, elderly women (70). This increased skewing with age seems to indicate that analysis of X-chromosome inactivation patterns cannot be used as a diagnostic test for clonality of chronic myeloproliferative disorders in elderly women (71).

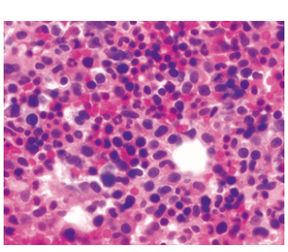


FIGURE 45.10. Bone marrow clot section of polycythemia vera in the proliferative stage showing hypercellular marrow with marked erythroid hyperplasia.

Another interesting observation is spontaneous *in vitro* erythroid colony formation without the addition of exogenous

erythropoietin (72). Supporting this observation is the generally low levels of serum erythropoietin in PV patients (73). Subsequent studies have actually found that the erythroid precursors in PV are hypersensitive to several different growth factors (e.g., stem-cell factor, interleukin-3, and granulocyte macrophage-colony stimulating factor), suggesting possible defects in the signal transduction pathway (65).

Proliferative Stage

Clinical Findings

About 5 to 10 cases of PV occur per million people per year, about half the incidence of CML. PV most commonly occurs in older people in their seventh decade. No sexual predilection is seen. About 20% of the patients with PV present with an acute thrombotic episode, and an additional 30% develop thromboses later on (74). Both arterial and venous thromboses occur with equal frequency. Arterial thrombi can involve cerebral, coronary, and peripheral blood vessels. The risk of thromboses increases with higher hematocrit levels from the increased red cell mass and the associated increased whole blood viscosity (74, 75). Other risk factors for thrombotic complications include older age over 70 years, previous history of thrombi, and history of frequent phlebotomy (most likely related to increased severity of the disease that requires more phlebotomies to maintain stable hematocrit levels) (76).

PV patients also experience microvascular, occlusive symptoms in the form of erythromelalgia and transient neurologic and ocular defects (75). Erythromelalgia presents as burning, painful, swollen toes and fingers. The neurologic and visual disturbances include attacks of aphasia, unstable gait, temporary interruption of ongoing activities, transient blindness, flashing lights, and blurred vision. These microvascular complications are attributed to increased platelet counts and are common in essential thrombocythemia and PV associated with thrombocytosis (76).

In additon, intractable pruritus is a frequent and bothersome symptom in PV (75, 77). About two thirds of the patients have a plethoric complexion. Splenomegaly is present in 70% of the patients, and a palpable liver is appreciated in 40% (77).

Peripheral Blood Findings

The key abnormalities in the peripheral blood are an elevated red cell count (usually greater than 6×10^{12}/L), increased hemoglobin level, and elevated hematocrit. No significant poikilocytosis usually is seen in the peripheral blood smear. The hemoglobin, not uncommonly, may be within normal range because oftentimes the patient is also iron deficient with decreased mean corpuscular volume. The red cell parameters in the complete blood count may indicate PV, but increased total red cell mass must be documented by radioisotopic studies for definitive diagnosis (see “Diagnostic Criteria and Other Studies,” below). Before considering PV as the underlying cause for the erythrocytosis, relative and secondary erythrocytosis need to be excluded. Relative erythrocytosis refers to spurious erythrocytosis as a result of plasma volume contraction from vomiting, dehydration or diarrhea. Secondary erythrocytosis is discussed below in “Diagnostic Criteria and Other Studies.”

The white cell counts generally range from 10 to 20 × 10^9/L with a slight shift to immaturity. Modest absolute basophilia is seen in about 70% of the patients (46). The platelet counts usually are increased, ranging from 400 to 800 × 10^9/L with counts greater than 1,000 × 10^9/L not uncommon (77). In fact, thromobocytosis and leukocytosis in the absence of fever and infection are cited as B criteria in the initial diagnostic criteria for PV (77).

Bone Marrow Findings

The morphologic findings in the bone marrow core biopsy are not specific or diagnostic for PV. The marrow is hypercellular for age with panmyelosis or trilineal hyperplasia, particularly of the erythroid (Fig. 45.10) and megakaryocytic series. Erythroid precursors do not show evidence of dysplasia. Megakaryocytes may cluster and may be large and hyperlobulated, resembling the megakaryocytes in essential thrombocythemia. The granulocytic elements are unremarkable with no increase in blasts. An iron stain is important, as iron stores are almost always absent. A reticulin stain shows normal reticulin content in most patients, but up to one third may have some degree of increased reticulin fibrosis. This finding at the proliferative stage does not appear to have prognostic significance (78).

Diagnostic Criteria and Other Studies

PV is a clinicopathologic diagnosis. In the 1970s, the Polycythemia Vera Study Group (PVSG) proposed a set of criteria for the diagnosis of PV (77). Table 45.2 lists the PVSG criteria. A diagnosis of PV was considered acceptable if the red cell mass exceeded 36 mL/kg in men or 32 mL/kg in women, arterial oxygen saturation was normal, and splenomegaly was present. If splenomegaly was not present, then two additional B criteria must be met.

TABLE 45.2. PVSG CRITERIA FOR POLYCYTHEMIA VERA

A Criteria (Major)		B Criteria (Minor)	
A1	Increased red cell mass Male ≥ 36 mL/kg Female ≥ 32 mL/kg	B1	Thrombocytosis (> 400 × 10^9/L)
A2	Normal arterial oxygen saturation ≥ 92%	B2	Leukocytosis (> 12 × 10^9/L) in absence of fever or infection
A3	Splenomegaly	B3	Increased neutrophil alkaline phosphatase score in absence of fever or infection
		B4	Increased B_{12} (≥ 900 mg/L) or increased unsaturated B_{12} binding capacity (>2200 mg/L)

Diagnosis of PV if A1 + A2 + A3 or A1 + A2 + 2B criteria

(From Berlin NI. Diagnosis and classification of the polycythemias. *Semin Hematol* 1975;12:339–351, reprinted with permission.)

With the development of new techniques and new scientific findings in PV, revised diagnostic criteria have been proposed (79) (Table 45.3). Total red-cell mass determined by radioisotopic studies involves *in vitro* labeling of patient's erythrocytes

with 51-chromium, reinjecting the labeled red cells into the patient, and measuring the radioactivity after adequate equilibration. It is felt that the measurement of red-cell mass in milliliters per kilogram was insufficient, but that the red cell mass should be compared to the normal predicted value (80).

TABLE 45.3. PROPOSED REVISED DIAGNOSTIC CRITERIA FOR POLYCYTHEMIA VERA (PV)

A Criteria		B Criteria	
A1	Increased red-cell mass (> 25% above mean normal predicted value)	B1	Thrombocytosis (> 400 × 10^9/L)
		B2	Neutrophilia (> 10 × 10^9/L)
		B3	Splenomegaly by isotope/ultrasound scan
A2	Secondary polycythemia excluded	B4	Spontaneous BFU-E growth or decreased serum erythropoietin
A3	Palpable splenomegaly		
A4	Evidence of clonality		

Diagnosis of PV if A1 + A2 + (A3 or A4) or A1 + A2 + 2B criteria

(From Pearson TC, Messinezy M. The diagnostic criteria of polycythaemia rubra vera. *Leuk Lymphoma* 1996;22:87–93, reprinted with permission.)

The initial PVSG criterion of normal arterial oxygen saturation was replaced by “absence of cause of secondary polycythemia” to be more encompassing and to include causes of inappropriate erythropoietin production by certain neoplasms. The differential diagnosis for secondary polycythemia is a long list (Table 45.4) (77). A major category is increased erythropoietin as a physiologic response to hypoxia of various causes, including living at high altitudes, smokers, congenital heart disease, chronic obstructive pulmonary disease, and high oxygen affinity hemoglobinopathy. In the latter entity, the abnormal hemoglobin binds tightly to the oxygen molecules, and oxygen delivery to tissue sites is decreased. The second major category is inappropriate erythropoietin production, as seen in some renal-cell carcinomas, hepatocellular carcinomas, and cerebellar hemangioblastomas.

TABLE 45.4. DIFFERENTIAL DIAGNOSIS OF SECONDARY POLYCYTHEMIA

Response to Hypoxia
High altitude
Congenital heart disease
Chronic obstructive pulmonary disease
Smoking
High oxygen affinity hemoglobinopathy
Inappropriate erythropoietin production
Tumors (renal-cell carcinoma, hepatocellular carcinoma, hemangioblastoma)
Benign renal lesions (cysts, hydronephrosis)

An additional primary criterion of cytogenetic abnormality was included, as discussed in “Pathogenesis.” The importance of splenomegaly was again emphasized with palpable splenomegaly a major criterion and splenomegaly by radiologic studies a minor criterion. A new minor criterion was added: the spontaneous growth of erythroid colonies without the addition of erythropoietin (81) or decreased serum erythropoietin level, in contrast to increased erythropoietin level in secondary erythrocytosis.

Progression of PV

In 7 to 8 years after diagnosis (78), about 10% of PV cases will progress to the stage known as postpolycythemic myeloid metaplasia or the “spent phase” (82). From the pathologic standpoint, the morphologic findings are essentially identical to chronic idiopathic myelofibrosis with marked fibrosis in the bone marrow and presence of teardrop red cells in the peripheral blood. Histologic examination of spleens at this stage shows extramedullary hematopoiesis as the cause for the increasing splenomegaly, whereas splenomegaly in the proliferative phase is secondary to marked congestion (83).

Progression to acute leukemia is considered part of the natural history with an incidence of 1% to 2% in untreated patients who received phlebotomy only (82). Patients who were treated with alkylating agents or radioactive phosphorus had an increased incidence of 10% to 15%. The leukemogenic effect of hydroxyurea is still an issue that needs to be resolved (84). In addition, an antecedent postpolycythemic myeloid metaplasia also causes an increased risk for the development of acute leukemia (82). The onset of blastic transformation is usually abrupt without a previous myelodysplastic stage. The blastic progression is mostly myeloid in origin with occasional cases of acute lymphoblastic leukemia (82). Acute leukemia generally shows a poor clinical response to chemotherapy.

Therapy/Prognosis

PV is a progressive disease, and patients need to be treated, in contrast to essential thrombocythemia. If untreated, patients have a 50% chance of dying within 18 months after the onset of first symptom (85). To control the vascular complications in PV, therapy is aimed at maintaining hematocrit levels below 45% and platelet counts below 400 × 10^9/L (76, 86). Phlebotomy alone does not prevent the microcirculatory complications of erythromelalgia and the transient ischemia of cerebral and ocular blood vessels because the thrombocytosis persists.

In the 1970s, treatment primarily consisted of radiophosphorus and alkylating agents such as busulfan, melphalan, and chlorambucil. The increased risk of developing therapy-related malignancies from alkylating agents was recognized, and hydroxyurea became first-line therapy in 1980. Hydroxyurea has been effective in controlling thrombocytosis and elevated hematocrit; thereby, decreasing the risk of thrombosis (84). However, it causes side effects, including myelosuppression and possibly

acute leukemia. Long-term safety of hydroxyurea is still not known.

Newer treatment modalities have evolved. Low-dose aspirin is currently recommended for the treatment of erythromelalgia and other microvascular circulatory problems (76, 87). Possible contraindications to the use of aspirin are history of gastrointestinal bleeding or peptic ulcer disease and extremely high platelet counts. Interferon-α has been tried in a limited number of patients and has been found to be effective in controlling the disease by reducing the degree of splenomegaly, platelet counts, and requirements for phlebotomy (88). In particular, this drug has been especially effective in the treatment of PV-associated pruritus, but side effects remain a limiting factor in up to one-third of the patients treated (88). Allogeneic bone marrow transplantation for PV patients who progress to myelofibrosis has only been attempted in very few patients (89).

Median survival is about 10 years, and a common cause of death in treated and untreated patients is thrombosis in the coronary or cerebral vessels (74). Other causes of death include acute leukemia and marrow failure from fibrosis.

ESSENTIAL THROMBOCYTHEMIA

Part of "45 - Chronic Myeloproliferative Disorders"

ET is characterized by marked thrombocytosis in the peripheral blood; however, thrombocytosis is often present in other CMPDs. ET is primarily a diagnosis of exclusion because it lacks a diagnostic genetic marker. In addition, its clinical, laboratory, and pathologic features significantly overlap with other CMPDs.

Etiology/Pathogenesis

ET is generally thought to be a clonal stem cell disorder. However, only about 5% of the cases analyzed by cytogenetics show clonal abnormalities, and no consistent karyotypic anomalies have been found (66, 90). Other ways to determine clonality have been attempted, and Fialkow in 1981 demonstrated clonality by glucose-6-phosphate dehydrogenase isoenzyme analysis in three ET patients (91).

Studies based on the human androgen receptor gene (HUMARA) showed conflicting results. While El-Kassar et al. demonstrated clonality in 27 of 48 patients (92), Harrison et al. (93) demonstrated clonality in only 10 of 46 female patients. Evaluation of clonality may be useful in determining which patients are at a higher risk for progression to acute leukemia or at greater risk for thrombotic or hemorrhagic complications (94).

A few cases of familial ET have been reported with generally an autosomal-dominant mode of inheritance (95,96 and 97). A molecular abnormality in the thrombopoietin gene was discovered in one family (97). A one-base deletion in the 5′ end of the thrombopoietin gene resulted in increased levels of thrombopoietin in affected family members. However, sporadic cases of ET do not display increased thrombopoietin levels (98), but one study found decreased expression of the thrombopoietin receptor (c-Mpl) in ET platelets (99). Therefore, it has been proposed that the underlying molecular mechanisms for familial and sporadic cases of ET are most likely different (97).

Clinical Findings

ET is less common than CML or PV with less than five new cases per million people per year. In fact, one study found ET in only 4% of 280 patients with a platelet count greater than 1,000 × 10^9/L (100). The age at diagnosis is variable with a median age of about 60 years and a range of 21 to 85 years of age in the group of 91 patients with ET followed by the PVSG (101). About 10% to 25% of the patients present under the age of 40 (74). No sexual predilection is noted.

Presenting clinical symptoms are variable, ranging from asymptomatic to unexplained thrombotic/hemorrhagic episodes, as seen in PV. Up to one third of the patients are discovered incidentally (74). Twenty-five percent to 75% of ET patients present with a thrombotic event, most commonly affecting the microcirculation in the form of erythromelalgia, characterized by painful, swollen, burning toes and fingers (75, 102). Erythromelalgia is caused by platelet microthrombi in the arterial vasculature of the extremities (102). Venous thrombosis is less common than arterial thrombosis in ET. Arterial thrombosis mainly involve the cerebral vasculature, peripheral vessels, and coronary arteries (103). Many ET patients experience neurologic symptoms such as headaches, dizziness, or visual problems as a result of cerebrovascular ischemia (103). Bleeding events at the time of presentation are apparently less common [4% in one study (104)]. Hemorrhage occurs with excessively high platelet counts (75, 104).

The overall risk for thrombosis in ET patients is around 10%. Major thrombotic events occur in about 20% to 30% of the patients (74, 105). Significant risk factors for the development of thrombotic episodes as determined in a study by Cortelazzo et al. (106) included age (over 60 years old), previous thrombotic event and long duration of thrombocytosis. Smoking, hypertension, diabetes mellitus, and hyperlipidemia apparently did not affect the risk for thrombosis. It generally is accepted that neither the degree of thrombocytosis nor qualitative abnormalities of the platelets correlate well with the risk of bleeding or thrombosis in ET (103).

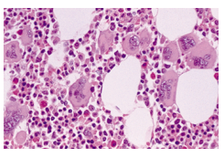


FIGURE 45.11. Bone marrow biopsy of essential thrombocythemia with megakaryocytic hyperplasia. The megakaryocytes are clustering and show characteristic morphology: large size and hyperlobated nuclei.

Peripheral Blood Findings

ET is a diagnostic possibility whenever a significant thrombocytosis is seen in a complete blood count. It is important to remember that ET is a diagnosis of exclusion. What is a significant thrombocytosis? A platelet count greater than 600 × 10^9/L has been cited in the past as a diagnostic criterion for ET; however, that has been questioned (107). The differential diagnosis for thrombocytosis falls into two major categories (Table 45.5): reactive versus malignant (ET or other types of CMPD). Reactive thrombocytosis includes postsplenectomy, chronic iron deficiency (platelet count as high as 1,000 × 10^9/L has been reported), hemorrhage, chronic inflammatory states such as systemic lupus erythematosus and other collagen vascular disorders, chronic infections, nonhematopoietic malignancies, and drug effects such as vincristine and high dose erythropoietin. Other types of CMPD need to be excluded before a diagnosis of ET can be rendered. In particular, CML and PV often present with some degree of thrombocytosis. Peripheral basophilia is expected in

CML, and the morphology of the megakaryocytes in CML is different from that seen in ET. Ultimately, the presence or absence of the Ph chromosome will provide the definitive answer. PV can often masquerade as ET with concomitant iron deficiency and apparently normal hemoglobin levels. A trial of iron therapy may be necessary before such a distinction can be made.

TABLE 45.5. DIFFERENTIAL DIAGNOSIS OF THROMBOCYTOSIS

Hemorrhage
Chronic iron deficiency
Postsplenectomy
Drugs
Chronic inflammation/infection
Neoplastic
Essential thrombocythemia
Other CMPDs (CML, PV, CIMF)

CIMF, chronic idiopathic myelofibrosis
CML, chronic myelogenous leukemia
CMPD, chronic myeloproliferative disorders
PV, polycythemia vera

The peripheral blood smear is often unremarkable, except for confirming the thrombocytosis. Morphology of the platelets is variable, ranging from normal to highly atypical forms. The white blood cell count is in general modestly increased without basophilia, and the hemoglobin level is normally or mildly decreased.

Bone Marrow Findings

The marrow is usually hypercellular with megakaryocytic hyperplasia (Fig. 45.11). The megakaryocytes may or may not cluster and are preferentially located close to bony trabeculae. They are characteristically very large with hyperlobulated nuclei, in contrast to the small, hypolobulated megakaryocytes of CML. There may be erythroid and granulocytic hyperplasia.

Diagnostic Criteria and Other Studies

Similar to PV, the PVSG proposed a set of criteria for the diagnosis of ET (108) (Table 45.6). These criteria primarily excluded other major types of CMPDs that can present with thromobocytosis as well as any causes of reactive thromobocytosis. A normal red-cell mass and the presence of stainable iron exclude PV. The absence of Ph chromosome eliminates the possibility of CML. Absence of significant fibrosis makes chronic idiopathic myelofibrosis (agnogenic myeloid metaplasia) very unlikely. Therefore, in addition to clinical correlation, an appropriate pathologic workup on a potential ET patient includes cytogenetic analysis, iron stain on bone marrow aspirate, and reticulin or trichrome stain to assess degree of marrow fibrosis. Morphologic findings of ET and proliferative phase of PV are very similar and sometimes can be indistinguishable.

TABLE 45.6. PVSG CRITERIA FOR ESSENTIAL THROMBOCYTHEMIA

1.	Platelet count > 600 × 10^9/L
2.	Hemoglobin ≤ 13 gm/dL or normal red cell mass (males < 36 mL/kg, females < 32 mL/kg)
3.	Stainable iron in marrow
4.	No Philadelphia chromosome
5.	Collagen fibrosis of bone marrow
	absent or
	< 1/3 biopsy area, no splenomegaly, no leukoerythroblastosis
6.	Reactive thrombocytosis excluded

(From Murphy S, Iland H, Rosental D, et al. Essential thrombocythemia: an interim report from the Polycythemia Vera Study Group. *Semin Hematol* 23:177–182, reprinted with permission.)

Like PV, revised diagnostic criteria for ET have been proposed, with the advent of new scientific techniques and findings. Table 45.7 lists the revised diagnostic criteria (109). In addition to the retention of some “old” criteria, characteristic bone marrow findings, as described above, are now part of the diagnostic criteria. Although prominent splenomegaly generally is not characteristic of ET, splenomegaly detected on scan or ultrasound favors ET over reactive thrombocytosis. Spontaneous erythroid and/or megakaryocytic colony formation without addition of cytokines also supports ET (110, 111), provided the red cell mass is normal.

Therapy

Although ET generally is considered to have an excellent prognosis with regards to overall survival, it can cause a certain degree of morbidity. Therefore, the decision whether or not to treat must be carefully evaluated, comparing the risks and benefits of treatment and no treatment. In general, a conservative approach of just observation is advocated for younger (under 40 years of

age), asymptomatic patients (105). Treatment is recommended for patients at high risk for thrombosis.

TABLE 45.7. PROPOSED REVISED DIAGNOSTIC CRITERIA FOR ESSENTIAL THROMBOCYTHEMIA

A Criteria		B Criteria	
A1	Platelet count > 600 × 10^9/L	B1	Splenomegaly on isotopic scan or echogram
A2	No increase in red cell mass		
A3	No Philadelphia chromosome	B2	Spontaneous growth of BFU-E and/or CFU-Meg
A4	Characteristic bone marrow	B3	Normal ESR/fibrinogen findings (increased numbers of large megakaryocytes with no collagen fibrosis)

Diagnosis of ET if A1 + A2 + A3 + A4 + or A1 + A2 + A3 + 2B criteria

ET, essential thrombocythemia

(Kutti J, Wadenvik H. Diagnostic and differential criteria of essential thrombocythemia and reactive thrombocytosis. *Leuk Lymphoma* 1996;22(Suppl 1):41–45, reprinted with permission.)

In the 1970s, patients were treated with radioactive phosphorus and alkylating agents such as melphalan and chlorambucil. Because of the increased risk of blastic transformation from these agents, hydroxyurea was introduced in the late 1970s. Hydroxyurea has been shown to be effective and is currently recommended for older patients at risk for thromboses (84).

The PVSG did not find hydroxyurea to be leukemogenic if used alone, after more than 8 years of follow-up of hydroxyurea-treated patients (101). However, there was a high incidence of acute leukemia (five of seven patients) in patients initially treated with hydroxyurea, but then switched to radioactive phosphorus or alkylating agents. A different study found an association between the development of acute myelogenous leukemia and myelodysplastic syndrome in hydroxyurea-treated ET patients and deletion of the short arm of chromosome 17 (112). The typical morphologic characteristics of the 17p- syndrome (dysgranulopoiesis in the form of pseudo Pelger-Huet changes and vacuolation in neutrophils) also were seen.

Other therapeutic agents have been tried, including low-dose aspirin, anegrelide, and interferon-α. Low-dose aspirin has been shown to be effective in ameliorating the microvascular symptoms of erythromelalgia and neurological and visual symptoms (84). Anegrelide is a platelet-lowering drug that has only recently been approved by the Food and Drug Administration. Clinical studies have found 80% to 90% response rate in controlling thrombocytosis (113,114 and 115). Anagrelide currently is recommended for younger, symptomatic patients or as an alternative to other forms of therapy. Because of the high degree of intolerance and cost, interferon-α usually is reserved for high-risk pregnant women or women of childbearing age and for those patients who have failed other forms of therapy (84, 116). The major advantage of interferon-α is its lack of teratogenicity and mutagenicity.

Prognosis/Clinical Course

In general, ET has an excellent prognosis with median survival of more than 10 years. It is important to distinguish ET from reactive thrombocytosis because a patient with ET is more likely to experience thrombohemorrhagic complications. Because most patients with ET can expect to have near normal life expectancy (117), it is important to distinguish ET from other CMPDs.

Transformation to polycythemia vera or myelofibrosis has been documented (74). Blastic transformation of ET occurs in a very small percentage of patients (less than 5%) (90). Usually the blasts are of myeloid origin, but lymphoblastic transformation has been described. Untreated patients are at very low risk for leukemic transformation. However, therapeutic agents such as radioactive phosphorus, and alklyating agents (118) increase the risk of development of acute leukemia.

CHRONIC IDIOPATHIC MYELOFIBROSIS

Part of "45 - Chronic Myeloproliferative Disorders"

This uncommon CMPD, first reported in 1879, has been described under at least 37 different names (119). The more common names include “myelosclerosis or myelofibrosis with myeloid metaplasia (MMM), agnogenic myeloid metaplasia, chronic idiopathic or primary myelofibrosis (CIMF), and megakaryocytic myelosis with osteosclerosis.” Bone marrow fibrosis, extramedullary hematopoiesis as evidenced by massive splenomegaly, leukoerythroblastic blood picture, and osteosclerosis characterize this disorder.

Etiology/Pathogenesis

The exact etiology of CIMF is unknown; however, it is a clonal stem cell disorder. Glucose-6-phosphate dehydrogenase isoenzyme analyses have demonstrated the same isoenzyme in erythrocytes, granulocytes, and platelets in a heterozygous female patient with CIMF (120). Studies using X-linked restriction fragment length polymorphism found clonality in all bone-marrow cells at various stages of disease in heterogzygous female patients (121). Cytogenetic abnormalities have been detected in 30% to 50% of the cases with deletions of the long arm of chromosome 13 and 20 being the most common (122,123,124 and 125).

In contrast, cytogenetic studies of the fibroblasts do not show clonality (126, 127). The marked myelofibrosis that is so characteristic of this disease is believed to be secondary to the stimulation of the nonclonal fibroblasts by several cytokines produced by the clonal, abnormal megakaryocytes (122, 128). The major cytokines or growth factors include platelet-derived growth factor (PDGF), transforming growth factor-ß (TGF-ß), epidermal growth factor (EGF), calmodulin, and basic fibroblast growth factor (bFBF) (122). PDGF as well as EGF, calmodulin, and bFBF are strong fibroblast mitogens. TGF-ß, synthesized by megakaryocytes and stored in the alpha granules of the platelets, upregulates extracellular matrix synthesis, including collagen type I, III, and IV, and hinders degradation of stromal matrix by inhibiting metalloproteinases that are collagenase-like proteases (122, 128). In addition, TGF-ß also promotes angiogenesis.

Extramedullary hematopoiesis that is so characteristic of this disorder is thought to be secondary to the marrow fibrosis that alters and distorts the marrow sinusoids (9, 129). Hematopoietic elements enter the sinusoids, morphologically manifested as intrasinusoidal hematopoiesis. They are transported to the peripheral blood, as evidenced by the leukoerythroblastic blood picture; and then become lodged in extramedullary sites such as the spleen and liver, resulting in organomegaly and extramedullary hematopoiesis.

Clinical Findings

CIMF is less common than CML and PV; roughly three new cases per million people are diagnosed each year. CIMF generally affects older patients in their seventh decade of life without sexual predilection. Not infrequently the patient is asymptomatic at the time of diagnosis with splenomegaly detected incidentally during a routine examination. However, some patients may present with constitutional symptoms (fever, night sweats, and weight loss), fatigue from anemia, hemorrhage secondary to thrombocytopenia or functional defects in platelets, or variceal bleeding from portal hypertension.

The most important clinical finding is the presence of splenomegaly secondary to extramedullary hematopoiesis (Fig. 45.12). A diagnosis of CIMF should be made with great caution

if splenomegaly is absent. The spleens in CIMF are among the largest, comparable to kala-azar, Gaucher's disease, and schistosomiasis. Hepatomegaly is seen in about two thirds of the patients (130). In addition, a significant number of patients (one third to one half) exhibit osteosclerosis, which can be detected in radiographic studies as symmetrically increased bone density (119, 130).

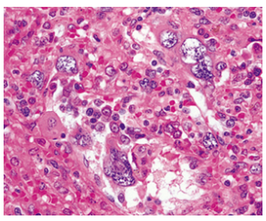


FIGURE 45.12. Extramedullary hematopoiesis in the spleen of a patient with chronic idiopathic myelofibrosis showing large, pleomorphic megakaryocytes in the red pulp (the case is courtesy of the College of American Pathologists, Chicago, IL; Performance Improvement Program, case 98-80).

Peripheral Blood Findings

A diagnosis of CIMF is usually considered when the bone marrow shows prominent myelofibrosis. Supporting morphologic findings in the peripheral blood for this diagnosis include the following: (i) anisopoikilocytosis with numerous teardrop red blood cells, (ii) a leukoerythroblastic blood picture with circulating nucleated red cells and immature granulocytes (Fig. 45.13), and (iii) atypical platelets with circulating micromegakaryocytes. A mild degree of dysplasia may be seen.

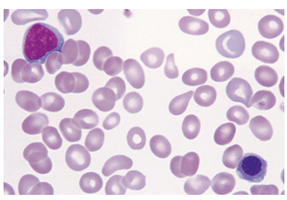


FIGURE 45.13. Typical peripheral blood smear of chronic idiopathic myelofibrosis with leukoerythroblastosis and teardrop red blood cells.

Patients may be anemic (normochromic, normocytic) as a result of ineffective erythropoiesis and splenic sequestration. The white cell count is usually mildly increased, between 10 to 20 × 10^9/L, and circulating myeloblasts may be seen. Up to 10% blasts may be present, but they do not necessarily indicate progression of disease (119, 131). About 20% of CIMF patients are leukopenic, particularly with disease progression (130). A mild basophilia and eosinophilia may be seen in a small proportion of the patients (10% to 30% of the cases) (119). Platelet counts are variable with about half of the patients with thrombocytosis early in the course of the disease. The degree of thrombocytosis may reach the levels seen in ET (131). Thrombocytopenia often results with increasing splenomegaly.

Bone Marrow Findings

Aspiration is generally not successful secondary to the marked myelofibrosis. There are two basic stages in CIMF with a cellular and fibrotic phase. In the cellular or early phase, the bone marrow core biopsy is variably hypercellular with panmyelosis and atypical megakaryocytic hyperplasia accompanied by marked reticulin fibrosis. Hematopoietic elements may show a characteristic streaming pattern with increasing fibrosis (Fig. 45.14). Sinusoids are dilated with intrasinusoidal hematopoiesis. Osteosclerosis characterized by thickened bony trabeculae is often present.

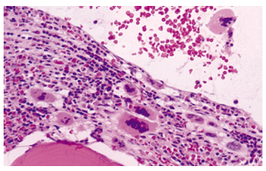


FIGURE 45.14. Bone marrow biopsy from a patient with chronic idiopathic myelofibrosis showing the characteristic streaming pattern secondary to marrow fibrosis.

Megakaryocytes are large, atypical, pleomorphic and often cluster. Megakaryocytic hyperplasia is often seen in CMPDs, but they assume different morphology. Fig. 45.15 compares the variable morphology with small megakaryocytes in CML; relatively normal megakaryocytes in PV; very large, hyperlobulated ones in ET; and large, pleomorphic forms in CIMF. Granulocytic and erythroid series generally are unremarkable, although occasional granulocytic hyperplasia and rare erythroid hypoplasia are seen.

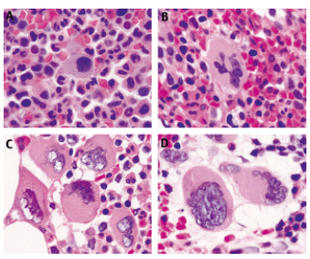


FIGURE 45.15. Composite photomicrograph comparing megakaryocytic morphology in: (**A**) chronic myelogenous leukemia (small, hypolobated); (**B**) polycythemia vera (relatively normal); (**C**) essential thrombocythemia (large, hyperlobated); (**D**) chronic idiopathic myelofibrosis (large, pleomorphic).

In the advanced, fibrotic stage, the bone marrow shows increasing fibrosis with collagen deposition and progressive loss of hematopoietic elements with only clusters of megakaryocytes remaining.

Diagnostic Criteria

In 1975, Laszlo, under the auspices of the Polycythemia Vera Study Group, defined CIMF by the following criteria: splenomegaly, fibrosis involving greater than one third of the

bone marrow biopsy, leukoerythroblastic blood picture, no evidence of an increased red-cell mass, absence of Ph chromosome, and exclusion of causes of secondary myelofibrosis (131).

In 1995, an Italian Cooperative Study Group was formed because of the lack of consistent, standardized criteria for the diagnosis of CIMF (132). Based on the consensus of 12 experts, the preferred name for this disease was myelofibrosis with myeloid metaplasia (MMM). A set of necessary and optional criteria was proposed with the realization that the criteria were limited by the lack of a specific genetic marker that defines this disorder (132). Table 45.8 lists the proposed diagnostic criteria. For a diagnosis of CIMF or MMM, all patients must demonstrate diffuse fibrosis in the bone marrow core biopsy, and the genetic marker for CML must be absent (no evidence of the BCR-ABL rearrangement). Depending on the presence or absence of splenomegaly, two or four of the optional criteria also must be met before such a diagnosis can be rendered.

TABLE 45.8. ITALIAN COOPERATIVE STUDY GROUP PROPOSED CRITERIA FOR MYELOFIBROSIS WITH MYELOID METAPLASIA (CIMF)

Necessary Criteria

1. Diffuse bone marrow fibrosis
2. No Philadelphia chromosome and no evidence of BCR-ABL rearrangement

Optional Criteria

1. Splenomegaly
2. Anisopoikilocytosis with dacrocytes
3. Immature myeloid cells in peripheral blood
4. Erythroblasts in peripheral blood
5. Clustering of megakaryoblasts and anomalous megakaryocytes in bone marrow
6. Myeloid metaplasia

Diagnosis of CIMF if

2 necessary criteria + 2 optional criteria (splenomegaly present)

2 necessary criteria + 4 optional criteria (splenomegaly absent)

From Barosi G, Ambrosetti A, Finelli C, et al. The Italian Consensus Conference on diagnostic criteria for myelofibrosis with myeloid metaplasia. *Br J Haematol* 1999;104:730–737. Reprinted with permission.

Differential Diagnosis

Normal bone marrow has an underlying, delicate reticulin framework with more reticulin around blood vessels and lymphoid aggregates. Marrow fibrosis is initially reticulin fibrosis with collagen fibrosis occurring later. Reticulin is detected by silver stains (Fig. 45.16), and bone marrow contains cells that produce both type III and type I collagen (133). To detect collagen fibrosis, a trichrome stain is required to detect the predominant type I collagen. To determine the degree of reticulin fibrosis, the

number and thickness of reticulin fibers need to be examined, whereas any degree of collagen deposition is considered pathologic and is usually associated with marked reticulin fibrosis and paucity of hematopoiesis (134).

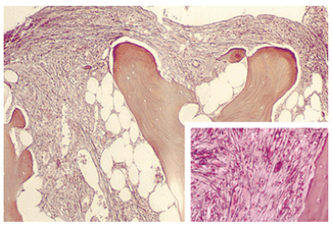


FIGURE 45.16. Silver stain of bone marrow biopsy in chronic idiopathic myelofibrosis showing marked reticulin fibrosis. Bone trabeculae are significantly thickened. The insert illustrates higher magnification.

The differential diagnosis of myelofibrosis is a long list (Table 45.9). Nonneoplastic causes of myelofibrosis include drugs, toxins, radiation, certain chronic infections (such as tuberculosis), Paget's disease or hyperparathyroidism. The neoplastic causes can be separated into hematopoietic and nonhematopoietic malignancies. Metastatic carcinoma should be excluded, including a cytokeratin immunostain, especially in a patient with a history of carcinoma.

TABLE 45.9. DIFFERENTIAL DIAGNOSIS OF MARROW FIBROSIS

Nonneoplastic
– Drugs/toxins
– Infections
– Granulomas
– Miscellaneous
Nonhematopoietic neoplasms (metastatic carcinoma)
Hematopoietic neoplasms
– Chronic lymphoproliferative disorders (hairy cell leukemia, lymphoma)
– Hodgkin's disease
– Mast cell disease
– Myeloproliferative disorders
• Acute myelofibrosis
• Acute myelogenous leukemia (AML-M7)
• Myelodysplastic syndrome
• CIMF
• Other chronic myeloproliferative disorders (PV in spent phase)

CIMF, chronic ideopathic myelofibrosis

PV, polycythemia vera

Hematopoietic neoplasms inciting marrow fibrosis can be classified into four major categories, including marrow involvement by Hodgkin's disease, systemic mast cell disease, chronic lymphoproliferative disorders (e.g., hairy cell leukemia and marrow involvement by non-Hodgkin's lymphoma) and myeloproliferative disorders of the bone marrow. Myeloproliferative disorders include acute myelofibrosis and acute myelogenous leukemia with fibrosis (AML-M7), myelodysplastic syndrome with fibrosis, and CMPDs.

Some degree of marrow fibrosis is usually seen in all CMPDs, but to a much lesser degree than CIMF. The only exception is the postpolycythemic myeloid metaplasia or spent phase of PV, which morphologically can be indistinguishable from CIMF. To differentiate from CML, peripheral blood smears from CML patients in accelerated phase usually do not display teardrop red blood cells. The bone marrow biopsy in CML shows predominant granulocytic hyperplasia with smaller megakaryocytes that are different from the large, pleomorphic, and bizarre megakaryocytes in CIMF. Ultimately, the presence of the Ph chromosome by cytogenetics or evidence of the BCR-ABL rearrangement by molecular methods confirms the diagnosis of CML.

Thrombocytosis above 1,000 × 10^9/L in CIMF may mimic ET. Again, ET does not show the characteristic peripheral blood findings of leukoerythroblastosis and teardrop red blood cells. The morphology of megakaryocytes in ET is different with much less pleomorphism and atypia. In occasional cases, only the clinical course will determine the type of disease. The proliferative stage of PV may be hard to distinguish from the cellular phase of CIMF in occasional cases. Again, PV will not demonstrate teardrop red blood cells, and determination of the red cell mass will oftentimes differentiate between the two, except in PV with iron deficiency (9).

Acute myelofibrosis is a rare disease that behaves in a clinically aggressive manner like acute leukemia, but this disorder is incompletely understood and somewhat controversial (135,136 and 137). However, clinically these patients do not present with splenomegaly, and the characteristic red blood cell changes of CIMF are lacking. The bone marrow in acute myelofibrosis is hypercellular with trilineage hyperplasia accompanied by marked reticulin fibrosis, mimicking CIMF in the cellular phase.

Acute myeloid leukemia with fibrosis, particularly AML-M7, may enter the differential diagnosis. Acute megakaryoblastic leukemia may show a spectrum of undifferentiated blasts to immature or mature megakaryocytes and may appear as marked megakaryocytic hyperplasia with marked reticulin fibrosis on the bone marrow core biopsy (137, 138). The erythroid and granulocytic elements may comprise a small component of the marrow. AML-M7 is favored over CIMF if a predominant blast population is seen at the time of initial disease presentation. Mediastinal germ cell tumors in young men are associated with AML-M7, and the incidence of AML-M7 is also increased in Down's syndrome patients (139).

CIMF does not usually show significant dysplasia. Some cases of myelodysplasia with fibrosis may show significant myelofibrosis and mimic CIMF. However, myelodysplasia with fibrosis is characterized by trilineal dysplasia and pancytopenia, and lacks organomegaly (140).

Clinical Course/Prognosis

The clinical course of CIMF is variable with reported survivals ranging from 1 year to greater than 30 years, but median survival is 5 years after diagnosis. Common causes of death are infection, thrombohemorrhagic episodes, cardiac failure (exacerbated by anemia and hemorrhage), or blastic transformation (141,142 and 143).

Overall, 5% to 10% of patients (9) will undergo blast crisis. Progression to acute leukemia is part of the natural history of this disease and is not limited to patients who received prior alkylating agents (9).

Various prognostic indicators have been implicated in different studies. Poor prognostic indices that have been reported include severe anemia, especially if accompanied by reticulocytopenia; thrombocytopenia; increased numbers of circulating, immature granulocytes; marked marrow fibrosis; constitutional symptoms of fever, night sweats, and weight loss; increased age; and clonal karyotypic abnormalities (141,142,143 and 144). Spleen size has not been shown to significantly affect prognosis (141).

Therapy

Standard therapy is primarily palliative with blood transfusions for the cytopenias; androgen therapy for erythropoiesis; hydroxyurea to improve hepatosplenomegaly and to provide relief of the constitutional symptoms; and splenectomy for refractory cytopenias, symptomatic splenomegaly, and portal hypertension from massive splenomegaly. Allogeneic bone marrow transplantation offers the possibility of a cure, particularly in young patients under the age of 40 (145, 146). Experimental therapeutic agents such as vitamin D, α- and γ-interferons, and erythropoietin have been tried, but show conflicting results (122).

HYPEREOSINOPHILIC SYNDROME/CHRONIC EOSINOPHILIC LEUKEMIA

Part of "45 - Chronic Myeloproliferative Disorders"

Eosinophilia is defined as abnormal accumulation of eosinophils in peripheral blood or tissue. Blood eosinophilia often is associated with tissue eosinophilia that can cause end organ damage. Peripheral blood eosinophilia is present when the absolute eosinophil count exceeds 0.4×10^9/L, and it has been classified into mild (less than 1.5×10^9/L), moderate (greater than 1.5×10^9/L, but less than 5×10^9/L), and severe (greater than 5×10^9/L) (147).

Differential Diagnosis of Eosinophilia

The differential diagnosis of peripheral blood eosinophilia (147,148 and 149) comprises a long and varied list that may be broadly classified into two major categories of reactive, nonclonal causes and neoplastic etiologies. The most common causes of reactive eosinophilia are allergies/atopic diseases (including asthma, rhinitis, eczema, atopic dermatitis), parasitic infections (predominantly helminthic parasites), and drug reactions. Connective tissue diseases such as rheumatoid arthritis and polyarteritis nodosa, cutaneous disorders, and certain neoplasms (e.g., breast and ovarian cancer, Hodgkin's disease) also can cause reactive eosinophilia.

The neoplastic or reactive nature of the eosinophilia in hematopoietic neoplasms may sometimes be difficult to determine. Tumor-associated tissue eosinophilia is seen in extramedullary myeloid tumors (granulocytic sarcoma), peripheral T-cell lymphomas, Hodgkin's disease, and mast cell disease. Primary leukemic and marrow disorders associated with eosinophilia include chronic myelogenous leukemia, chronic eosinophilic leukemia (Fig. 45.17), acute myelogenous leukemia-M4 with increased eosinophils (AML-M4 EO) associated with chromosome 16 abnormalities, and acute lymphoblastic leukemia with eosinophilia. In most cases of AML with eosinophilia, the eosinophils are thought to be part of the malignant clone (150). In contrast, the eosinophilia associated with acute lymphoblastic leukemia is generally reactive (150,151 and 152).

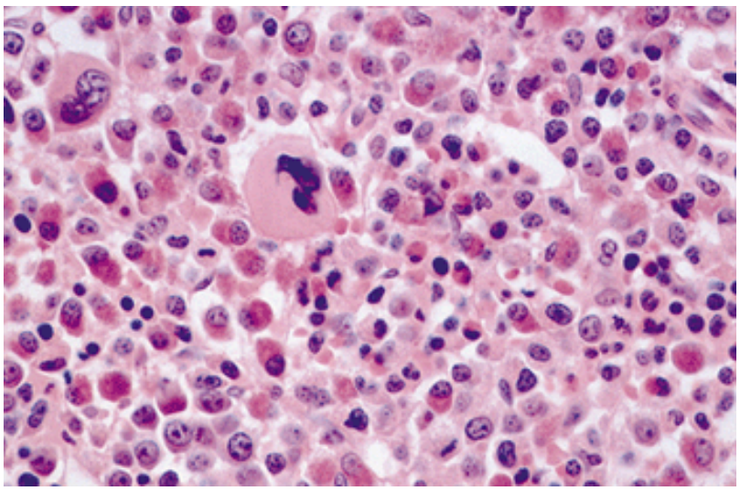


FIGURE 45.17. Bone marrow biopsy of chronic eosinophilic leukemia showing myeloid hypercellularity with marked eosinophilic hyperplasia.

In addition, some diseases are associated with accumulation of eosinophils limited to specific organs (149). Examples include Well's syndrome (eosinophilic cellulitis), Loffler's syndrome (eosinophilic pneumonia), and Shulman's syndrome (eosinophilic fasciitis).

Hypereosinophilic Syndrome (HES)

This diagnosis is reserved for patients with unexplained, persistent blood eosinophilia causing organ damage. This disorder is seen predominantly in men and will progress if not adequately treated. The diagnostic criteria for this disorder include persistent blood eosinophilia of greater than 1.5×10^9/L for longer than 6 months, exclusion of all known causes of eosinophilia, and evidence of organ/tissue damage presumably from the eosinophilia (153, 154). HES is difficult to differentiate from chronic eosinophilic leukemia, as a recent study found clonal eosinophils in a HES patient using clonality assays directed to the human androgen receptor gene (HUMARA) (155). However, such clonality studies require a larger cohort of HES patients.

Organ damage secondary to eosinophilia can occur, regardless of the reactive or clonal nature of the eosinophilia. The heart is the most commonly involved organ and may manifest as constrictive pericarditis, myocarditis, cardiomyopathy, endomyocardial fibrosis, intramural thrombi, and valve regurgitation (147, 156). Neurologic, cutaneous, gastrointestinal, pulmonary, ocular, renal, and articular involvements also have been reported. The tissue damage is believed to be secondary to the released contents of the eosinophil granules together with thrombosis and eosinophilic infiltration of the tissue (147). Eosinophils contain several proteins that can damage endothelium

and cause thrombosis (157). The eosinophil major basic protein and cationic protein can also cause significant tissue damage (158).

Chronic Eosinophilic Leukemia (CEL)

Diagnostic criteria for CEL have been proposed by Weide et al. (159) and Oliver et al. (160). First, the blood eosinophil count is greater than 1.5×10^9/L. All other causes of eosinophilia are excluded. There is no evidence of the Ph chromosome or BCR/ABL translocation to exclude the possibility of eosinophilic variant of CML (159). In addition, there is no evidence of inv (16), t(16;16), or CBFβ/MYH11 rearrangement to exclude AML-M4 EO (159). Finally, eosinophils demonstrate clonal chromosomal abnormalities or two of the following criteria have to be met: (i) immature eosinophil precursors account for greater than 25% of the bone marrow or periperal blood differential count, (ii) myeloblasts comprise greater than 5%, but less than 30% of the bone marrow or peripheral blood differential count, or (iii) eosinophils demonstrate positivity for chloroacetate esterase (160).

Some cases of CEL are diagnosed later on with follow-up, as the patient develops granulocytic sarcoma or blastic transformation, or subsequent studies show evidence of clonality. Morphologic features of the eosinophils usually are not reliable in distinguishing between reactive and neoplastic disorders (150). Hypogranularity/agranularity, cytoplasmic vacuolation, and nuclear abnormalities (hyperlobation, hypolobation, and ring forms) of the eosinophils have been described in CEL (147, 150, 159).

Clonality can only be reliably resolved by cytogenetic and molecular studies. Although most cases have normal karyotype, numerous cytogenetic abnormalities have been reported in CEL (150, 159, 160). Among the more common chromosomal anomalies are trisomy 8, isochromosome 17, monosomy 7, translocations involving the long arm of chromosome 5, and loss of the Y chromosome. Fluorescence *in situ* hybridization also has been used to demonstrate cytogenetic abnormalities in the eosinophils (161). Clonality assays utilizing glucose-6-phosphate dehydrogenase isoenzyme analysis or X-linked gene inactivation patterns in heterozygous females have been used (147, 150). Some chromosomal abnormalities have been associated with specific clinical presentations. For example, t(5;12) has been associated with a chronic myeloproliferative disorder with marrow eosinophilia (150, 162). These patients are mostly men with hepatosplenomegaly, neutrophilia, and low neutrophil alkaline phosphatase score. In addition, an association between t(8;13), T-lineage lymphoblastic lymphoma and myeloid hyperplasia/malignancy in the bone marrow with eosinophilia has been described (163, 164).

It has been proposed that CEL, similar to CML, has a chronic phase, but blastic transformation can occur (147, 159, 160). Some patients, however, die soon after diagnosis without an increase in blasts in their bone marrow. The early death is thought to be primarily from HES (150). Treatment options for CEL include hydroxyurea, interferon, aggressive chemotherapy, or allogeneic bone marrow transplantation (159).

REFERENCES

1. Harris NL, Jaffe ES, Diebold J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee Meeting-Arlie House, Virginia, November 1997. *J Clin Oncol* 1999;17:3835–3849.

2. Bain, BJ. *Leukaemia Diagnosis*, 2nd ed. Oxford, UK: Blackwell Science Ltd., 1999:143.

3. Faderl S, Talpaz M, Estrov Z, et al. The biology of chronic myeloid leukemia. *N Engl J Med* 1999;341:164–172.

4. Sawyers CL. Chronic myeloid leukemia. *N Engl J Med* 1999;340:1330–1340.

5. Verfaillie CM. Biology of chronic myelogenous leukemia. *Hematol Oncol Clin North Am* 1998;12:1–29.

6. Gordon MY, Dowdling CR, Riley GP, et al. Altered adhesive interactions with marrow stroma of haematopoietic progenitor cells in chronic myeloid leukemia. *Nature* 1987;328:342–343.

7. Gishizky ML, Witte DN. Initiation of deregulated growth of multipotent progenitor cells by bcr-abl in vitro. *Science* 1992:256:836–839.

8. Kantarjian HM, Giles FJ, O'Brien SM, Talpaz M. Clinical course and therapy of chronic myelogenous leukemia with interferon-alpha and chemotherapy. *Hematol Oncol Clin North Am* 1998;12:31–69.

9. Vardiman JW. Chronic myelogenous leukemia and the myeloproliferative disorders. In: Knowles DM, ed. *Neoplastic Hematopathology*. Baltimore, MD: Williams & Wilkins, 1992:1405–1438.

10. Kantarjian HM, Talpaz M, Gutterman JU. Chronic myelogenous leukemia: past, present, and future. *Hematol Pathol* 1988:2:91–120.

11. Meyerson HJ, Farhi DC, Rosenthal NS. Transient increase in blasts mimicking acute leukemia and progressing myelodysplasia in patients receiving growth factor. *Am J Clin Pathol* 1998;109:675–681.

12. Spiers ASD, Bain BJ, Turner JE. The peripheral blood in chronic granulocytic leukaemia: study of 50 untreated Philadelphia-positive cases. *Scand J Haematol* 1977;18:25–38.

13. Dickstein JI, Vardiman JW. Issues in the pathology and diagnosis of the chronic myeloproliferative disorders and the myelodysplastic syndromes. *Am J Clin Pathol* 1993;99:513–525.

14. Kurzrock R, Gutterman JU, Talpaz M. The molecular genetics of Philadelphia chromosome-positive leukemias. *N Engl J Med* 1988;319:990–998.

15. Nowell PC, Hungerford DA. A minute chromosome in human chronic myelogenous leukemia. *Science* 1960;132:1497–1499.

16. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 1973;243:290–293.

17. Roth MS, Antin JH, Ash R, et al. Prognostic significance of Philadelphia chromosome-positive cells detected by the polymerase chain reaction after allogeneic bone marrow transplant for chronic myelogenous leukemia. *Blood* 1992;79:276–282.

18. Miyamura K, Tahar T, Tanimoto M, et al. Long persistent BCR-ABL positive transcript detected by polymerase chain reaction after marrow transplant for chronic myelogenous leukemia without clinical relapse: a study of 64 patients. *Blood* 1993;81:1089–1093.

19. Radich JP, Gehly G, Gooley T, et al. Polymerase chain reaction detection of the *BCR-ABL* fusion transcript after allogeneic marrow transplantation for chronic myeloid leukemia: results and implications in 346 patients. *Blood* 1995;85:2632–2638.

20. Biernaux C, Loos M, Sels A, et al. Detection of major BCR-ABL gene expression at a very low level in blood cells of some healthy individuals. *Blood* 1995;86:3118–3122.

21. Bose S, Deininger M, Gora-Tybor J, et al. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood* 1998;92:3362–3367.

22. Faderl S, Talpaz M, Kantarjian HM, Estrov Z. Should polymerase chain reaction analysis to detect minimal residual disease in patients with chronic myelogenous leukemia be used in clinical decision making? *Blood* 1999;93:2755–2759.

23. Cox MC, Maffei L, Buffolino S, et al. A comparative analysis of FISH, RT-PCR, and cytogenetics for the diagnosis of *BCR-ABL*-positive leukemias. *Am J Clin Pathol* 1998;109:24–31.

24. Tbakhi A, Pettay J, Sreenan JJ, et al. Comparative analysis of interphase FISH and RT-PCR to detect *BCR-ABL*translocation in chronic myelogenous leukemia and related disorders. *Am J Clin Pathol* 1998;109:16–23.

25. Third International Workshop on Chromosomes in Leukemia. Clinical significance of chromosomal abnormalities in acute lymphoblastic leukemia. *Cancer Genet Cytogenet* 1981;4:101–137.

26. Bloomfield CD, Goldman AI, Alimena G, et al. Chromosomal abnormalities identify high-risk and low-risk patients with acute lymphoblastic leukemia. *Blood* 1986;67:415–420.

27. Peterson LC, Parkin JL, Arthur DC, et al. Acute basophilic leukemia: a clinical, morphologic and cytogenetic study of eight cases. *Am J Clin Pathol* 1991;96:160–170.

28. Matutes E, Morilla R, Farahat N, et al. Definition of acute biphenotypic leukemia. *Haematologica* 1997;82:64–66.

29. Foucar K. Neutrophilia. In: Kjeldsberg CR, ed. *Practical Diagnosis of Hematologic Disorders*, 2nd ed. Chicago, IL: ASCP Press, 1995:225–235.

30. DePalma L, Delgado P, Werner M. Diagnostic discrimination and cost effective assay strategy for leukocyte alkaline phosphatase. *Clin Chim Acta* 1996;244:83–90.

31. Gay JC, Athens JW. Variations of leukocytes in disease. In: Lee RG, Foerster J, Lukens J, et al, eds. *Wintrobe's Clinical Hematology*. Baltimore, MD: Williams & Wilkins, 1999:1836–1861.

32. Galton DAG. Haematological differences between chronic granulocytic leukaemia, atypical chronic myeloid leukaemia, and chronic myelomonocytic leukaemia. *Leuk Lymphoma* 1992;7:343–350.

33. Bennett JM, Catovsky D, Daniel MT, et al. The chronic myeloid leukaemias: guidelines for distinguishing chronic granulocytic, atypical chronic myeloid, and chronic myelomonocytic leukaemia. Proposals by the French-American-British Cooperative Leukaemia Group. *Br J Haematol* 1994;87:746–754.

34. Dickstein JI, Vardiman JW. Hematopathologic findings in the myeloproliferative disorders. *Semin Oncol* 1995;22:355–373.

35. Pugh WC, Pearson M, Vardiman JW, Rowley JD. Philadelphia chromosome-negative chronic myelogenous leukaemia: a morphological reassessment. *Br J Haematol* 1985;60:457–467.

36. Martiat P, Michaux JL, Rodhain J for the Groupe Français de Cytogénétique Hématologique. Philadelphia-Negative (Ph^-) chronic myeloid leukemia (CML): Comparison with Ph^+ CML and chronic myelomonocytic leukemia. *Blood* 1991;78:205–211.

37. Germing U, Gattermann N, Minning H, et al. Problems in the classification of CMML—dysplastic versus proliferative type. *Leuk Res* 1998;22:871–878.

38. Niemeyer CM, Fenu S, Hasle H, et al. Response: differentiating juvenile myelomonocytic leukemia from infectious disease. *Blood* 1998;91:365–367.

39. Aricó M, Biondi A, Pui C-H. Juvenile myelomonocytic leukemia. *Blood* 1997;90:479–488.

40. Luna-Fineman S, Shannon KM, Atwater SK, et al. Myelodysplastic and myeloproliferative disorders of childhood: a study of 167 patients. *Blood* 1999;93:459–466.

41. Niemeyer CM, Aricó M, Basso G, et al. Chronic myelomonocytic leukemia in childhood: a retrospective analysis of 110 cases. *Blood* 1997;89:3534–3543.

42. Emanuel PD, Bates LJ, Castleberry RP, et al. Selective hypersensitivity to granulocyte-macrophage colony-stimulating factor by juvenile chronic myeloid leukemia hematopoietic progenitors. *Blood* 1991;77:925–929.

43. Castro-Malaspina H, Schaison G, Passe S, et al. Subacute and chronic myelomonocytic leukemia in children (Juvenile CML). Clinical and hematologic observations, and identification of prognostic factors. *Cancer* 1984;54:675–686.

44. Hess JL, Zutter MM, Castleberry RP, Emanuel PD. Juvenile chronic myelogenous leukemia. *Am J Clin Pathol* 1996;105:238–248.

45. Luna-Fineman S, Shannon KM, Lange BJ. Childhood monosomy 7: epidemiology, biology, and mechanistic implications. *Blood* 1995;85:1985–1999.

46. Brunning RD, McKenna RW. Chronic myeloproliferative diseases. In: Rosai J and Sobin LH, eds. *Atlas of Tumor Pathology: Tumors of the Bone Marrow*. Washington, DC: Armed Forces Institute of Pathology, 1994:195–254.

47. Matano S, Nakamura S, Kobayashi K, et al. Deletion of the long arm of chromosome 20 in a patient with chronic neutrophilic leukemia: cytogenetic findings in chronic neutrophilic leukemia. *Am J Hematol* 1997;54:72–75.

48. Zittoun R, Réa D, Hoang Ngoc L, et al. Chronic neutrophilic leukemia. A study of four cases. *Ann Hematol* 1994;68:55–60.

49. Kantarjian HM, Dixon D, Keating MJ, et al. Characteristics of accelerated disease in chronic myelogenous leukemia. *Cancer* 1988;6l:1441–1446.

50. Muehleck SD, McKenna RW, Arthur DC, et al. Transformation of chronic myelogenous leukemia: clinical, morphologic, and cytogenetic features. *Am J Clin Pathol* 1984;82:1–14.

51. Rosenthal S, Canellos GP, DeVita VT Jr, et al. Characteristics of blast crisis in chronic granulocytic leukemia. *Blood* 1977;49:705–714.

52. Kantarjian HM, Keating MJ, Talpaz M, et al. Chronic myelogenous leukemia in blast crisis. Analysis of 242 patients. *Am J Med* 1987;83:445–454.

53. Fioretos T, Strömbeck B, Sandberg T, et al. Isochromosome 17q in blast crisis of chronic myeloid leukemia and in other hematologic malignancies is the result of clustered breakpoints in 17p11 and is not associated with coding *TP53* mutations. *Blood* 1999;94:225–232.

54. Derderian PM, Kantarjian HM, Talpaz M, et al. Chronic myelogenous leukemia in the lymphoid blastic phase: characteristics, treatment response, and prognosis. *Am J Med* 1993;94:69–74.

55. Urbano-Ispizua A, Cervantes F, Matutes E, et al. Immunophenotypic characteristics of blast crisis of chronic myeloid leukaemia: correlations with clinico-biological features and survival. *Leukemia* 1993;7:1349–1354.

56. Cervantes F, Villamor N, Esteve J, et al. 'Lymphoid' blast crisis of chronic myeloid leukaemia is associated with distinct clinicohaematological features. *Br J Haematol* 1998;100:123–128.

57. Dorfman DM, Longtine JA, Fox EA, et al. T-cell blast crisis in chronic myelogenous leukemia. Immunophenotypic and molecular biologic findings. *Am J Clin Pathol* 1997;107:168–176.

58. Menasce LP, Banerjee SS, Beckett E, et al. Extra-medullary myeloid tumor (granulocytic sarcoma) is often misdiagnosed: a study of 26 cases. *Histopathology* 1999;34:391–398.

59. Silver RT, Woolf SH, Hehlmann R, et al. An evidence-based analysis of the effect of busulfan, hydroxyurea, interferon, and allogeneic bone marrow transplantation in treating the chronic phase of chronic myeloid leukemia: developed for the American Society of Hematology. *Blood* 1999;94:1517–1536.

60. Bhatia R, Verfaillie CM, Miller JS, et al. Autologous transplantation therapy for chronic myelogenous leukemia. *Blood* 1997;89:2623–2634.

61. Rege-Cambrin G, Mecucci C, Tricot G, et al. A chromosomal profile of polycythemia vera. *Cancer Genet Cytogenet* 1987;25:233–245.

62. Mertens F, Johansson B, Heim S, et al. Karyotypic patterns in chronic myeloproliferative disorders: report on 74 cases and review of the literature. *Leukemia* 1991;5:214–220.

63. Swolin B, Weinfeld A, Westin J. A prospective long-term cytogenetic study in polycythemia vera in relation to treatment and clinical course. *Blood* 1988;72:386–395.

64. Diez-Martin JL, Graham DL, Petitt RM, et al. Chromosome studies in 104 patients with polycythemia vera. *Mayo Clin Proc* 1991;66:287–299.

65. Hinshelwood S, Bench AJ, Green AR. Pathogenesis of polycythaemia vera. *Blood Rev* 1997;11:224–232.

66. Dewald GW, Wright PI. Chromosome abnormalities in the myeloproliferative disorders. *Semin Oncol* 1995;22:341–354.

67. Adamson JW, Fialkow PJ, Murphy S, et al. Polycythemia vera: stem-cell and probable clonal origin of the disease. *N Engl J Med* 1976;295:913–916.

68. Anger B, Janssen JWG, Schrezenmeier H, et al. Clonal analysis of chronic myeloproliferative disorders using X-linked DNA polymorphisms. *Leukemia* 1990;4:258–261.

69. Lucas GS, Padua RA, Masters GS, et al. The application of X-chromosome gene probes to the diagnosis of myeloproliferative disease. *Br J Haematol* 1989;72:530–533.

70. Champion KM, Gilbert JGR, Asimakopoulos FA, et al. Clonal haemopoiesis in normal elderly women: implications for the myeloproliferative disorders and myelodysplastic syndromes. *Br J Haematol* 1997;97:920–926.

71. Mitterbauer G, Winkler K, Gisslinger H, et al. Clonality analysis using X-chromosome inactivation at the human androgen receptor gene (HUMARA): evaluation of large cohorts of patients with chronic myeloproliferative diseases, secondary neutrophilia, and reactive thrombocytosis. *Am J Clin Pathol* 1999;112:93–100.

72. Prchal JF, Axelrad AA. Bone-marrow responses in polycythemia vera. *N Engl J Med* 1974;290:1382.

73. Birgegar G, Wide L. Serum erythropoietin in the diagnosis of polycythaemia and after phlebotomy treatment. *Br J Haematol* 1992;81:603–606.

74. Murphy S. Diagnostic criteria and prognosis in polycythemia vera and essential thrombocythemia. *Semin Hematol* 1999;36:9–13.

75. Michiels JJ. The myeloproliferative disorders. An historical appraisal and personal experiences. *Leuk Lymphoma* 1996;22:1–14.

76. Barbui T, Finazzi G. Risk factors and prevention of vascular complications in polycythemia vera. *Semin Thromb Hemost* 1997;23:455–461.

77. Berlin NI. Diagnosis and classification of the polycythemias. *Semin Hematol* 1975;12:339–351.

78. Ellis JT, Peterson P, Geller SA, et al. Studies of the bone marrow in polycythemia vera and the evolution of myelofibrosis and second hematologic malignancies. *Semin Hematol* 1986;23:144–155.

79. Pearson TC, Messinezy M. The diagnostic criteria of polycythaemia rubra vera. *Leuk Lymphoma* 1996;22:87–93.

80. Najean Y, Dresch C, Rain J-D, et al. Radioisotope investigations for the diagnosis and follow-up of polycythemic patients. In: Wasserman LR, Berk PD, Berlin NI, eds. *Polycythemia Vera and the Myeloproliferative Disorders*. Philadelphia: W.B. Saunders Company, 1995:79–90.

81. Westwood NB, Pearson TC. Diagnostic applications of haemopoietic progenitor culture techniques in polycythaemias and thrombocythaemias. *Leuk Lymphoma* 1996;22:95–103.

82. Landaw SA. Acute leukemia in polycythemia vera. *Semin Hematol* 1986;23:156–165.

83. Wolf BC, Banks PM, Mann RB, et al. Splenic hematopoiesis in polycythemia vera. A morphologic and immunohistologic study. *Am J Clin Pathol* 1988;89:69–75.

84. Gilbert HS. Historical perspective on the treatment of essential thrombocythemia and polycythemia vera. *Semin Hematol* 1999;36:19–22.

85. Chievitz E, Thiede T. Complications and causes of death in polycythemia vera. *Acta Med Scand* 1962;172:513–523.

86. Michiels JJ. Erythromelalgia and vascular complications in polycythemia vera. *Semin Thromb Hemost* 1997;23:441–454.

87. Willoughby S, Pearson TC. The use of aspirin in polycythaemia vera and primary thrombocythaemia. *Blood Rev* 1998;12:12–22.

88. Taylor PC, Dolan G, Ng J-P, et al. Efficacy of recombinant interferon-alpha (rIFN-α) in polycythaemia vera: a study of 17 patients and an analysis of published data. *Br J Haematol* 1996;92:55–59.

89. Anderson JE, Sale G, Appelbaum FR, et al. Allogeneic marrow transplantation for primary myelofibrosis and myelofibrosis secondary to polycythaemia vera or essential thrombocytosis. *British J Haematol* 1997;98:1010–1016.

90. Fenaux P, Simon M, Caulier MT, et al. Clinical course of essential thrombocythemia in 147 cases. *Cancer* 1990;66:549–556.

91. Fialkow PJ, Faguet GB, Jacobson RJ, et al. Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. *Blood* 1981;58:916–919.

92. El-Kassar N, Hetet G, Briere J, et al. Clonality analysis of hematopoiesis and thrombopoietin levels in patients with essential thrombocythemia. *Leuk Lymphoma* 1998;30:181–188.

93. Harrison CN, Gale RE, Machin SJ, et al. A large proportion of patients with a diagnosis of essential thrombocythemia do not have a clonal disorder and may be at lower risk of thrombotic complications. *Blood* 1999;93:417–424.

94. Nimer SD. Essential thrombocythemia: another "heterogeneous disease" better understood? *Blood* 1999;93:415–416.

95. Eyster ME, Saletan SL, Rabellino EM, et al. Familial essential thrombocythemia. *Am J Med* 1986;80:497–502.

96. Kikuchi M, Tayama T, Hayakawa H, et al. Familial thrombocytosis. *Br J Haematol* 1995;89:900–902.

97. Kondo T, Okabe M, Sanada M, et al. Familial essential thrombocythemia associated with one-base deletion in the 5′-untranslated region of the thrombopoietin gene. *Blood* 1998;92:1091–1096.

98. Cripe LD, Hromas R. Malignant disorders of megakaryocytes. *Semin Hematol* 1998;35:200–209.

99. Horikawa Y, Matsumura I, Hashimoto K, et al. Markedly reduced expression of platelet c-mpl receptor in essential thrombocythemia. *Blood* 1997;90:4031–4038.

100. Buss DH, Cashell AW, O'Connor ML, et al. Occurrence, etiology, and clinical significance of extreme thrombocytosis: a study of 280 cases. *Am J Med* 1994;96:247–253.

101. Murphy S, Peterson P, Iland H, et al. Experience of the Polycythemia Vera Study Group with essential thrombocythemia: a final report on diagnostic criteria, survival, and leukemic transition by treatment. *Semin Hematol* 1997;34:29–39.

102. Van Genderen PJJ, Michiels JJ. Erythromelalgia: a pathognomonic microvascular thrombotic complication in essential thrombocythemia and polycythemia vera. *Semin Thromb Hemost* 1997;23:357–363.

103. Ravandi-Kashani F, Schafer AI. Microvascular disturbances, thrombosis, and bleeding in thrombocythemia: current concepts and perspectives. *Semin Thromb Hemost* 1997;23:479–488.

104. Colombi M, Radaelli F, Zocchi L, et al. Thrombotic and hemorrhagic complications in essential thrombocythemia. A retrospective study of 103 patients. *Cancer* 1991;67:2926–2930.

105. Barbui T, Finazzi G. Clinical parameters for determining when and when not to treat essential thrombocythemia. *Semin Hematol* 1999;36:14–18.

106. Cortelazzo S, Viero P, Finazzi G, et al. Incidence and risk factors for thrombotic complications in a historical cohort of 100 patients with essential thrombocythemia. *J Clin Oncol* 1990:8:556–562.

107. Lengfelder E, Hochhaus A, Kronawitter U, et al. Should a platelet limit of $600 \times 10^9/1$ be used as a diagnostic criterion in essential thrombocythaemia? An analysis of the natural course including early stages. *Br J Haematol* 1998;100:15–23.

108. Murphy S, Iland H, Rosenthal D, et al. Essential thrombocythemia: an interim report from the Polycythemia Vera Study Group. *Semin Hematol* 1986;23:177–182.

109. Kutti J, Wadenvik H. Diagnostic and differential criteria of essential thrombocythemia and reactive thrombocytosis. *Leuk Lymphoma* 1996;22 (suppl 1):41–45.

110. Juvonen E, Ikkala E, Oksanen K, et al. Megakaryocyte and erythroid colony formation in essential thrombocythaemia and reactive thrombocytosis: diagnostic value and correlation to complications. *Br J Haematol* 1993;83:192–197.

111. Rolovic Z, Basara N, Gotic M, et al. The determination of spontaneous megakaryocyte colony formation is an unequivocal test for discrimination between essential thrombocythaemia and reactive thrombocytosis. *Br J Haematol* 1995;90:326–331.

112. Sterkers Y, Preudhomme C, Laï J-L, et al. Acute myeloid leukemia and myelodysplastic syndromes following essential thrombocythemia treated with hydroxyurea: high proportion of cases with 17p deletion. *Blood* 1998;91:616–622.

113. Anagrelide Study Group. Anagrelide, a therapy for thrombocythemic states: experience in 577 patients. *Am J Med* 1992;92:69–76.

114. Tefferi A, Silverstein MN, Petitt RM, et al. Anagrelide as a new platelet-lowering agent in essential thrombocythemia: mechanism of action, efficacy, toxicity, current indications. *Semin Thromb Hemost* 1997;23:379–383.

115. Silverstein MN, Tefferi A. Treatment of essential thrombocythemia with anagrelide. *Semin Hematol* 1999;36(suppl 2):23–25.

116. Elliott MA, Tefferi A. Interferon-α therapy in polycythemia vera and essential thrombocythemia. *Semin Thromb Hemost* 1997;23:463–472.

117. Rozman C, Giralt M, Feliu E, et al. Life expectancy of patients with chronic nonleukemic myeloproliferative disorders. *Cancer* 1991;67:2658–2663.

118. Killick S, Matutes E, Swansbury J, et al. Teaching cases from the Royal Marsden Hospital. Case 17: Essential thrombocythaemia with inversion 3 terminating in acute leukaemia. *Leuk Lymphoma* 1998;30:661–664.

119. Ward HP, Block MH. The natural history of agnogenic myeloid metaplasia (AMM) and a critical evaluation of its relationship with the myeloproliferative syndrome. *Medicine* 1971;50:357–420.

120. Jacobson RJ, Salo A, Fialkow PJ. Agnogenic myeloid metaplasia: a clonal proliferation of hematopoietic stem cells with secondary myelofibrosis. *Blood* 1978;51:189–194.

121. Kreipe H, Jaquet K, Felgner J, et al. Clonal granulocytes and bone marrow cells in the cellular phase of agnogenic myeloid metaplasia. *Blood* 1991;78:1814–1817.

122. Reilly JT. Idiopathic myelofibrosis: pathogenesis, natural history and management. *Blood Rev* 1997;11:233–242.

123. Reilly JT, Snowden JA, Spearing RL, et al. Cytogenetic abnormalities and their prognostic significance in idiopathic myelofibrosis: a study of 106 cases. *Br J Haematol* 1997;98:96–102.

124. Demory JL, Dupriez B, Fenaux P, et al. Cytogenetic studies and their prognostic significance in agnogenic myeloid metaplasia: a report on 47 cases. *Blood* 1988;72:855–859.

125. Miller JB, Testa JR, Lindgren V, Rowley JD. The pattern and clinical significance of karyotypic abnormalities in patients with idiopathic and postpolycythemic myelofibrosis. *Cancer* 1985;55:582–591.

126. Greenberg BR, Woo L, Veomett IC, et al. Cytogenetics of bone marrow fibroblastic cells in idiopathic chronic myelofibrosis. *Br J Haematol* 1987;66:487–490.

127. Wang JC, Lang H-D, Lichter S, et al. Cytogenetic studies of bone marrow fibroblasts cultured from patients with myelofibrosis and myeloid metaplasia. *Br J Haematol* 1992;80:184–188.

128. Martyré M-C. TGF-ß and megakaryocytes in the pathogenesis of myelofibrosis in myeloproliferative disorders. *Leuk Lymphoma* 1995;20:39–44.

129. O'Keane JC, Wolf BC, Neiman RS. The pathogenesis of splenic extramedullary hematopoiesis in metastatic carcinoma. *Cancer* 1989;63:1539–1543.

130. Weinstein IM. Idiopathic myelofibrosis: historical review, diagnosis and management. *Blood Rev* 1991;5:98–104.

131. Laszlo J. Myeloproliferative disorders (MPD): myelofibrosis, myelosclerosis, extramedullary hematopoiesis, undifferentiated MPD, and hemorrhagic thrombocythemia. *Semin Hematol* 1975;12:409–432.

132. Barosi G, Ambrosetti A, Finelli C, et al. The Italian Consensus Conference on diagnostic criteria for myelofibrosis with myeloid metaplasia. *Br J Haematol* 1999;104:730–737.

133. Bentley SA, Alabaster O, Foidart JM. Collagen heterogeneity in normal human bone marrow. *Br J Haematol* 1981;48:287–291.

134. Foucar, K. Miscellaneous disorders of bone marrow, including stromal and bone abnormalities. In: Foucar K, ed. *Bone Marrow Pathology*. Chicago: ASCP Press, 1995:431–474.

135. Fabich DR, Raich PC. Acute myelofibrosis: a report of three cases. *Am J Clin Pathol* 1977;67:334–338.

136. Bearman RM, Pangalis GA, Rappaport H. Acute ("malignant") myelosclerosis. *Cancer* 1979;43:279–293.

137. Hruban RH, Kuhajda FP, Mann RB. Acute myelofibrosis: immunohistochemical study of four cases and comparison with acute megakaryocytic leukemia. *Am J Clin Pathol* 1987;88:578–588.

138. Gassmann W, Loffler H. Acute megakaryoblastic leukemia. *Leuk Lymphoma* 1995;18(suppl 1):69–73.

139. DeMent SH. Association between mediastinal germ cell tumors and hematologic malignancies: an update. *Hum Pathol* 1990;21:699–703.

140. Lambertenghi-Deliliers G, Orazi A, Luksch R, et al. Myelodysplastic syndrome with increased marrow fibrosis: a distinct clinico-pathological entity. *Br J Haematol* 1991;78:161–166.

141. Visani G, Finelli C, Castelli U, et al. Myelofibrosis with myeloid metaplasia: clinical and haematological parameters predicting survival in a series of 133 patients. *Br J Haematol* 1990;75:4–9.

142. Cervantes F, Pereira A, Esteve J, et al. Identification of 'short-lived' and 'long-lived' patients at presentation of idiopathic myelofibrosis. *Br J Haematol* 1997;97:635–640.

143. Cervantes F, Barosi G, Demory J-L, et al. Myelofibrosis with myeloid metaplasia in young individuals: disease characteristics, prognostic factors and identification of risk groups. *Br J Haematol* 1998;102:684–690.

144. Dupriez B, Morel P, Demory JL, et al. Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood* 1996;88:1013–1018.

145. Guardiola P, Esperou H, Cazals-Hatem D, et al. Allogeneic bone marrow transplantation for agnogenic myeloid metaplasia. *Br J Haematol* 1997;98:1004–1009.

146. Guardiola P, Anderson JE, Bandini G, et al. for the International Collaboration for Transplantation in Agnogenic Myeloid Metaplasia. Allogeneic stem cell transplantation for agnogenic myeloid metaplasia: a European group for blood and marrow transplantation, Société Française de Greffe de Moelle, Gruppo Italiano per il Trapianto del Midollo Osseo, and Fred Hutchinson Cancer Research Center Collaborative Study. *Blood* 1999;93:2831–2838.

147. Brito-Babapulle F. Clonal eosinophilic disorders and the hypereosinophilic syndrome. *Blood Rev* 1997;11:129–145.

148. Foucar K. Eosinophilia. In: Kjeldsberg CR, ed. *Practical Diagnosis of Hematologic Disorders*. Chicago: ASCP Press, 1995:237–244.

149. Rothenberg ME. Eosinophilia. *N Engl J Med* 1998;338:1592–1600.

150. Bain BJ. Review: Eosinophilic leukaemias and the idiopathic hypereosinophilic syndrome. *Br J Haematol* 1996;95:2–9.

151. Spitzer G, Garson OM. Lymphoblastic leukemia with marked eosinophilia. *Blood* 1973:42:377–384.

152. Knuutila S, Alitalo R, Ruutu T. Power of the MAC (morphology-antibody-chromosomes) method in distinguishing reactive and clonal cells: report of a patient with acute lymphatic leukemia, eosinophilia, and t(5;14). *Genes Chromosomes Cancer* 1993;8:219–223.

153. Chusid MJ, Dale DC, West BC, et al. The hypereosinophilic syndrome: analysis of fourteen cases with review of the literature. *Medicine* 1975;54:1–27.

154. Fauci AS, Harley JB, Roberts WC, et al. The idiopathic hypereosinophilic syndromes: clinical, pathophysiologic, and therapeutic considerations. *Ann Intern Med* 1982;97:78–92.

155. Chang HW, Leong KH, Koh DR, et al. Clonality of isolated eosinophils in the hypereosinophilic syndrome. *Blood* 1999;93:1651–1657.

156. Parrillo JE, Borer JS, Henry WL, et al. The cardiovascular manifestations of the hypereosinophilic syndrome: prospective study of 26 patients, with review of the literature. *Am J Med* 1979;67:572–582.

157. Rohrbach MS, Wheatley CL, Slifman NR, et al. Activation of platelets by eosinophil granule proteins. *J Exp Med* 1990;172:1271–1274.

158. Young JDE, Peterson CGB, Venge P, et al. Mechanism of membrane damage mediated by human eosinophil cationic protein. *Nature* 1986;321:613–616.

159. Weide R, Rieder H, Mehraein Y, et al. Chronic eosinophilic leukemia (CEL): a distinct myeloproliferative disease. *Br J Haematol* 1997;96:117–123.

160. Oliver JW, Deol I, Morgan DL, et al. Chronic eosinophilic leukemia and hypereosinophilic syndromes. Proposal for classification, literature review, and report of a case with a unique chromosomal abnormality. *Cancer Genet Cytogenet* 1998;107:111–117.

161. Kwong YL, Chan LC. Involvement of eosinophils in acute myeloid leukaemia with monosomy 7 demonstrated by fluorescence in situ hybridization. *Br J Haematol* 1994;88:389–391.

162. Hanson CA, Kurtin PJ, Persons DL, et al. t(5;12)(q31;p13): a chronic myeloproliferative disorder with marrow eosinophilia. *Mod Pathol* 1995;8:112A.

163. Abbruzzo LV, Jaffe ES, Cotelingam JD, et al. T-lymphoblastic lymphoma with eosinophilia associated with subsequent myeloid malignancy. *Am J Surg Pathol* 1992;16:236–245.

164. Inhorn RC, Aster JC, Roach SA, et al. A syndrome of lymphoblastic lymphoma, eosinophilia, and myeloid hyperplasia/malignancy associated with t(8;13)(p11;q11): description of a distinctive pathological entity. *Blood* 1993;85:1881–1887.

Section 8 Coagulation

Coagulation - Introduction

Douglas A. Triplett

Section Chief

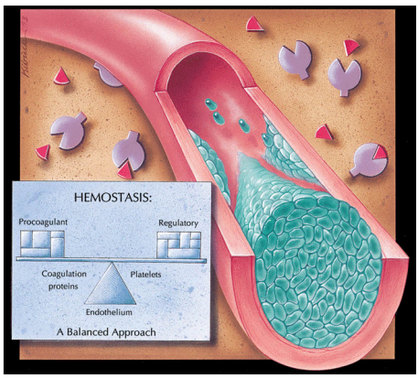


Figure.

46 Overview of Hemostasis

47 Laboratory Evaluation of Platelet Disorders

48 Coagulation Abnormalities

49 Thrombophilia

Our understanding of hemostatic process has increased greatly over the past decade. With the introduction of sophisticated biochemical and molecular biology techniques, the components of this process have been characterized in great detail. This section is designed to provide an overview of the hemostatic process and of disorders resulting from abnormalities of this system. In Chapter 46, the normal mechanisms of hemostasis are reviewed, including primary hemostasis (platelet plug formation), secondary hemostasis (fibrin clot formation), and regulation of hemostasis. Included in this chapter is an introduction to the evaluation of patients with hemostatic problems.

Disorders of primary hemostasis are discussed in Chapter 47. Congenital and acquired disorders of platelet function, including heparin-induced thrombocytopenia and thrombotic thrombocytopenic purpura, are summarized in this chapter. The characteristic laboratory findings in these disorders are reviewed.

Disorders of fibrin clot formation are discussed in Chapter 48. These include the common congenital disorders hemo-philia A and B as well as acquired abnormalities such as factor inhibitors and lupus anticoagulants. Laboratory techniques for diagnosis of these various disorders are reviewed in detail.

Chapter 49 is devoted to problems associated with an increased risk of thrombosis. Emphasis is placed on congenital disorders associated with thrombosis (thrombophilia) and the role of the laboratory in the diagnosis of these conditions. The potential for utilizing markets of hemostasis activation is also reviewed.

46
Overview of Hemostasis

John T. Brandt

Hemostasis may be defined as the process that maintains flowing blood in a fluid state and prevents loss of blood from sites of vascular disruption. This definition implies two major components: first, a potent procoagulant mechanism that is capable of forming stable hemostatic plugs at sites of vascular disruption; and second, regulatory systems that confine normal hemostatic plug formation to sites of vascular disruption. Active components in this system include circulating platelets, procoagulant and regulatory plasma proteins, and endothelial cells lining the vascular wall. Normal hemostasis is a dynamically modulated process in which a delicate balance between the procoagulant and regulatory components of hemostasis is maintained.

The procoagulant mechanism normally is initiated by vascular damage or disruption. This leads to deposition of platelets at the damaged site, a process that is sometimes referred to as primary hemostasis. The platelet plug is reinforced and anchored by the fibrin clot, which is composed of a meshwork of fibrin strands. The formation of the fibrin clot is sometimes referred to as secondary hemostasis. Formation of the hemostatic plug normally is halted as it extends into areas of the vessel lined by intact endothelial cells (Fig. 46.1). The mechanisms involved in turning off the process comprise the regulatory systems of hemostasis. This chapter reviews the major mechanisms involved in primary hemostasis, secondary hemostasis, and regulation of hemostasis. It also outlines an approach to the evaluation of patients with hemostatic disorders. Specific disorders affecting hemostasis are discussed in subsequent chapters in this section.

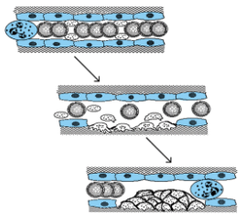


FIGURE 46.1. The formation of a hemostatic plug usually begins when the vessel wall is damaged, exposing the thrombogenic subendothelial tissue to blood. The initial response is platelet adhesion, with the responding platelets forming a monolayer over the exposed subendothelial tissue (**B**). During this response the platelets are activated and additional platelets are recruited to form the platelet plug. This phase of platelet-platelet interaction is known as platelet aggregation. Fibrin formation begins as the responding platelets are activated, resulting in deposition of a fibrin meshwork, which envelops the platelet plug (**C**). Hemostatic plug formation is normally limited to sites of vascular disruption by an active regulatory mechanism that is dependent upon intact endothelial cells.

PRIMARY HEMOSTASIS

Part of "46 - Overview of Hemostasis"

Platelet Production and Morphology

Platelets, an essential component of hemostasis, are anucleate cytoplasmic fragments derived from bone marrow megakaryocytes that are themselves derived from the pluripotent hematopoietic stem cell (1,2,3,4,5 and 6). At least two sequential stages of megakaryocyte progenitors have been demonstrated by *in vitro* culture techniques; the burst-forming unit-megakaryocyte and the more mature colony-forming unit-megakaryocyte. Each of these stages appears to be sensitive to a different set of cytokines and hematopoietic growth factors that may up- or down-regulate megakaryopoiesis. The major growth factor that drives megakaryocyte development is thrombopoietin (TPO), a 332 amino-acid glycoprotein. Thrombopoietin has the capacity to affect both stages of megarkaryocyte development; it can induce production of burst-forming unit-megakaryocytes from pluripotent stem cells and drive maturation and differentiation of colony-forming unit-megakaryocytes. The effect of TPO is mediated by interaction with a specific receptor, cMpl, on progenitor cells. The importance of this ligand-receptor interaction is demonstrated by the severe thrombocytopenia that occurs in experimental animals in which the gene for cMpl is inactivated by homologous recombination. Other cytokines and growth factors, such as IL3, IL6, IL11, erythropoietin and CSF-GM also appear to actively modulate the effect of TPO on megakaryocyte development.

Maturation of the colony-forming unit-megakaryocyte leads to formation of recognizable megakaryocytes. Megakaryocytes are large cells with multilobated nuclei derived from a process of endomitosis. Normal megakaryocytes are usually 8N-32N, but ploidy can vary depending on the rate of platelet production and consumption. Maturation of megakaryocytes is associated with development of platelet specific proteins both within cytoplasmic granules and on the membrane surface. Platelets are formed and released from megakaryocytes by a process of cytoplasmic fragmentation along lines of demarcation formed by infolding of the cytoplasmic membrane. Platelet size thus depends on the size of the zone of demarcation. When platelet production is increased, the zones of demarcation are typically larger, resulting in an increase in the mean platelet volume (MPV), a common feature of thrombocytopenia secondary to increased consumption.

The nonactivated platelet normally has a discoid shape that is maintained by a circumferential band of microtubules (Fig. 46.2). The platelet membrane is physically and biochemically complex. There are numerous infoldings of the external plasma membrane to form the surface-connected open canalicular system. The membrane contains a variety of receptors and functional proteins, some of which are linked to membrane-associated cytoplasmic enzyme systems such as phospholipase A2, phospholipase C, and adenylate cyclase (Table 46.1). Beneath the cytoplasmic membrane is a sol-gel zone rich in actin and myosin. These proteins are involved in the shape change and contraction characteristic of platelet activation and aggregation.

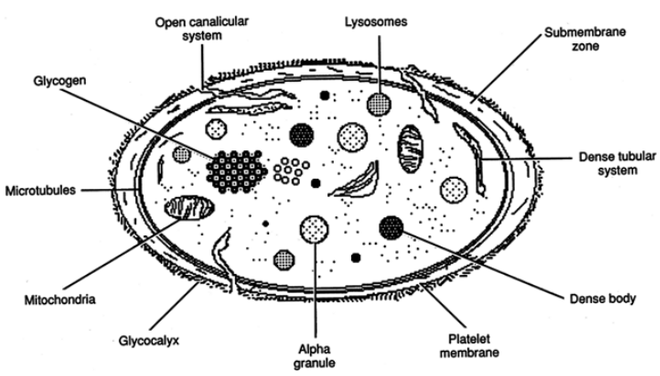


FIGURE 46.2. Key structural features of normal resting platelets are depicted. The open canalicular system is formed by invaginations of the plasma membrane. This system becomes dilated upon platelet activation and serves as a route for platelet secretion. Closely juxtaposed to the open canalicular system is the dense tubular system, which serves as a storage site for calcium. Platelet α-granules contain a variety of proteins, while the major functional constituents of the dense granules are small organic compounds such as ADP, ATP, and serotonin. The discoid shape of the resting platelet is maintained by a circumferential band of microtubules.

Juxtaposed to the surface-connected open canalicular system is a series of membrane-bound tubules referred to as the dense tubular system. The dense tubular system corresponds to the endoplasmic reticulum of smooth muscle cells and functions as a storage site for calcium. Within the cytoplasm are several

organelles, including mitochondria, peroxisomes, α-granules, dense granules and lysosomes; storage pools of glycogen also are frequently present in resting platelets. The α granules contain a variety of proteins, many of which are homologues of plasma proteins (Table 46.2). A long-standing question has been whether or not such proteins are derived from endocytosis by the platelets or active synthesis by the megakaryocyte. Evidence to date suggests that both mechanisms are involved. Platelet IgG, albumin, and fibrinogen concentrations are directly proportional to the plasma concentration of these proteins, suggesting that their presence in the α-granule represents endocytosis from plasma rather than endogenous platelet synthesis. Other proteins, including platelet factor 4, ß-thromboglobulin, and von Willebrand factor (vWF), are clearly synthesized by megakaryocytes.

TABLE 46.1. PLATELET SURFACE PROTEINS AND RECEPTORS

	Protein/Receptor	Alternate Designation	Function
Expressed on nonactivated platelets	GP Ia/IIa	CD49b/CD29	Collagen receptor
	GP Ic/IIa	CD49e/CD29	Fibronectin receptor
	GP Ib/IX/V	CD42b/CD42c/CD42a/CD42d	von Willebrand factor receptor
	GP IIb/IIIa	CD41/CD61	Fibrinogen receptor
	GP IV	CD36	Thrombospondin receptor
	GP Ic'/IIa	CD49f/CD29	Laminin receptor
	Vn Rα/IIIa (α_v $ß_3$)	CD51/CD61	Vitronectin receptor
	GP VI		Collagen receptor
	Thrombin receptor (Protease-activated receptor)		Mediates thrombin activation of platelets
	P2Y1 ADP receptors		ADP stimulation of calcium mobilization
	P2X1 ADP receptors		ADP stimulation of cation influx
	$P2T_{AC}$ ADP receptors		ADP-mediated macroscopic aggregation
	α_2-adrenergic receptors		Mediate epinephrine-induced platelet aggregation
	PECAM-1	CD31	Adhesion molecule (Platelet-endothelial cell adhesion molecule)
	FcγRIIa	CD32	Receptor for immune-complex IgG
	Membrane cofactor protein	CD46	Cofactor for factor I-mediated cleavage of C3b/C4b
	Decay accelerating factor	CD55	Glycosylphosphatidylinositol linked protein that accelerates decay of C3 and C5 convertase
	Membrane inhibitor of reactive lysis (MIRL)	CD59	Glycosylphosphatidylinositol linked inhibitor of complement-mediated cell lysis
	Apolipoprotein E		Binding of ApoE stimulates nitric oxide synthetase
	Receptor 2		
Expression limited to activated platelets	P-Selectin	CD62P	Mediates interaction with leukocytes
	CD40 Ligand	CD154	Induces endothelial-cell activation
	Thrombospondin		Stablizes fibrinogen-GP IIb/IIIa interaction
	Multimerin		Function of surface expression unknown
	GP53	CD63	Lysosomal protein
	LAMP-1		Lysosomal protein

Some α-granule proteins are specific for platelets. Activation of platelets *in vivo* can lead to secretion of these proteins into plasma. As these proteins have a relatively short half-life in circulating blood, the plasma level of proteins such as platelet factor 4 and ß-thromboglobulin can serve as an indicator of *in vivo* platelet activation. Thrombospondin, a major constituent of α-granules, has a limited tissue distribution and has also been used as a marker of platelet activation. Some granule-associated proteins (Table 46.1) redistribute to the platelet surface after platelet activation; measurement of platelet surface expression of such proteins also can be used as an indicator of *in vivo* platelet activation.

The platelet-dense granules contain a variety of small organic compounds, including ADP, ATP, calcium, and serotonin. Perhaps the most important component of the dense granules is ADP, an essential platelet agonist. Patients who have platelets that are deficient in dense granules have a bleeding disorder related to abnormal platelet function. As circulating platelets are anucleate, their ability to synthesize new proteins is relatively limited. Thus, once a platelet has undergone release of its granule contents or inactivation of one or more of its enzymes (e.g., aspirin effect), it remains functionally impaired until it is removed from the circulation.

The Platelet Response to Vascular Injury

The typical platelet response to vascular injury can be divided into three major phases. These are: *platelet adhesion*, the interaction of platelets with nonplatelet surfaces; *platelet activation*, a phase during which biochemical pathways are activated, cytoplasmic calcium concentration increases, and platelets undergo

shape change and secrete granule constituents, including ADP; and *platelet aggregation*, the interaction of platelets with each other to form platelet aggregates.

TABLE 46.2. PLATELET α-GRANULE PROTEINS

Homologues of Plasma Proteins	Platelet-Specific Proteins	Platelet-Associated Proteins
Fibrinogen	Platelet factor 4	Thrombospondin
Factor V, VIII	ß-Thromboglobulin	Multimerin
von Willebrand factor		P-Selectin
High molecular weight Kininogen		Platelet-derived growth factor
Protein S		Endothelial-cell growth factor
Plasminogen		
Plasminogen activator inhibitor-1		
α_2-Antiplasmin		
α_1-Antiprotease		
α_2-Macroglobulin		
Fibronectin		
Vitronectin		
Osteonectin		
Albumin		
Factor D		
ß-1H-Globulin		

Platelet Adhesion

The adhesion of platelets to exposed subendothelial structures is mediated primarily by the interaction between vWF and its cellular receptor, glycoprotein Ib/IX/V (GP Ib/IX/V) (Fig. 46.3)(7,8,9,10,11 and 12). vWF is a very large glycoprotein that is normally released into the circulation as an unbranched polymer consisting of a variable number of homodimeric subunits (Fig. 46.4). vWF is synthesized by both endothelial cells and megakaryocytes; plasma vWF is primarily a product of endothelial cell synthesis and release, whereas platelet α-granule vWF is primarily a product of megakaryocyte synthesis. In addition, endothelial cells secrete vWF into the subendothelial cell space so that vWF is present in the subendothelial matrix itself. All three sources of vWF are necessary for an adequate platelet response to vascular injury.

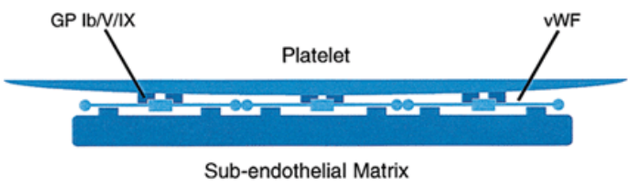


FIGURE 46.3. Platelet adhesion is a "molecular gluing" process mediated by von Willebrand factor (vWF) interaction with the platelet glycoprotein Ib/IX/V receptor complex (GP Ib/IX/V) and the subendothelial matrix. HMW multimers of vWF are necessary for effective platelet adhesion, perhaps because this maximizes the contact between platelet surfaces and the tissues to which the platelets adhere. Under normal flow conditions, vWF does not bind to GP Ib/IX/V. Apparently, binding of vWF to matrix materials such as collagen induces a conformational change in vWF that allows it to bind to GP Ib/IX/V.

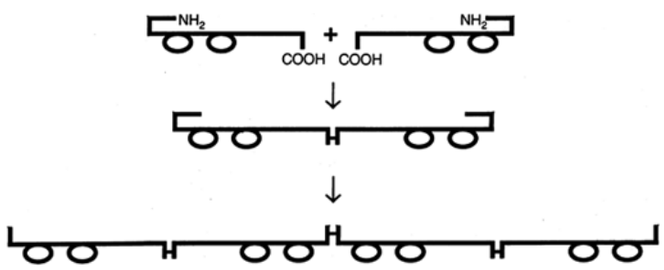


FIGURE 46.4. The synthesis of von Willebrand factor (vWF) is a complex, multi-step process in which 275 kd monomeric subunits are assembled into long, unbranched polymers. Monomeric subunits initially form dimers through the formation of disulfide linkages between the carboxyl terminal ends of monomers. The amino-terminal ends of dimers are then linked together by disulfide bonds. During the later stages of processing, a large peptide is removed from the amino-terminal end of the subunits; this fragment has been designated vWF antigen II. Only the large polymers (HMW multimers) are effective in mediating platelet adhesion.

The GP Ib/IX/V receptor complex is composed of four proteins: Ibα, Ibß, V and IX, each coded for by a distinct gene. Expression of the receptor complex on the platelet surface is dependent on coordinated synthesis of these protein components, as defects in either the Ibα, Ibß, or IX gene may lead to decreased expression of the receptor (Bernard-Soulier syndrome, see Chapter 47). The binding site for vWF is present in the N-terminal extracellular domain of GP Ibα. Intracellular domains of the receptor complex are associated with the platelet cytoskeleton and phospholipase A2. Consequently, ligand binding to the receptor is able to initiate biochemical signaling pathways.

Under normal conditions, vWF multimers do not bind to the platelet GP Ib/IX/V receptor. Current evidence suggests that a conformational change in vWF and/or GP Ib/IX/V is necessary for binding. Such a conformational change may be induced by interaction between vWF and subendothelial components such as collagen, high sheer stress associated with blood flow, or various exogenous agents such as ristocetin or certain venoms. Only the longer polymers [high-molecular-weight (HMW) multimers] of vWF are capable of effectively mediating platelet adhesion, perhaps because the larger forms maximize surface contact between the platelet and vessel wall. In addition, the HMW multimers of vWF may enhance clustering of GP Ib/IX/V receptors on the platelet surface, a step which may be critical for transmitting a ligand binding signal into the cytoplasm of the responding platelets.

Platelet Activation

Platelet activation is associated with a change in platelet shape, marked by extension of pseudopodia, secretion of dense and α-granule contents, activation of the ligand-binding site on GP IIb/IIIa, and translocation of negatively charged aminophospholipids to the external surface of the platelet(13,14,15,16,17 and 18). The cascade of these events is driven by a rapid increase in cytoplasmic calcium concentration. The rise in calcium is triggered by activation of biochemical response pathways resulting from interaction of platelet agonists such as thrombin, ADP, thromboxane A_2, collagen, epinephrine, and vWF with specific membrane receptors. Some of the membrane receptors, such as the thrombin receptor, are coupled to a G protein response mechanism that directly activates phospholipase Cγ while other receptors interact with the cytoskeleton or other response mechanisms that trigger

activation of other phospholipase C isoforms. For example, the GP Ib/IX/V receptor complex is linked to the platelet cytoskeleton and vWF binding induces clustering of these receptors, transmitting a signal to the cytoplasm and activating the biochemical response pathways.

Activation of phospholipase C by any one of a number of membrane signals leads to cleavage of phosphatidylinositol-bisphosphate (PIP_2) and release of two critical intracytoplasmic mediators, diacylglycerol (DG) and inositol trisphosphate (IP_3) (Fig. 46.5). Inositol trisphosphate causes a rapid release of calcium from its storage sites into the platelet cytoplasm, leading to the increase in cytoplasmic calcium. This activates calmodulin-dependent phosphokinases and leads directly to phosphorylation of myosin light chain. Diacylglycerol binds to and activates protein kinase C, which then phosphorylates a number of proteins, including pleckstrin, a 47-kd protein essential for propagating the activation signal. In addition, DG can lead to activation of the phospholipase A2 pathway, further amplifying the activation signal.

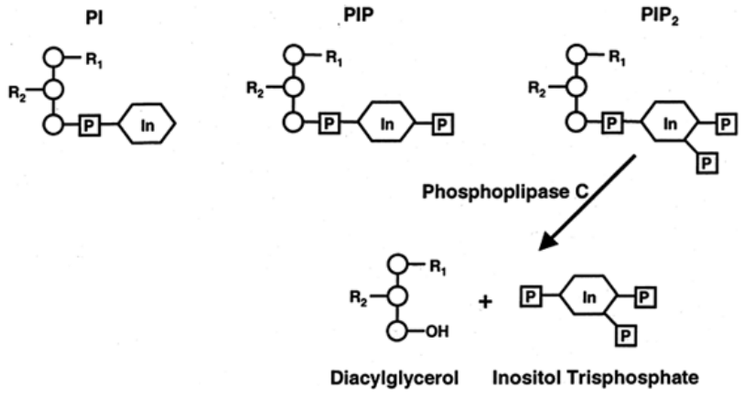


FIGURE 46.5. Platelet activation involves a number of biosynthetic pathways in the platelet. The phospholipase C (PL C) pathway is responsible for production of diacylglycerol (DG) and inositol trisphosphate (IP_3) from phosphatidylinositol bisphosphate (PIP_2). Diacylglycerol activates protein kinase C (PK C), leading to phosphorylation of pleckstrin and activation of a number of downstream pathways. IP_3 releases calcium from storage sites, leading to a sudden increase in cytoplasmic calcium concentration and activation of calmodulin-dependent protein kinases. This pathway leads to myosin light chain phosphorylation and platelet contraction.

Phospholipase A2 releases arachidonic acid from membrane phospholipids, particularly phosphatidylcholine (Fig. 46.6). Arachidonic acid can be converted into cyclic endoperoxides by cyclo-oxygenase or metabolized by 12-lipoxygenase. In the platelet, the cyclo-oxygenase pathway leads to formation of thromboxane A_2, another potent stimulator of phospholipase C. Aspirin and other nonsteroidal antiinflammatory agents have an antiplatelet effect because of their ability to inhibit cyclo-oxygenase and block formation of thromboxane A_2.

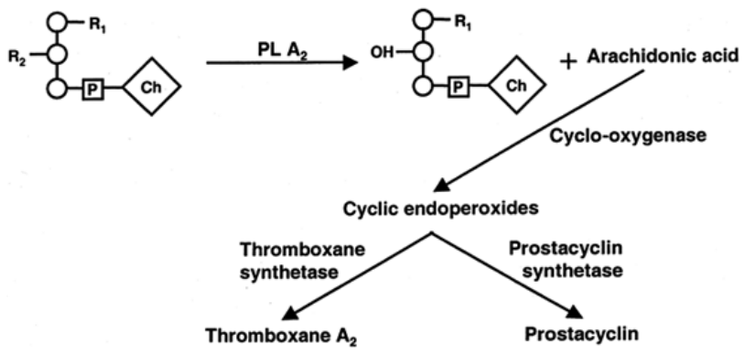


FIGURE 46.6. The phospholipase A2 (PL A2) pathway releases arachidonic acid (AA) from membrane phospholipids, particularly phosphatidylcholine. Arachidonic acid is then converted into cyclic endoperoxides by cyclo-oxygenase. In the platelets these cyclic endoperoxides are converted into thromboxane A_2 (TxA_2), a potent platelet agonist. In the vessel wall, the cyclic endoperoxides are converted into prostacyclin, a potent inhibitor of platelet function.

TABLE 46.3. CONTACT SYSTEM ENZYMES

Protein	Molecular Weight (kd)	Biological Half-Life (hr)	Activated By	Function
Factor XII (Hageman factor)	80	50	Kallikrein	Activates prekallikrein, factor XI
Prekallikrein (Fletcher factor)	85	50	Factor XIIa	Activates factor XI, Prourokinase, HMW Kininogen→bradykinin
Factor XI	160 (dimer)	40	Factor XIIa	Activates factor IX

The phosphatidylinositol-3 kinase (PI3-K) pathway leads to synthesis of intracellular lipid messengers, including phosphatidylinositol-3,4 bisphosphate and phosphatidylinositol-3,4,5 triphosphate (Fig. 46.7). These mediators interact with a variety of proteins bearing pleckstrinlike domains, initiating downstream activities, including exposure of the ligand-binding site on GP IIb/IIIa. Inhibition of the PI3-K pathway by specific inhibitors blocks a normal platelet aggregation response to many platelet agonists, indicating that this pathway plays an important role in modulating the signal pathways in platelets.

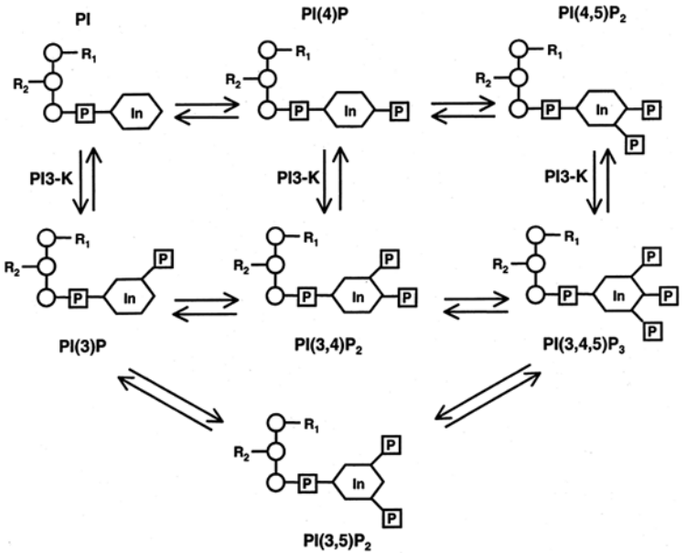


FIGURE 46.7. The phosphatidylinositol-3-kinase (PI3-K) pathway generates important intermediate messengers in the platelet activation pathway. The various phosphatidylinositol-3-PO4 species [PI(3,X)Px] bind to pleckstrinlike domains on responder molecules, modulating their activity. The PI3-K pathway is essential for platelet aggregation, as inhibition of PI3-K by specific inhibitors can block the aggregation response to certain platelet agonists.

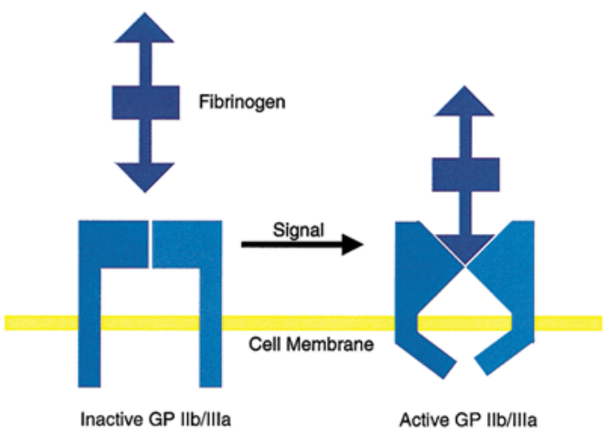


FIGURE 46.8. Platelet aggregation is mediated by the cross-linking of glycoprotein IIb/IIIa (GP IIb/IIIa) on adjacent platelet membranes. A number of adhesive proteins may mediate this process, but fibrinogen is the major mediator of this process *in vivo*. The structural symmetry of fibrinogen makes it an excellent molecule for this process. Under resting conditions, the binding site on GP IIb/IIIa for adhesive proteins is not exposed. Upon platelet activation, GP IIb/IIIa undergoes a conformational change, allowing binding of fibrinogen or other adhesive proteins. The conformational change in GP IIb/IIIa may be initiated by changes in the intracytoplasmic domains of IIb and/or IIIa (inside-out signaling) or by extracellular factors. Binding of fibrinogen to GP IIb/IIIa also induces a change in the cytoplasmic domain of the molecules, so that extracellular activation of the receptor and subsequent ligand binding can also be a signaling process (outside-in signaling).

One of the principal results of activation of these various biochemical pathways is exposure of the ligand-binding site on GP IIb/IIIa, the receptor that mediates the final stage of the platelet response, platelet aggregation. In its nonactivated form, GP IIb/IIIa cannot bind to its ligands and thus cannot participate in cross-linking of platelets. However, platelet activation is associated with a conformational change in GP IIb/IIIa that permits ligand binding. Activation of the ligand binding sites on GP IIb/IIIa may occur through biochemical modification of the cytoplasmic component of the complex (inside-out-signaling) or interaction of GP IIb/IIIa with extracelllular factors. As with platelet adhesion, ligand binding to the activated GP IIb/IIIa receptor can also activate other platelet response pathways (outside-in-signaling).

Platelet activation also leads to secretion of the contents of the dense and α-granules. The secreted substances propagate the platelet aggregation response and accelerate thrombin formation by the coagulation cascade. Another critical event associated with platelet activation is the translocation of phosphatidylserine from the inner surface of the platelet membrane to the exterior surface. This process is dependent on a “scramblase” enzyme that is activated by the increase in cytoplasmic calcium concentration. Expression of phosphatidylserine on the external surface facilitates the formation of coagulation activation complexes that mediate thrombin formation. Translocation of phosphatidylserine also is necessary for the formation of platelet microparticles associated with platelet activation.

Platelet Aggregation

Platelet aggregation is another example of “molecular gluing.” The activated form of GP IIb/IIIa can be cross-linked by a variety of adhesive proteins bearing the amino-acid sequence RGD (19). However, the major protein involved in cross-linking platelets *in vivo* is fibrinogen. Intact fibrinogen is a dimeric molecule that allows reciprocal interaction of its terminal ends (D domains) with GP IIb/IIIa receptors on adjacent platelets once the receptor has been activated (Fig. 46.8). Thrombospondin, a major component of platelet α-granules, appears to stabilize and augment the platelet-platelet interaction mediated by fibrinogen. Ligand binding to GP IIb/IIIa activates biochemical response pathways within the platelet cytoplasm, a process referred to as “outside-in signaling.” Under conditions of high shear stress, vWF also can bind directly to the GP IIb/IIIa complex, initiating a platelet aggregation response through the response pathways.

In summary, the primary hemostatic response consists of three major stages. Platelet adhesion is the first step and is mediated by vWF interaction with GP Ib/IX/V and subendothelial receptors. The subsequent shape change and secretion are dependent on activation of multiple biochemical pathways within the platelet resulting in a rapid increase in cytoplasmic calcium. Activation of the platelet leads to expression of the ligand binding site on GP IIb/IIIa and cross-linking by intact fibrinogen. Defects in various platelet and plasma components of this response mechanism have been described and constitute many of the platelet disorders described in Chapter 47. In addition, components of these response pathways have been the target of antithrombotic drugs aimed at inhibiting the platelet component of the hemostatic response.

SECONDARY HEMOSTASIS

Part of "46 - Overview of Hemostasis"

Secondary hemostasis is marked by the conversion of a soluble protein, fibrinogen, into an insoluble gel. Formation of the fibrin meshwork results in reinforcement and stabilization of the platelet plug at sites of vascular disruption. The process of fibrin clot formation can be divided into two major phases; the first phase consists of formation of the potent procoagulant enzyme thrombin; the second phase consists of the conversion of fibrinogen into fibrin with subsequent polymerization and stabilization. The first phase takes place on negatively charged phospholipid surfaces while the second phase occurs in the fluid phase.

Thrombin is formed as a result of multiple enzymatic steps, collectively known as the coagulation cascade. Most of the enzymatic steps involve formation of an “enzyme activation complex” composed of an active enzyme, a proenzyme/substrate protein, a catalytic cofactor, and a negatively charged phospholipid surface (activated platelet with translocated phosphatidylserine) that serves to localize the process and concentrate the reactants (Fig. 46.9). The coagulation cascade can be viewed as the formation of a series of these key activation complexes. As each enzyme activates many substrate proenzymes, each activation complex represents an important amplification step. The result of this amplification is that a small initiating signal can be transformed into a high concentration of thrombin at the site of clot formation.

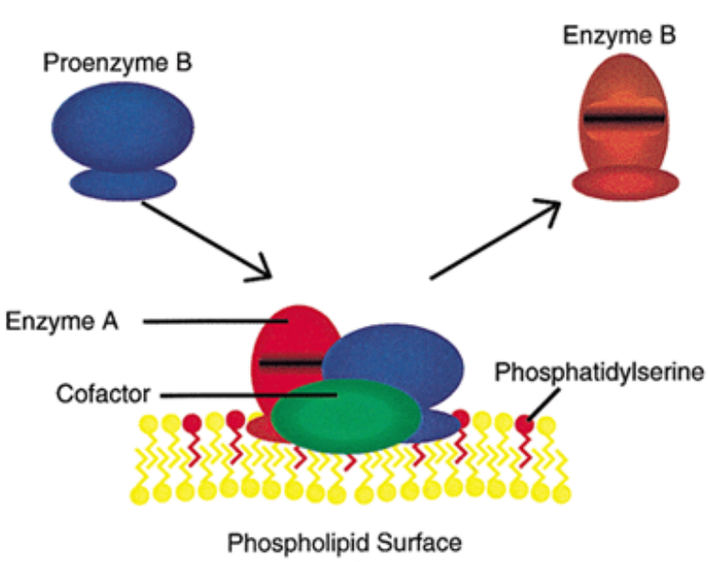


FIGURE 46.9. The basis for most reactions involving the coagulation cascade is the formation of an enzyme activation complex composed of an active enzyme (serine protease); a substrate that is usually the zymogen (proenzyme) of another active enzyme; a catalytic cofactor protein that serves to localize the enzyme and substrate to a phospholipid surface; and a negatively charged phospholipid surface. Thrombin is formed by such a complex in which the active enzyme is factor Xa, the substrate is prothrombin, the cofactor is factor Va, and the phospholipid is provided by a cell surface (e.g., activated platelet). As each enzyme/cofactor is capable of activating many substrate/proenzyme molecules, the formation of these activation complexes allows rapid amplification of the original activation signal.

Coagulation Proteins

The enzymes of the coagulation cascade are serine proteases; that is, they are enzymes that split specific peptide bonds in selected

proteins through formation of unstable intermediate bonds between the target peptide bond and the reactive serine at the enzyme's active site (20,21). Coagulation serine proteases can be divided into two families, the contact system enzymes and the vitamin K-dependent proteases (Table 46.3 and Table 46.4). The contact system serine proteases share several characteristics: the molecular weight of each monomeric unit is 80 to 85 kd; (*b*) they do not require vitamin K for complete synthesis; (*c*) they are not adsorbed by $BaSO_4$ or $Al(OH)_3$; (*d*) they are not dependent on ionic calcium for activation and (*e*) they are not completely consumed during *in vitro* coagulation of blood. Factor XI is unique among these serine proteases in that it circulates as a homodimer linked by disulfide bonds and it is the only protein of this family that is associated with a bleeding tendency when the factor is absent.

TABLE 46.4. VITAMIN K DEPENDENT PROTEINS

Protein	Molecular Weight (kd)	Gla Residues	Biological Half-Life (h)	Activated By	Function
Factor II (prothrombin)	72	10	60	Factor Xa	Cleaves fibrinogen, activates V, VIII, XIII, protein C, platelets
Factor VII	48	10	4-6	Factors IIa, Xa	Activates factors IX, X
Factor IX	57	12	25	Factors XIa, VIIa	Activates factor X
Factor X	55	11	24	Factors VIIa, IXa	Activates factor II
Protein C	62	9	6	Thrombin	Inactivates factors Va and VIIIa
Protein S	69	10	60	Not activated	No enzymatic activity; cofactor for activated protein C

The vitamin K dependent factors are all characterized by the presence of 9-12 γ carboxyglutamic acid residues near the amino terminal end of the protein. Vitamin K is necessary for the post-transcriptional carboxylation of these glutamic acid residues (Fig. 46.10). In the absence of γ-carboxylation, these proteins do not bind calcium and do not interact with phospholipid surfaces. Such descarboxy forms, which are found in vitamin K deficiency, oral anticoagulant therapy and liver disease, are ineffective at mediating secondary hemostasis. Vitamin K dependent proteins are heat stable and, with the exception of prothrombin, are not entirely consumed during coagulation *in vitro*. Some of the proteins (protein C, protein S, factor IX, and factor X) are characterized by an additional post-transcriptional modification, the hydroxylation of an aspartic acid or asparagine in epidermal growth factorlike domains to form ß-hydroxyaspartic acid or ß-hydroxyasparagine. This hydroxylation step creates additional calcium binding sites, which are necessary for normal function of the proteins. The catalytic domain necessary for trypsinlike enzyme activity is found at the carboxyl-terminal end of the vitamin K dependent proteases. Protein S differs from the other vitamin K dependent coagulation proteins in that it has no catalytic domain at the carboxyl terminus. In contrast to the proteolytic activity of the other vitamin K dependent factors, protein S functions as a cofactor, accelerating the activity of activated protein C (APC).

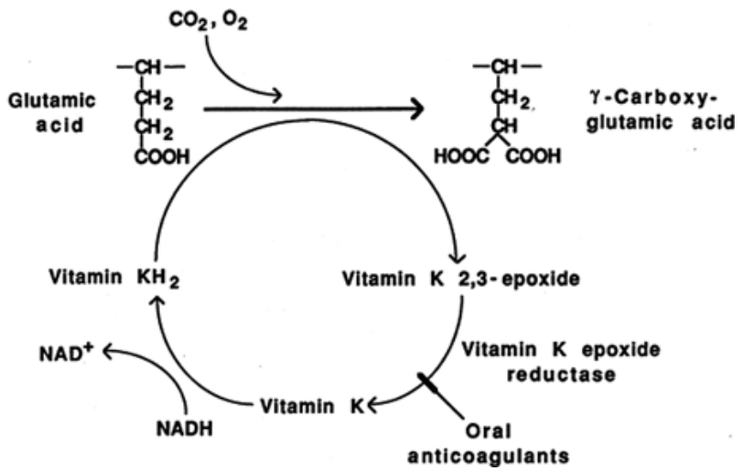


FIGURE 46.10. Vitamin K is necessary for the post-translational modification of glutamic acid residues located near the amino terminal end of the vitamin K dependent proteins. In the presence of reduced vitamin K, these residues are converted to the γ-carboxyglutamic acid residues necessary for calcium binding and interaction with negatively charged phospholipid surfaces. The vitamin K epoxide formed in this process is then metabolized to regenerate the reduced form of vitamin K. Oral anticoagulants inhibit the regeneration of reduced vitamin K, trapping it in the inactive epoxide form.

The remaining coagulation proteins can be grouped into a family designated as thrombin-sensitive proteins and cofactors (Table 46.5). These proteins are all large, with molecular weights greater than 110 kd. Factor V, factor VIII, and HMW kininogen function as catalytic cofactors in enzyme activation complexes, while fibrin functions as a cofactor for the activation of plasminogen by tissue plasminogen activator (tPA). Factor XIII is the only enzyme in this family; it differs from other coagulation enzymes in that it is a transpeptidase and is responsible for cross-linking and stabilizing the fibrin clot. HMW kininogen is the only one of these proteins that is not activated by thrombin.

TABLE 46.5. THROMBIN-SENSITIVE PROTEINS AND COFACTORS

Protein	Molecular Weight (kd)	Biological Half-Life	Activated By	Function
Fibrinogen	340	4-5 days	Thrombin	Fibrin clot, cofactor for activation of plasmin
Factor V	330	12-36 h	Thrombin	Cofactor for activation of prothrombin
Factor VIII	240	12 h	Thrombin	Cofactor for activation of factor X
High molecular kininogen (Fitzgerald factor)	110	6-7 days	Not activated	Cofactor for activation of factor XI, prekallikrein; source of bradykinin
Factor XIII	320	3-5 days	Thrombin	Transpeptidase that stabilizes fibrin clot

Thrombin Formation

There are at least two mechanisms for initiating thrombin formation; these have been designated the intrinsic and extrinsic pathways (22,23,24 and 25). Their names derive from the fact that all components of the intrinsic system are found in circulating blood, while the extrinsic pathway requires the presence of tissue factor, a component that is normally extrinsic to circulating blood. The extrinsic pathway is the major physiological mechanism for initiating the coagulation cascade and generating thrombin for fibrin clot formation. Activation of the intrinsic pathway is not necessary for generation of fibrin clots, but may play a role in fibrinolysis and inflammatory response mechanisms and may be involved in pathological activation of the coagulation cascade. The two pathways correspond to two commonly used tests to screen the coagulation system, the prothrombin time (PT), which measures the extrinsic pathway,

and the activated partial thromboplastin time (APTT), which measures the intrinsic pathway.

Extrinsic Coagulation Pathway

The extrinsic pathway begins with exposure of blood to tissue factor (Fig. 46.11). In the presence of tissue factor and calcium, VIIa activates both factor IX and factor X. Just where the initial factor VIIa comes from is a matter of continuing debate. However, it appears that the coagulation system is dynamically modulated with a constant low level of activation. It is likely that there is always a small amount of factor VIIa around, but in the absence of tissue factor it is incapable of initiating an effective response. Factor IXa, in the presence of factor VIIIa and phospholipid, also converts factor X to Xa. Factor Xa then converts prothrombin to thrombin in the presence of factor Va and phospholipid. The prothrombinase complex (Xa/Va/II/phospholipid) is often designated the “final common pathway” as it is common to both the intrinsic and extrinsic pathways.

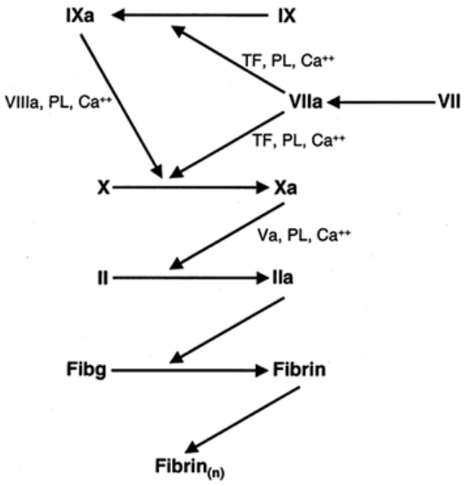


FIGURE 46.11. The coagulation cascade may be initiated by one of two mechanisms; the contact system, that activates the intrinsic system, and the tissue factor pathway (shown here), which initiates the extrinsic system. The extrinsic pathway is activated whenever blood is exposed to tissue factor (TF), an ubiquitous membrane protein found on most cells except circulating blood cells and quiescent endothelial cells. Factor VIIa binds to TF to form an “enzyme activation complex” that converts factor IX to IXa and factor X to Xa. Factor IXa interacts with factor VIIIa to form an additional activation complex that converts factor X to Xa. Factor Xa then interacts with factor Va to form the prothrombinase complex that converts prothrombin into thrombin. Thrombin then converts fibrinogen into fibrin, which spontaneously polymerizes if the concentration of monomers is sufficiently high. Both factors V and VIII must be activated by thrombin before they express effective procoagulant activity (Fig. 46.12). Therefore, initial thrombin generation is slow until there is sufficient thrombin formed to activate these cofactors.

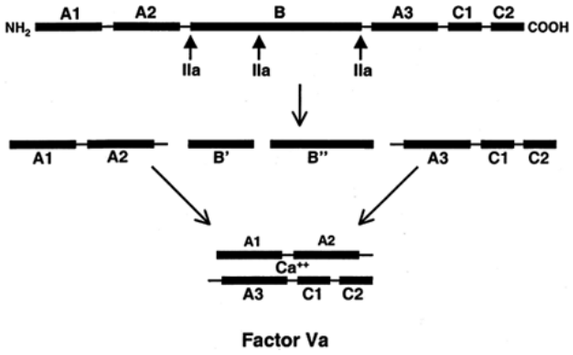


FIGURE 46.12. Factors V and VIII are critical cofactors for the generation of thrombin. These two factors are quite similar in structure, with each composed of 3 homologous “A” domains, a large, heavily glycosylated “B” domain and two homologous “C” domains. These domains are organized into the sequence A1·A2·B·A3·C1·C2. Thrombin activates factors V (shown) and VIII by splitting peptide bonds that remove the large B domain. The A1·A2 fragment (heavy chain) and the A3·C1·C2 fragment form a noncovalently linked, calcium dependent complex that functions as the catalytic cofactor. Factors Va and VIIIa are degraded by activated protein C-mediated cleavage of peptide bonds in the A1·A2 domain.

The role of factors VIIIa and Va is to accelerate the interaction between active enzymes and their target proenzymes. Both factors V and VIII must be activated by thrombin for expression of catalytic cofactor activity (Fig. 46.12). Therefore, the rate of thrombin generation is initially very slow until sufficient thrombin to activate factors V and VIII is formed. The rate of thrombin formation rises exponentially once these cofactors are activated, leading to explosive generation of thrombin at the site of clot formation. The rate of thrombin formation is significantly reduced in the absence of these critical cofactors.

Intrinsic Coagulation Pathway

Classically, activation of the intrinsic system is thought to occur when factor XII binds to negatively charged surfaces, resulting in a conformational change in the molecule and expression of factor XIIa enzymatic activity (Fig. 46.13). Factor XIIa can activate both prekallikrein and factor XI; HMW kininogen serves as a cofactor for both of these activation steps. Indeed, prekallikrein and factor XI each bind to HMW kininogen in circulating plasma. The kallikrein formed by surface-activated XIIa is a potent activator of factor XII and thus the activation of prekallikrein initiates a positive feedback loop, permitting generation of large quantities of factor XIIa and increasing concentrations of factor XIa. Calcium is not required during these initial steps involving vitamin K independent proteins.

More recently, it has been shown that prekallikrein/HMW kininogen complex can bind to specific multi-protein receptors on cell surfaces, including endothelial cells. Cell-associated prekallikrein can be converted to kallikrein independent of factor XII. The kallikrein thus formed can generate bradykinin and activate factor XII and single chain prourokinase. Activation of factor XII can then lead to amplification of the pathway (Fig. 46.13). In addition, alternate mechanisms for the activation of factor XI have been described. Factor XI and XIa can bind avidly to the surface of activated platelets. This binding may be augmented in the presence of low concentrations of thrombin. On the activated platelet surface both thrombin and factor XIa can directly activate factor XI to factor XIa. These cell surface-based mechanisms for activation of the contact system of coagulation bypass the classic factor XII-mediated mechanism proposed for activation of the intrinsic pathway.

Once formed, factor XIa activates factor IX in the presence of calcium and a negatively charged phospholipid surface, particularly activated platelet membranes (Fig. 46.13). This is the only enzymatic step in thrombin formation for which there is no known protein cofactor. Factor IXa activates factor X in the presence of VIIIa, calcium, and phospholipid. Factor Xa then converts prothrombin to thrombin in the presence of factor Va, calcium, and phospholipid. Intrinsic coagulation thus involves two enzyme activation complexes — the factor IXa/VIIIa/X/phospholipid complex and the factor Xa/Va/prothrombin/phospholipid complex — that also are involved in the extrinsic pathway.

The key product of these pathways is thrombin, a potent enzyme that serves several key roles in hemostasis. These include conversion of fibrinogen to fibrin, activation of factors V and VIII, activation of factor XIII, activation of protein C, and activation of platelets. In addition, thrombin can interact and stimulate a variety of other cells, including endothelial cells. For example, interaction of thrombin with endothelial cells can lead to fundamental changes in the endothelial cell surface that increase

procoagulant activity and promote thrombosis. Critical to these functions of thrombin is the fact that conversion of prothrombin is accompanied by separation of the amino terminal, phospholipid-binding portion of the protein from its catalytic (enzymatic) domain. The large amino terminal fragment that is released is referred to as prothrombin fragment 1.2 ($F_{1.2}$). With release of $F_{1.2}$, thrombin becomes a soluble enzyme in contrast to the other vitamin K-dependent enzymes, which are only effective when associated with phospholipid surfaces.

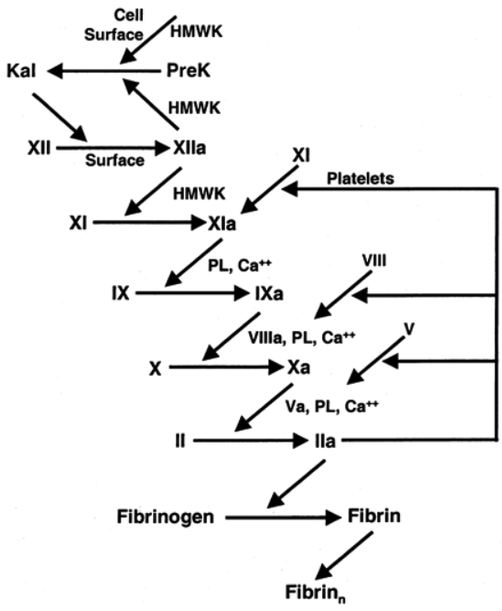


FIGURE 46.13. Classically, the intrinsic system is thought to begin when factor XII is activated by contact with appropriate negatively charged surfaces. Activated factor XII (XIIa) then activates prekallikrein (PreK) and factor XI, initiating the enzymatic cascade resulting in thrombin formation. Kallikrein (Kal), in turn, activates factor XII, resulting in a positive feedback loop in which a high concentration of factor XIIa and XIa can be generated. Factor XIa converts factor IX to IXa and the pathway then follows the reactions of the extrinsic pathway. More recently, it has been shown that there are at least two other mechanisms for initiating the intrinsic system. First, the HMW kininogen (HMWK)/prekallikrein complex that is present in plasma can bind to specific multimolecular cell receptors where prekallikrein is converted to kallikrein. Kallikrein then can activate factor XII, initiating the positive feedback loop of contact activation. Second, thrombin can directly activate factor XI on platelet surfaces, bypassing the need for factor XII and kallikrein. As with the extrinsic pathway, thrombin must activate factors V and VIII before they are effective cofactors.

Fibrin Formation

Fibrinogen is a large homodimeric protein having a molecular weight of 340 kd. It is composed of six peptide chains: two Aα, two Bß, and two γ chains (26). Each half of the fibrinogen molecule is composed of one Aα, one Bß, and one γ chain linked together by extensive disulfide bonds that are concentrated in the amino and carboxyl-terminal ends. The extensive cross-linking gives these regions of the molecule a globular structure, while the intervening sequences are arranged in a long helical structure. The two halves are linked together by reciprocal disulfide bonds between the amino terminal ends of each γ chain (Fig. 46.14).

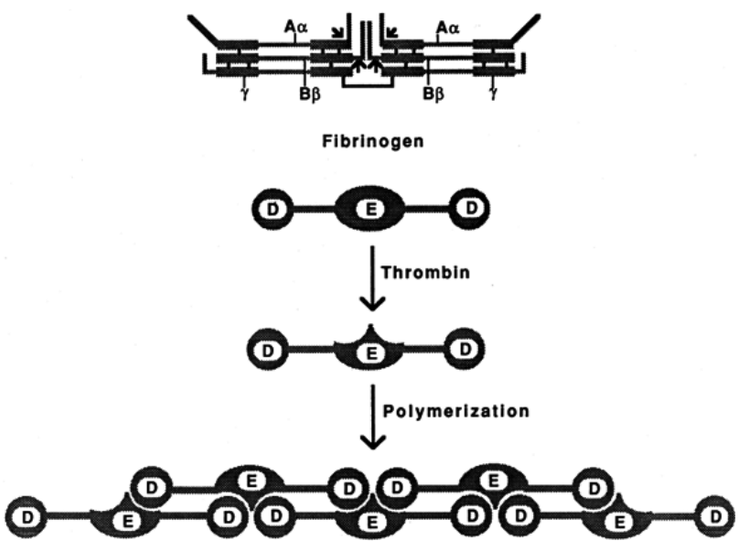


FIGURE 46.14. Fibrinogen is a large dimeric protein composed of six peptide chains. Each half of the dimer is composed of an Aα, a Bß, and a γ chain. There are extensive interchain disulfide bonds near the amino termini, resulting in a central nodular region known as the E domain. Interchain disulfide bonds at the carboxyl termini also result in nodular regions, known as the D domains. A fibrinogen molecule thus consists of a central E domain linked by helical segments to two D domains. Thrombin removes a short peptide from each Aα and Bß chain (fibrinopeptides A and B), altering the conformation of the central E domain. This exposes binding sites for D domains within the central E domain and permits polymerization of the fibrin monomers to form the fibrin gel.

Thrombin cleaves a 16 amino-acid peptide (fibrinopeptide A) from the amino terminal end of each Aα chain and a 14 amino acid peptide (fibrinopeptide B) from each amino terminal end of the Bß chains. Loss of these peptides from the central (E) domain of the fibrinogen molecule alters the charge distribution and exposes binding sites for the carboxyl-terminal ends of fibrin, known as the D domains. As each E domain binds two D domains, the fibrin monomers polymerize in a staggered overlap, with D domains of adjacent fibrin monomers interacting with a central E domain of a third fibrin monomer (Fig. 46.14). Lateral association then leads to gelation of the fibrin. Formation of the fibrin gel is dependent on having a high concentration of fibrin monomers; the gel does not form if the fibrin monomers are present in low concentration. As the gel forms, a large number of proteins are entrapped in the gel matrix, including thrombin, factor XIII, plasminogen, tPA, plasminogen activator inhibitor-

1 (PAI-1) and α_2-antiplasmin. The relative activity of these incorporated proteins ultimately determines the fate of the fibrin clot.

The fibrin gel is stabilized by factor XIIIa, which covalently cross-links fibrin monomers. Plasma factor XIII is a large tetrameric molecule composed of two A and two B subunits. The A and B subunits are under separate genetic control. The B subunit serves as a plasma carrier protein for the A subunit, which contains the catalytic domain. Factor XIII is activated by thrombin, but full enzyme activity is expressed only after the A subunit

separates from the B subunit, a process that occurs within the fibrin matrix. Thrombin activation of factor XIII is accelerated by the presence of non–cross-linked fibrin but is inhibited by fully cross-linked fibrin. This autoregulation serves to help limit factor XIIIa activity to sites of fibrin clot formation. Factor XIIIa covalently links glutamine to lysine amino acids. The principal sites of cross-linkage are between D domains of adjacent fibrin monomers and between the carboxyl-terminal ends of the α chains of distant fibrin monomers (Fig 46.15). Covalent cross-linkage greatly stabilizes the fibrin clot and makes it resistant to fibrinolysis. In addition to cross-linking and stabilizing the fibrin clot, factor XIIIa can also catalyze the formation of covalent bonds between fibrin and other proteins incorporated in the clot matrix, including α_2-antiplasmin.

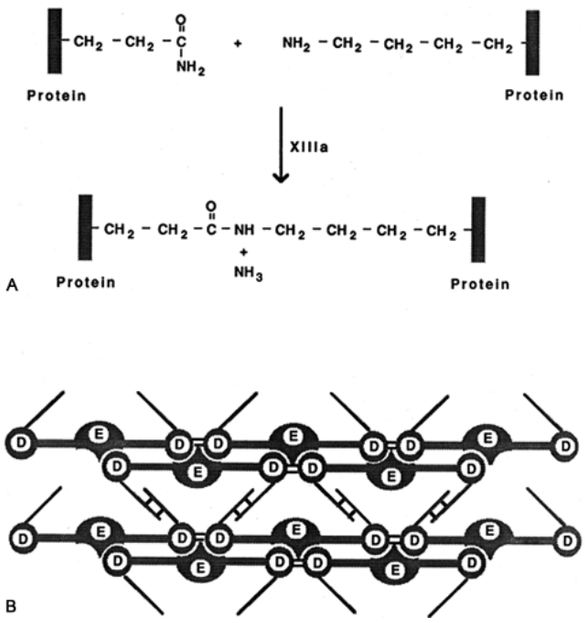


FIGURE 46.15. Factor XIIIa is a transpeptidase that cross-links peptide chains by inducing covalent linkage between glutamine and lysine amino acids (**A**). There are multiple cross-linkage sites in fibrin monomers; some are located near the carboxyl termini of γ chains and result in cross-linkage of adjacent D domains, while others are located in the carboxyl termini of γ chains and result in cross-linkage of distant fibrin monomers. The net effect is an extensively cross-linked fibrin gel that is resistant to fibrinolysis.

REGULATION OF HEMOSTASIS

Part of "46 - Overview of Hemostasis"

The procoagulant mechanisms described in the previous sections lead to rapid formation of hemostatic plugs at sites of vascular damage. Under normal circumstances, coagulation is limited to sites of vascular disruption and is not propagated throughout the vessel. However, if the procoagulant mechanisms are not held in check, thrombus formation can propagate within the vessel, leading to clinically significant occlusive thrombi. There are multiple systems that work in a synergistic manner to regulate the extent of clot formation. An underlying theme to these various systems is the critical role of intact endothelium. Although the endothelium was once thought of as a relatively inert barrier between the circulating blood and the procoagulant extravascular tissues, it is now clear that endothelial cells play an active role in the hemostatic response. Furthermore, it is becoming apparent that endothelial cell function can vary in important ways between different tissues and in response to various inflammatory mediators. Thus the relative importance of the various regulatory systems described below may differ depending on the tissue in which hemostatic activation occurs or the underlying disease process (27,28 and 29).

TABLE 46.6. REGULATION OF PLATELET RESPONSE

Component	Action	Effect
Prostacyclin	Increases cyclic AMP	Inhibits aggregation
Nitric oxide	Increases cyclic GMP	Inhibits aggregation and adhesion
ADPase	Metabolizes ADP	Inhibits aggregation
Endothelial glyocalyx	Electrostatic repulsion	Inhibits platelet adhesion

Regulation of the Platelet Response in Hemostasis

A number of factors that contribute to the regulation of the platelet response have been identified and most are related to intact endothelial cells (Table 46.6) (30,31). One of the most important inhibitors of the platelet response is prostacyclin, which is formed by metabolism of arachidonic acid through the prostaglandin pathway in endothelial and vascular wall cells (Fig. 46.6). Prostacyclin is a short-lived, but potent inhibitor of platelet aggregation. It stimulates platelet adenylate cyclase, which leads to an increase in platelet cytoplasmic cyclic AMP concentration. Cyclic AMP activates Protein Kinase A, which in turn phosphorylates target proteins, leading to inhibition of the phospholipase C pathway and inhibition of the release of calcium from cytoplasmic storage sites.

TABLE 46.7. REGULATION OF FIBRIN CLOT FORMATION

Component	Action	Effect
Tissue factor pathway inhibitor	Inhibits tissue factor/VIIa	Inhibits thrombin formation
Serine protease inhibitors	Neutralize thrombin, factor Xa	Inhibit thrombin formation and activity
Protein C system	Degradation of factors Va and VIIIa	Inhibits thrombin formation
Fibrinolytic system	Degradation of fibrin	Removes excess fibrin clot

Nitric oxide (NO), also known as endothelial-derived relaxing factor, is a short-lived metabolite that has profound effects on platelet responsiveness to vascular injury. NO is a potent vasodilator that causes an increase in platelet cytoplasmic cyclic GMP levels. The increase in cyclic GMP activates Protein Kinase G, which in turn phosphorylates target proteins resulting in inhibition of both platelet adhesion and aggregation. Increased levels of cyclic GMP also lead to inhibition of the phosphodiesterase that degrades cyclic AMP. Consequently, the NO-induced increase in cyclic GMP leads to increased levels of cyclic AMP. In a reciprocal fashion, increased levels of cyclic AMP lead to inhibition of the phosphodiesterase that degrades cyclic GMP. This reciprocal effect on the concentration of these cyclic nucleotides leads to a synergism between the effects of NO and prostacyclin; thus, the low basal levels of these two factors may be sufficient to minimize platelet activation under resting conditions.

There are multiple forms of NO synthetase, including constitutive and inducible isoenzymes. The activity of the constitutive enzyme usually is regulated by calcium concentration and thus is limited to periods of cellular stimulation associated with transient increases in cytoplasmic calcium. The inducible isoform of the enzyme is not dependent on increased calcium concentration for activity and thus can provide a more persistent supply of NO. A number of inflammatory mediators can stimulate production of the inducible NO synthetase and thus NO is an important inflammatory mediator in the sepsis syndrome. Platelets contain a constitutive NO synthetase that can generate low levels of NO. This may be an important mechanism of modulating platelet responsiveness in the circulation. Endogenous platelet NO synthesis can be modulated by the interaction between apolipoprotein E and the apolipoprotein E receptor 2, a transmembrane signaling receptor found on platelet surfaces.

Endothelial cells also express an ADPase (apyrase) (CD39) on their cell surface. The function of ADPase is to convert ADP into inactive metabolites. ADP is essential for platelet aggregation and exposure of the GP IIb/IIIa ligand binding site. Degradation of ADP thus down-regulates propagation of platelet aggregation. Another barrier to platelet activation is the glycocalyx on the surface of intact endothelial cells. The glycocalyx has a negative charge and tends to repulse platelets under resting conditions. In addition, intact endothelial cells serve as a barrier between circulating platelets and subendothelial vWF. Thus, platelet adhesion is prevented in areas where the endothelium remains intact. Under normal circumstances, these components work in concert to limit platelet activation to sites of endothelial disruption.

TABLE 46.8. SERINE PROTEASE INHIBITORS (SERPINS)

Serpin	Target Enzymes
Antithrombin	Factor Xa, thrombin
Heparin cofactor II	Thrombin, chymotrypsin
α_2-Antiplasmin	Plasmin
Protein C inhibitor	Activated protein C, tissue plasminogen activator
Plasminogen activator inhibitor-1	Tissue plasminogen activator, urokinase, activated protein C
C-1-Esterase inhibitor	C-1-Esterase, factor XIIa, kallikrein
α_1-Antiprotease	Neutrophil elastase, trypsin

Regulation of Fibrin Clot Formation

At least four systems have been identified as having a role in the regulation of fibrin clot formation (Table 46.7). Each of these

component mechanisms appears to be critical for effective regulation of hemostasis, and isolated deficiency of components of these systems has been associated with clinical manifestations. These four systems work synergistically to regulate fibrin formation and stabilization.

Tissue Factor Pathway Inhibitor

Tissue factor pathway inhibitor (TFPI) is a 40-kd glycoprotein that has also been called lipid-associated coagulation inhibitor (LACI) and extrinsic pathway inhibitor (EPI) (32). Tissue factor pathway inhibitor is a Kunitz-type inhibitor with three Kunitz inhibitor domains. It rapidly combines with free factor Xa to form a heterodimeric complex; this interaction is further accelerated in the presence of heparin. In the process, TFPI Kunitz domain 2 binds to the active site of factor Xa, inhibiting its enzymatic activity. The TFPI/Xa complex then binds to the phospholipid membrane-associated tissue factor/factor VIIa complex. The Kunitz domain 1 on TFPI binds to the active site of factor VIIa, inhibiting VIIa enzymatic activity (Fig. 46.16). The γ-carboxyglutamic acid domain of factor Xa is necessary for the TFPI/Xa complex to bind to the phospholipid surface. The TFPI/Xa complex thus is responsible for inhibition of the major physiological initiator of hemostasis, the tissue factor/VIIa complex.

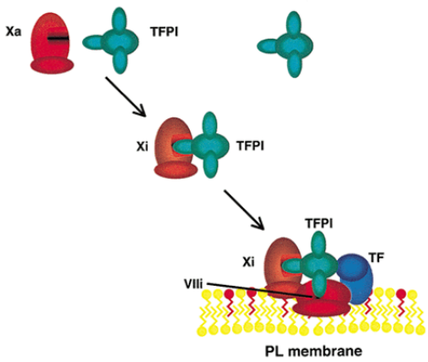


FIGURE 46.16. Tissue factor pathway inhibitor (TFPI) is a Kunitz-type inhibitor that initially forms a 1:1 complex with factor Xa (Xa). Binding of TFPI to Xa converts Xa into an inactive enzyme (Xi). The TFPI-Xi complex then binds to the tissue factor (TF)/factor VIIa(VIIa)-phospholipid complex, inhibiting the activity of TF/VIIa. This provides a mechanism for turning off the tissue factor pathway of initiating the coagulation cascade.

The vascular distribution of TFPI is complex. The majority (75% to 90%) of TFPI is bound to endothelial cells. It can be reversibly displaced into plasma by administration of exogenous heparin. Indeed, one of the potential antithrombotic mechanisms of heparin is its ability to increase the plasma level of TFPI and accelerate its interaction with factor Xa. Most of the circulating TFPI is found bound to plasma lipoproteins; this interaction is mediated by Kunitz domain 3 on TFPI. Only a small portion of total vascular TFPI (1% to 3%) actually is present in plasma as free TFPI. The distribution of TFPI into these various compartments has made accurate assessment of TFPI levels in clinical samples difficult.

Serine Protease Inhibitors

The **se**rine **p**rotease **in**hibitors (serpins) are substrates for specific serine proteases that form stable complexes with the enzymes following cleavage of a target peptide bond (33,34,35,36,37 and 38). In contrast to usual substrate proteins, cleavage of the target peptide bond results in a refolding of the serpin into a stable conformation, trapping the serine protease in the complex (Fig. 46.17). The enzyme-inhibitor complex is removed from the circulation after binding to specific receptors in the liver. Some serpins have a “latent” or inactive conformation in which the target site is not expressed. These serpins need to be “activated” (e.g., by binding to heparin in the case on antithrombin) before they can effectively inhibit their target serine protease. Cleavage of the reactive loop at other sites by nontarget proteases can induce a conformational change without trapping the protease. The conformational change stabilizes the serpin so that is no longer capable of exposing the reactive loop and inhibiting its target serine protease. The serpin family includes a large number of proteins that regulate several key response pathways (Table 46.8). A characteristic of the serpins is that, while they can inhibit many different serine proteases, they usually show a preference for only one or a few select enzymes.

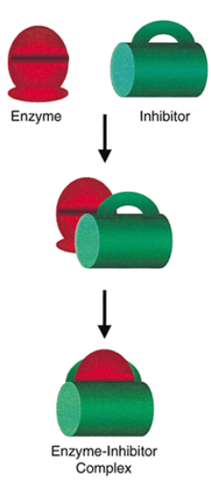


FIGURE 46.17. The general mechanism of action of serine protease inhibitors (serpins) is shown schematically. The inhibitors function as pseudosubstrates for the enzyme, which attacks a susceptible peptide bond in a reactive loop of the inhibitor. In a typical proteolytic reaction, the target protein separates from the enzyme following cleavage of the susceptible peptide bond. In the case of a serpin, cleavage of the susceptible bond in the reactive loop allows the serpin to refold into a stable conformation that traps the serine protease in a stable complex. The enzyme-inhibitor complexes then are rapidly removed from the circulation following interaction with a specific receptor in the liver. The reactive loop of some serpins is exposed under normal conditions; such inhibitors are able to rapidly inhibit their target serine protease. The reactive loop of other serine proteases appears to be in a latent conformation under resting conditions. For example, antithrombin requires interaction with heparin to expose the reactive loop before it can effectively interact with its target enzymes, factor Xa and thrombin. Cleavage of the serpin at sites away from the target peptide may alter the conformation of the reactive loop so that the serpin is no longer capable of inhibiting its target proteases. Such proteolytic cleavage basically inactivates the serpin.

The major serpins involved in the regulation of thrombin formation are antithrombin (AT) and heparin cofactor II. Antithrombin appears to be the most important of these, as heterozygous deficiency of AT is clearly associated with an increased risk of thrombosis (Chapter 48). Antithrombin is capable of neutralizing purified thrombin and factor Xa in the absence of a catalyst, but the rate of this reaction is too slow to account for its ability to regulate these coagulation enzymes *in vivo*. The interaction between AT and its target enzymes is greatly accelerated by heparinlike molecules. Heparin binds to AT, inducing a conformational change that results in exposure of the reactive loop. In addition, heparin simultaneously binds to thrombin and AT, resulting in approximation of the inhibitor to its target enzyme. Relatively long chains of heparin (≥ 18 saccharide units) are necessary to mediate this function. In contrast, only the heparin-induced conformational change is necessary for the inhibition of factor Xa. Accordingly, inhibition of factor Xa can be accelerated

by both short (e.g., low-molecular-weight) and long heparin chains (e.g., unfractionated heparin). Heparin is not normally found in circulating plasma; however, heparinlike glycosaminoglycans are present on the surface of endothelial cells. Thus, endothelial cells are probably the site of heparin-mediated neutralization of factor Xa and thrombin *in vivo*.

In contrast to AT, heparin cofactor II inhibits only thrombin among the coagulation enzymes. Like AT, it also requires a cofactor, dermatan sulfate, to accelerate the rate of inhibition. As the name implies, heparin also can accelerate the activity of heparin cofactor II, but high concentrations (greater than 1.0 U/mL) are necessary for effective activity. The precise physiologic role of heparin cofactor II is uncertain. As dermatan sulfate usually is found in tissues outside the vessel wall, it may be involved in extravascular regulation of thrombin. Levels of heparin cofactor II are decreased in patients with acute consumptive coagulopathies, suggesting that it does contribute to the normal regulation of hemostasis.

The Protein C System

The protein C system is necessary for regulating the major cofactors of the coagulation cascade, factors Va and VIIIa (39,40 and 41). The protein C system involves multiple protein components and can be divided into three phases: activation, functional, and regulatory (Table 46.9). Thrombin is the major activator of protein C. As with many of the reactions in the hemostatic system, the activation of protein C by thrombin is a slow process unless a catalytic cofactor, thrombomodulin, is present. Thrombomodulin is a membrane-associated protein that is found on the surface of most, but not all endothelial cells. Thrombin binds with high affinity to thrombomodulin and undergoes an important conformational change; thrombin associated with thrombomodulin loses its procoagulant activity and no longer converts fibrinogen to fibrin nor activates factors V, VIII, XIII, or platelets. Thus, binding of thrombin by thrombomodulin is itself an anticoagulant step, as it converts thrombin from a procoagulant enzyme into an anticoagulant enzyme. The altered thrombin, however, readily activates protein C and can still be inhibited by AT (Fig. 46.18). A receptor for protein C and APC is present on endothelial cells and appears to facilitate the activation of protein C by thrombomodulin.

TABLE 46.9. PROTEIN C SYSTEM

Phase	Component	Function
Activation	Thrombin	Activates protein C
	Thrombomodulin	Cofactor for thrombin activation of protein C
	Protein C	Substrate for thrombin-thrombomodulin
	Endothelial PC receptor	Facilitate protein C activation, endothelial signaling
Functional	Activated protein C	Proteolytic degradation of factors Va and VIIIa, endothelial signaling when bound to endothelial PC receptor
	Protein S	Cofactor for activated protein C
	Factors Va and VIIIa	Substrates for activated protein C
	Factor V	Cofactor for degradation of factor VIIIa by activated protein C and PS
Regulation	Protein C inhibitor	Neutralizes activated protein C
	Plasminogen activator inhibitor-1	Neutralizes activated protein C
	α_1-Antiprotease	Neutralizes activated protein C

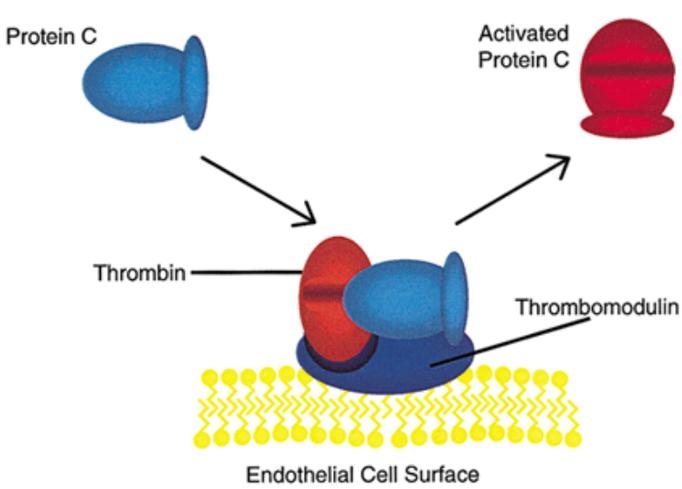


FIGURE 46.18. Protein C (PC) is activated by an endothelial cell (EC)-dependent process. Thrombin (IIa) binds to thrombomodulin (TM), a constituent membrane protein present on most ECs and undergoes a conformational change. Thrombin bound to TM loses its procoagulant activity but becomes an effective activator of protein C, releasing a short peptide from the heavy chain of PC. Activated protein C (APC) is then available to carry out its functions.

Activated protein C degrades factors Va and VIIIa by cleaving selected peptide bonds in the heavy chain of each cofactor. This step requires a phospholipid surface and a protein cofactor, protein S (Fig. 46.19). Protein S is produced by a variety of cells, including hepatocytes, endothelial cells, and megakaryocytes. The complement related protein, C4b binding protein (C4b BP), avidly binds protein S in plasma and thus the plasma pool of protein S is divided into C4b BP-bound and free protein S fractions. Protein S bound to C4b BP is unable to function as an active cofactor for APC and thus only the free protein S fraction is functionally active. Any stimulus that causes an increase in C4b BP without also increasing the total protein S concentration can lead to a functional deficiency of free protein S; this pattern may be observed in some inflammatory reactions.

The degradation of factor VIIIa by APC/protein S is accelerated in the presence of native, nonactivated factor V. A common mutation in factor V (Arg506→Gln, factor V_{Leiden}) does not affect the procoagulant activity of factor Va but does affect both

the ability of APC to inactivate factor Va_{Leiden} and the APC cofactor activity of nonactivated factor V_{Leiden}. This results in “resistance” to APC and a thrombophilic phenotype.

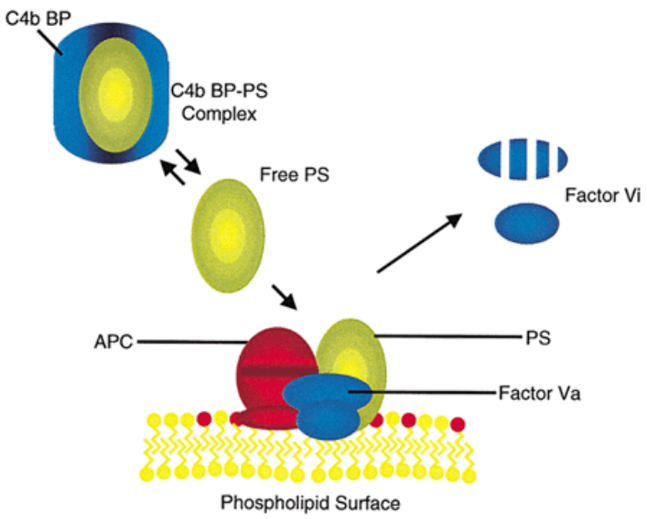


FIGURE 46.19. Activated protein C (APC) degrades the major cofactors of the coagulation cascade, factors Va and VIIIa by cleaving the heavy chains of Va and VIIIa. Protein S (PS) serves as a catalytic cofactor for this phospholipid-dependent process. Protein S binds to C4b binding protein (C4b BP) in plasma and the bound PS is not able to participate in the activity of APC. Only free protein S is capable of functioning as an effective cofactor.

The ability of APC to degrade factor Va is inhibited by prothrombin and thus a change in the balance between the concentration of prothrombin and APC may alter the effectiveness of the protein C system. For example, during the early phases of oral anticoagulant therapy, the concentration of protein C drops more rapidly than prothrombin. In this setting, prothrombin would be in relative excess to the APC formed, leading to relative inhibition of the protein C system. This may be one of the underlying mechanisms of warfarin-induced skin necrosis. Another setting in which this may be important is the relatively common prothrombin mutation at nucleotide 20210, which results in an increased plasma concentration of prothrombin and a thrombophilic phenotype.

Activated protein C is inhibited by several serpins, including protein C inhibitor (which is identical to plasminogen activator inhibitor-3), plasminogen activator inhibitor-1 (PAI-1) and α_1-antiprotease. Experimental evidence suggests that APC has profibrinolytic activity in addition to its effect on factors Va and VIIIa. There are at least two potential mechanisms for this profibrinolytic activity. First, there is competition between APC and tPA for protein C inhibitor and PAI-1. In the presence of increased amounts of APC, the inhibitors are utilized to neutralize APC, leaving tPA free to activate plasminogen. Second, activation of procarboxypeptidase B is dependent on thrombin activation. Inhibition of thrombin generation by degradation of factors Va and VIIIa decreases formation of carboxypeptidase B. As discussed below, carboxypeptidase B (also known as thrombin activatable fibrinolysis inhibitor, TAFI) inhibits the fibrinolytic pathway at multiple stages.

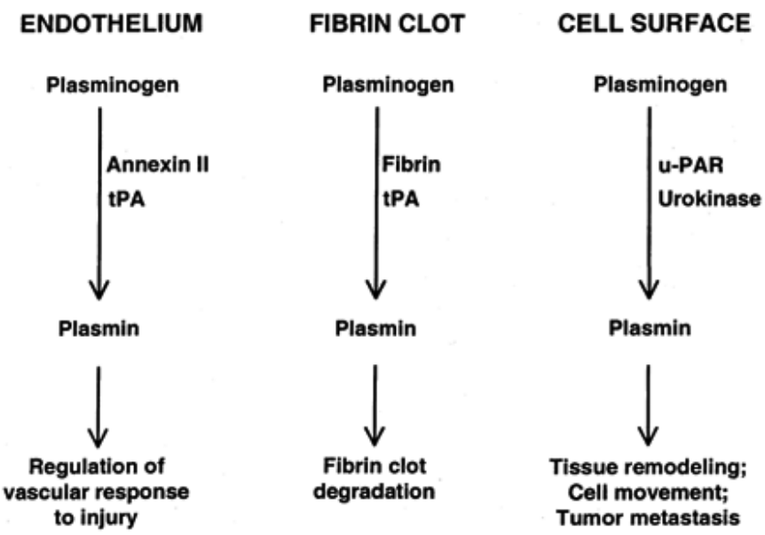


FIGURE 46.20. The key enzyme of the fibrinolytic system is plasmin, which serves a number of functions. Plasmin can be activated in a number of environments, with its functional role dependent on the site of activation. Plasmin can be activated on the surface of endothelial cells by tissue plasminogen activator (tPA) in the presence of annexin II. This pathway appears to play an important role in regulating the vascular response to injury. Plasminogen also is activated by tPA within the matrix of the fibrin clot, where it is capable of digesting excess clot. A third important site of plasmin activity is on the surface of cells. The plasminogen activator urokinase binds to a specific cell receptor, urokinaselike plasminogen activator receptor (u-PAR), where it converts plasminogen to plasmin. The cell-associated plasmin is then able to digest matrix constituents and facilitate tissue remodeling, cell movement through extracellular matrices, and tumor-cell invasion.

The Fibrinolytic System

The fibrinolytic system is also a complex, multicomponent system that can be divided into activation, functional, and regulatory phases (Table 46.10) (42,43,44,45,46,47,48 and 49). The key component of the fibrinolytic system is plasminogen, a 92-kd zymogen of a serine protease. The amino terminal portion of plasminogen contains multiple kringle domains with lysine binding sites that mediate the binding of plasminogen to other proteins. The catalytic domain is in the carboxyl-terminal domain of the molecule and is activated by limited proteolytic cleavage. Several mechanisms for the activation of plasminogen have been described, including clot-based activation, endothelial-mediated activation, cell-based activation and fluid phase activation (Fig. 46.20).

TABLE 46.10. THE FIBRINOLYTIC SYSTEM

Phase	Component	Function
Activation	Plasminogen	Substrate for plasminogen activators
	Tissue plasminogen activator (tPA)	Major physiological activator of plasminogen
	Fibrin	Cofactor for activation of plasminogen by tPA
	Annexin II	Cell surface receptor for plasminogen and tPA that mediates plasminogen activation
	Kallikrein	Converts prourokinase to urokinase
	Urokinase	Direct activator of plasminogen
	Prourokinase	Precursor of urokinase
	Urokinase-type Plasminogen Activator Receptor	Cellular receptor for urokinase
Functional	Plasmin	Potent proteolytic enzyme
	Fibrin	Major substrate for plasmin
	Fibrinogen	Substrate for free plasmin
	Factors V and VIII	Substrates for free plasmin
	Platelet membrane glycoproteins	Substrates for free plasmin
	Extracellular matrix	Substrate for cell-associated and free plasmin
Regulation	Plasminogen activator inhibitor-1	Inhibits tPA, urokinase and activated protein C
	Protein C inhibitor	Inhibits tPA and activated protein C
	α_2-Antiplasmin	Inhibits plasmin
	Carboxypeptidase B (Thrombin activated inhibitor of fibrinolysis - TAFI)	Inhibits fibrinolysis by cleavage of carboxyl-terminal lysine and arginine residues

Plasmin is a potent proteolytic enzyme that is capable of degrading fibrin clots. However, the ability of plasmin to cleave fibrin is impaired by factor XIIIa-mediated cross-linking of fibrin monomers. The more extensively the monomers are cross-linked, the more resistant they are to plasmin activity whereas clots formed in the absence of factor XIIIa are very susceptible to lysis by plasmin. Plasmin digestion of fibrinogen and fibrin monomers occurs in a relatively ordered pattern (Fig. 46.21). The initial sites of attack include the carboxyl end of the α chain and the amino terminal end of the ß chain. Cleavage at these sites generates a large residual molecule known as fragment X. Plasmin can then cleave bonds in a helical region, releasing a D domain; the remaining D/E component is known as fragment Y. Cleavage of the remaining helical region releases another D domain and the central E domain. If adjacent D domains have been cross-linked by factor XIIIa, then plasmin cleavage will generate fragments known as D-dimers, which can be measured as an indicator of *in vivo* fibrinolysis. Under normal circumstances, plasmin remains bound to the fibrin clot through interaction of its lysine binding sites with the fibrin matrix. As the clot is lysed and plasmin is released, the free plasmin is inhibited by α_2-antiplasmin. Thus plasmin activity is normally limited to the clot itself. However, if plasmin breaks free of the fibrin clot and is not inhibited, it is capable of digesting fibrinogen, factors V and VIII, as well as platelet membrane glycoproteins.

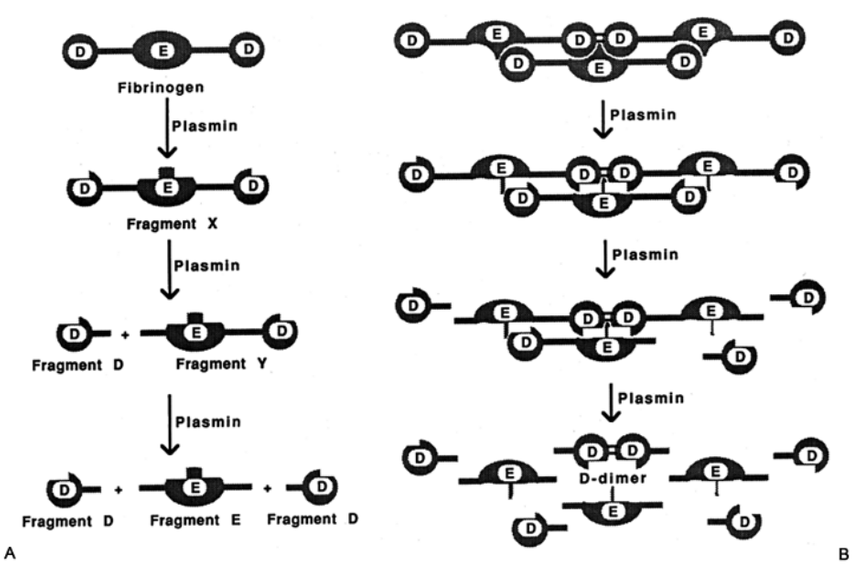


FIGURE 46.21. Plasmin can degrade either fibrinogen (**A**) or fibrin (**B**) in a stepwise fashion. Plasmin initially attacks the carboxyl terminus of the Aα chain and the amino terminus of the Bß chain, resulting in the formation of the large degradation product designated fragment X. Cleavage of an interconnecting helical segment releases a D domain from the residual fragment (fragment Y). Further proteolysis results in the release of fragments D and E. Analogous fragments are released from the degradation of fibrin. However, adjacent D domains covalently cross-linked by factor XIIIa remain cross-linked, forming D-dimers. These D-dimers are thus unique products of plasmin degradation of cross-linked fibrin gels.

Clot-based activation of plasminogen is mediated by tPA, a serine protease that is secreted by endothelial cells. Free tPA is capable of activating plasminogen in solution but the rate of acti

vation is very slow. However, activation of plasminogen by tPA is greatly accelerated in the presence of fibrin (but not intact fibrinogen). As both tPA and plasminogen have a relatively high affinity for fibrin, both proteins are incorporated into the matrix of the fibrin gel. The dependence on fibrin as a cofactor essentially limits this mechanism of plasminogen activation to sites of fibrin clot formation.

Plasminogen and tPA also bind avidly to intact endothelial cells. The receptor for both proteins is annexin II, which is expressed both within the cytoplasm and on the surface of endothelial cells (48). Binding of plasminogen to surface annexin II appears to be dependent on cleavage of the bond between lysine 307 and arginine 308 of annexin II, which exposes a carboxyl-terminal lysine residue for interaction with the lysine binding site on plasminogen. Lipoprotein (a) is a low-density lipoprotein particle that has a variable number of kringle domains that are structurally similar to the kringle domains of plasminogen. Lipoprotein (a) can successfully compete with plasminogen for binding to annexin II at lysine 307. Thus, the higher the concentration of lipoprotein (a), the less plasminogen will bind to endothelial cells and be activated by tPA. Of note, high levels of lipoprotein (a) are associated with an increased risk of cardiovascular disease. The ability of annexin II to bind plasminogen can also be destroyed by carboxypeptidase B, an enzyme that removes carboxyl-terminal lysine and arginine residues.

The binding of tPA to annexin II is dependent on a free cysteine found near the amino terminal end of annexin II. Site-directed mutagenesis of this cysteine significantly impairs the ability of tPA to bind to the modified annexin II. Homocysteine can also interact with this cysteine residue to form a disulfide-linked derivative of annexin II that lacks the ability to bind tPA. Formation of this modified form of annexin II increases as the concentration of homocysteine increases and tPA binding is significantly impaired at homocysteine levels commonly encountered in clinical practice and that have been associated with an increased risk of atherothrombotic disease. The association between increased levels of homocysteine or lipoprotein (a) with vascular disease suggests that this mechanism of plasminogen activation may play a critical role in regulating the hemostatic system and the response to vascular injury.

TABLE 46.11. COMMON PATTERNS ASSOCIATED WITH BLEEDING DISORDERS

	Platelet Disorders	Coagulation Disorders	Fibrinolytic Disorders
Clinical history	Mucocutaneous bleeding	Soft-tissue bleeding	Delayed bleeding
Screening laboratory tests	Long bleeding time and/or low platelet count	Long PT and/or APTT	Normal

APTT, activated partial thromboplastin time

PT, prothrombin time

Urokinase is another serine protease that can convert plasminogen into plasmin. In contrast to tPA, free urokinase readily activates plasminogen and thus urokinase can mediate fluid phase plasminogen activation. Additionally, urokinase is a key mediator of cell-associated plasminogen activation. A variety of cells express a glycosylphosphatidylinositol-linked membrane receptor known as urokinase-type plasminogen activator receptor (u-PAR, CD87). When urokinase binds to u-PAR it is capable of activating plasminogen in association with the cell surface. Such localized plasmin formation plays an essential role in a variety of cellular

processes, including tissue remodeling, cell migration, and tumor metastasis. This mechanism also appears to contribute to the bleeding tendency of patients with acute promyelocytic leukemia (AML-M3). The leukemic cells in this disorder frequently express u-PAR, leading to inappropriate plasmin activation and an enhanced risk of bleeding. Treatment with trans-retinoic acid induces maturation of the leukemic cells, down regulation of u-PAR expression and reversal of the bleeding tendency. Formation of the u-PAR/urokinase complex also initiates intracellular signaling that is probably mediated by proteins such as integrins that are juxtaposed to u-PAR on the cell surface. A linkage to a transmembrane signaling protein is needed because, as a glycosylphosphotidylinositol-linked surface protein, u-PAR has no intracytoplasmic signaling domain. Plasminogen-activator inhibitor-1 binds to and inhibits both tPA and urokinase. When PAI-1 binds to urokinase associated with u-PAR, the entire complex is internalized into the cell where the PAI-1/urokinase complex is degraded and u-PAR is recycled to the cell surface.

Regulation of the fibrinolytic system occurs on at least three levels. First, the major activators of plasmin are inhibited by serpins. Tissue plasminogen activator is inhibited by both PAI-1 and protein C inhibitor, whereas urokinase is inhibited by PAI-1. As discussed in the section on protein C, these inhibitors also inhibit APC, and their availability for regulation of tPA is dependent on the amount of APC in the environment. That is, the higher the concentration of APC, the less PAI-1 and protein C inhibitor will be available to regulate tPA. The second level of regulation involves the serpin α_2-antiplasmin. In the fluid phase, α_2-antiplasmin rapidly inhibits free plasmin. However, within the clot matrix it competes with fibrin for the same lysine binding sites on plasmin (Fig. 46.22). Thus, as long as plasmin is bound to the fibrin clot matrix (or cell surface) it is not inhibited by α_2-antiplasmin, but once plasmin breaks free of the fibrin clot, it is instantaneously inhibited by α_2-antiplasmin. Because of this mechanism, it is uncommon to encounter free plasmin in the circulation. When this does happen, it is often associated with a significant coagulopathy related to the nonspecific proteolytic activity of plasmin. For example, this is one of the key pathological changes underlying disseminated intravascular coagulation. One of the reasons that such pathological fibrinolysis can even occur is that there is normally twice as much plasminogen as α_2-antiplasmin on a molar basis. With marked activation of the fibrinolytic system, α_2-antiplasmin is consumed and this regulatory mechanism is overwhelmed.

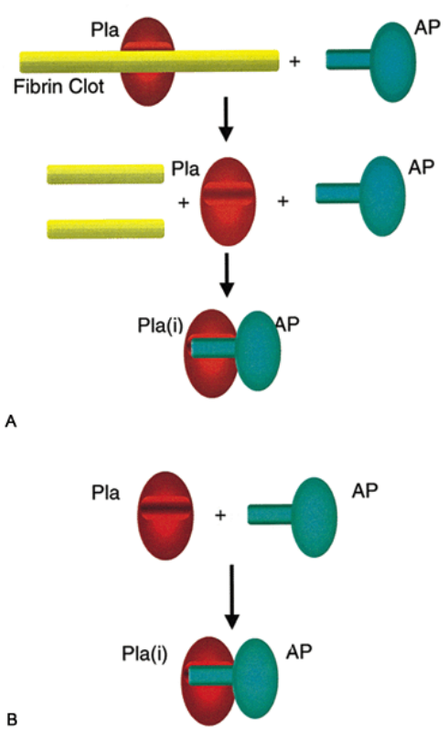


FIGURE 46.22. Plasmin (Pla) activity is regulated by the serine protease inhibitor α_2-antiplasmin (AP). Plasmin associated with a fibrin gel is relatively protected from inhibition because fibrin and AP compete for the same lysine binding sites on plasmin (**A**). As the fibrin gel is lysed, the released plasmin is neutralized by AP, effectively limiting plasmin action to the fibrin matrix. In contrast, plasmin formed in plasma or released from cell surfaces or dissolving clot is very rapidly neutralized by AP (**B**). The rapid inhibition of plasmin normally limits the sites where it can function. However, depletion of AP, as might occur with overwhelming activation of plasmin, can lead to free plasmin in the circulation.

A third mechanism of regulating fibrinolysis involves the enzyme carboxypeptidase B, also known as TAFI. Procarboxypeptidase B circulates as an inactive zymogen that is converted to its active form by thrombin or factor XIa. Thrombin-mediated activation of TAFI is accelerated by thrombomodulin on endothelial surfaces. Factor XIa appears to play a critical role in the activation of TAFI, either through generation of sufficient thrombin to mediate TAFI activation or direct activation of TAFI by platelet-associated factor XIa. Impaired activation of TAFI may account for many of the bleeding manifestations associated with congenital deficiency of factor XI.

TAFI removes carboxyl-terminal lysine and arginine amino-acid residues from a number of proteins. As lysine is often involved in the interaction between the lysine binding sites on plasmin(ogen) and other proteins, cleavage of such lysine residues impairs plasmin(ogen) binding to target proteins. In addition, removal of the carboxyl-terminal lysine from modified plasminogen or plasmin impairs the potential activity of the enzyme. Thus TAFI can induce changes in both plasmin(ogen) and related proteins that impair clot-based and endothelial-mediated fibrinolysis.

Summary of the Regulation of the Coagulation Cascade

These four systems interact in a synergistic manner to regulate activity of the coagulation cascade. As indicated in Fig. 46.23, virtually every step of the cascade is controlled by one or more of these systems. Combined with the factors regulating primary hemostasis, these systems are normally successful in limiting hemostatic plug formation to sites of vascular disruption. A

number of hereditary and acquired defects of the regulatory system have now been described. Alterations that result in a decrease in the effectiveness of any one of these regulatory components can lead to increased clot formation *in vivo*, giving rise to a thrombotic tendency.

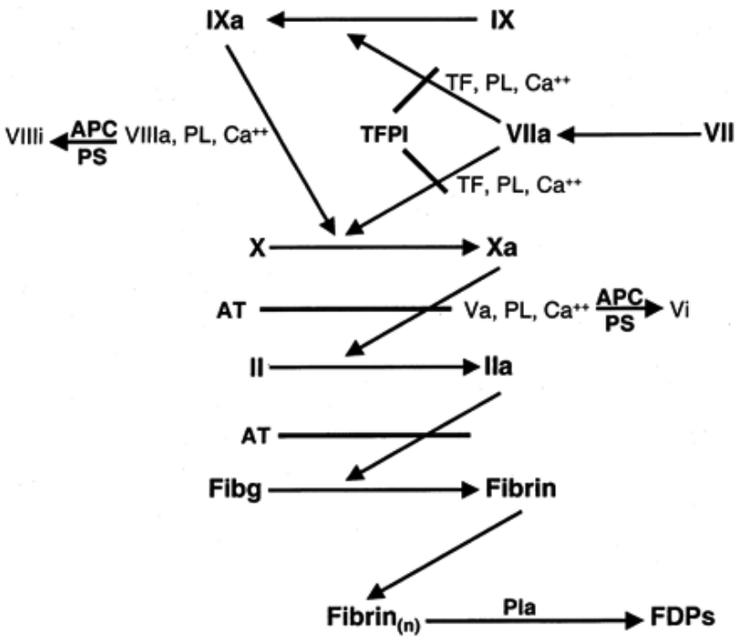


FIGURE 46.23. The four systems involved in the regulation of the coagulation cascade, tissue factor pathway inhibitor (TFPI), antithrombin (AT), activated protein C (APC) and plasmin (Pla) function in a synergistic way to control thrombin formation. The points of regulation of the extrinsic (tissue factor) pathway of coagulation are depicted in this figure. As can be seen, these four systems control almost every step of the extrinsic pathway of coagulation.

AN APPROACH TO HEMOSTATIC DISORDERS

Part of "46 - Overview of Hemostasis"

There are a number of reasons why the hemostatic system may be evaluated during the clinical management of patients. Common reasons include therapeutic drug monitoring, presurgical evaluation of hemostasis, evaluation of a possible bleeding tendency, evaluation of a possible thrombotic tendency, evaluation for the possibility of a circulating lupus anticoagulant, evaluation for the possibility of disseminated intravascular coagulation, and evaluation for other specific disorders. The hemostasis laboratory plays a key role in monitoring oral anticoagulant and unfractionated heparin therapy. In contrast, the laboratory has only a minor role in monitoring fibrinolytic therapy (e.g., tPA therapy) or antiplatelet therapy. In general, laboratory parameters measured during these latter therapeutic regimens provide little useful information, at least currently.

Routine presurgical evaluation of hemostasis remains a controversial topic. In general, routine preoperative bleeding times, prothrombin times (PTs) and activated partial thromboplastin times (APTTs) do not accurately predict the risk of bleeding during surgery. In fact, a thorough clinical history is probably more sensitive and specific for the identification of patients with a bleeding disorder. In the absence of a clinical history or medical disorder that suggests a compromised hemostatic system, routine preoperative hemostasis testing is probably not indicated.

Laboratory evaluation of hemostasis is indicated whenever the clinical and/or medical history suggests the possibility of altered hemostasis. As hemostasis is a balanced system with potent procoagulant and regulatory mechanisms, bleeding or thrombotic disorders can arise whenever this balance is disturbed. Two general mechanisms may lead to an increase in bleeding: any alteration that decreases procoagulant activity or any defect that increases regulatory activity. Von Willebrand's disease and hemophilia A provide examples of the first mechanism, whereas heparin therapy and fibrinolytic therapy provide examples of the second mechanism. Similarly, thrombosis may occur because of decreased regulatory activity, as with antithrombin deficiency, or increased procoagulant activity, as may occur during the course of metastatic carcinoma.

Evaluation of a Potential Bleeding Disorder

The clinical history of a patient with a potential bleeding disorder is used to determine whether the patient truly does have a bleeding problem, whether the problem is likely to be congenital or acquired, and whether the defect is related to altered primary hemostasis, secondary hemostasis, or regulation of hemostasis (50,51,52,53,54 and 55). As perceptions of the seriousness of a bleeding problem vary significantly from patient to patient, the clinician must establish the frequency of bleeding, sites of bleeding, and severity of bleeding to determine the actual presence of a significant bleeding diathesis. Age of onset and family history are very helpful in distinguishing between congenital and acquired disorders. An early age of onset is suggestive of a congenital bleeding disorder, whereas the sudden appearance of bleeding in a previously healthy adult suggests an acquired problem. Family history can help establish a pattern of hereditary transmission. Some disorders are characterized by sex-linked recessive inheritance while other disorders are characterized by either autosomal-dominant or autosomal-recessive inheritance.

The pattern of bleeding is very helpful in constructing a differential diagnosis. Mucocutaneous bleeding, characterized by epistaxis, ecchymosis, and genitourinary bleeding, is suggestive of a platelet disorder. Soft-tissue bleeding, including hemarthrosis, hematomas, and retroperitoneal bleeding, is more suggestive of a coagulation disorder such as hemophilia A. Delayed bleeding is a classic manifestation of fibrinolytic-type bleeding. The clinical history should include a thorough medical history to document any other medical problems that may predispose to a bleeding diathesis. In addition, a careful medication history, including over-the-counter medications, should be obtained. It is helpful to remember that the most common cause of abnormal platelet function is drug effect, particularly aspirin therapy.

Following the clinical history, routine screening tests are commonly performed. These include the platelet count and bleeding time to assess platelet function and the PT and APTT to assess fibrin clot formation. Note that these screening tests do not assess fibrin stabilization or fibrinolysis. The bleeding time is a commonly performed test to assess primary hemostasis. A number of techniques for determining the bleeding time have been described, but the most commonly performed method is the template bleeding time using a disposable device. Although this test appears to be simple and straightforward, a number of

technical variables can confound results. Therefore a standardized technique should be used when performing this procedure. Newer technologies for assessing platelet function, such as the PFA 100 (Dade Behring, Deerfield, Illinois), are emerging and may replace the bleeding time as a screening test of platelet function.

The PT and APTT are tests of fibrin clot formation. The PT evaluates the extrinsic system of coagulation, beginning with activation of coagulation by tissue factor/VIIa (Fig. 46.8), and is sensitive to defects in fibrinogen, prothrombin, factor V, factor X, and factor VII. Of interest, the tissue factor/VIIa activation of factor IX is not evident with the PT unless the PT reagent is first diluted. This is because the lack of dependence of the PT on factors VIII and IX is related to the concentration of tissue factor and phospholipid used in the assay rather than the biological importance of this part of the pathway. The APTT evaluates the intrinsic system of coagulation and is sensitive to defects in fibrinogen, prothrombin, factor V, factor X, factor VIII, factor IX, factor XI, factor XII, prekallikrein, and HMW kininogen. A number of technical variables can affect performance of the PT and APTT. Among these specimen acquisition, sample processing, and choice of reagent are key components. Care should be taken to obtain a clean venipuncture sample and to anticoagulate it promptly with an appropriate amount of citrate. The concentration of citrate used is important, particularly for monitoring anticoagulant therapy. It is now recommended that 0.109 M (3.2%) citrate be used to collect coagulation samples. The sample should be processed as soon as possible to yield platelet-poor plasma. PT and APTT reagents vary significantly in their sensitivity and responsiveness to various hemostatic disorders. Therefore, the laboratory should be aware of the performance characteristics of their reagents and should carefully construct their own reference ranges to adequately distinguish normal from abnormal.

Based on the results of the clinical history and screening laboratory tests, a limited number of patterns emerge (Table 46.11). Platelet disorders usually are characterized by mucocutaneous bleeding and a long bleeding time and/or a low platelet count. In contrast, coagulation disorders usually are characterized by a history of soft tissue bleeding and an abnormality of the APTT and/or PT. Fibrinolytic disorders are classically characterized by a history of delayed bleeding, often following trauma, and normal screening tests of hemostasis. Based on these initial results, further evaluation for a specific disorder can be undertaken. It is helpful to keep in mind the relative frequency of congenital disorders associated with bleeding when pursuing such an evaluation (Table 46.12). The three most common disorders are von Willebrand's disease, hemophilia A, and hemophilia B. Specific platelet disorders are reviewed in Chapter 47 and specific coagulation disorders are presented in Chapter 48.

TABLE 46.12. APPROXIMATE FREQUENCY OF CONGENITAL BLEEDING DISORDERS

Disorder or Deficiency	Frequency per 10^6 Population
Fibrinogen	1
Prothrombin	< 0.5
Factor V	< 0.5
Factor VII	< 0.5
Factor VIII	60-100
Factor IX	10-20
Factor X	< 0.5
Factor XI	1
Factor XIII	< 0.5
α_2-antiplasmin	< 0.5
von Willebrand's disease	> 100
Bernard-Soulier syndrome	< 0.5
Glanzmann's thrombasthenia	< 0.5
Hermansky-Pudlak syndrome	< 0.5

Evaluation of a Potential Thrombotic Tendency

As with the evaluation of bleeding disorders, the evaluation of a potential thrombotic tendency begins with a thorough medical history (56,57,58 and 59). One of the primary goals of the history is to determine the likelihood of a congenital defect. In this regard, it is important to note the age of onset of symptoms and any family history of thrombosis. Most congenital defects are associated with onset of symptoms prior to the age of 45, and the family history for most of these disorders also is positive. Congenital thrombophilia usually is associated with recurrent deep venous thrombosis but can, on occasion, be associated with arterial thrombosis. A complete medical history must be obtained to determine if there is an underlying medical illness that may predispose to thrombosis (metastatic tumor, autoimmune disorder, etc.). A thorough medication history is also important.

There are no screening assays that assess the overall function of the regulatory system of coagulation or the degree of activation of the procoagulant systems. Therefore, the laboratory approach to potential thrombotic disorders involves measuring selected individual components of the regulatory systems. In general, assays that reflect the biological activity of the protein in question should be used for screening purposes. In practice, such functional assays have been difficult to develop for some components of the system, notably protein S. In addition, assays for lupus anticoagulant activity and antiphospholipid antibodies may be helpful in the initial screening of a patient for a thrombotic tendency. A variety of markers of the level of activation of the hemostatic system have been described. These include sensitive D-

dimer, fibrinopeptide A, $F_{1.2}$, thrombin-antithrombin complex, protein C activation peptide, and plasminogen-antiplasmin complex assays. The usefulness of these markers in determining the risk of thrombosis, however, remains undetermined. Disorders associated with thrombosis are discussed in detail in Chapter 49.

REFERENCES

1. Caen JP, Han ZC, Bellucci S, et al. Regulation of megakaryocytopoiesis. *Haemostasis* 1999;29:27–40.

2. McNicol A, Israels SJ. Platelet dense granules: structure, function and implications for haemostasis. *Thromb Res* 1999;95:1–18.

3. Michelson AD, Furman MI. Laboratory markers of platelet activation and their clinical significance. *Curr Opin Hematol* 1999;6:342–348.

4. Nurden P, Poujol C, Nurden AT. The evolution of megakaryocytes to platelets. *Baillieres Clin Haematol* 1997;10:1–27.

5. Schmitz G, Rothe G, Ruf A, et al. European Working Group on Clinical Cell Analysis: consensus protocol for the flow cytometric characterisation of platelet function. *Thromb Haemost* 1998;79:885–896.

6. Wendling F, Han ZC. Positive and negative regulation of megakaryocytopoiesis. *Baillieres Clin Haematol* 1997;10:29–45.

7. Andrews RK, Shen Y, Gardiner EE, et al. The glycoprotein Ib-IX-V complex in platelet adhesion and signaling. *Thromb Haemost* 1999;82:357–364.

8. Ruggeri ZM. von Willebrand factor. *J Clin Invest* 1997;99:559–564.

9. Ruggeri ZM. Structure and function of von Willebrand factor. *Thromb Haemost* 1999;82:576–584.

10. Sadler JE. Biochemistry and genetics of von Willebrand factor. *Annu Rev Biochem* 1998;67:395–424.

11. Shen Y, Romo GM, Dong JF, et al. Requirement of leucine-rich repeats of glycoprotein (GP) Ibalpha for shear-dependent and static binding of von Willebrand factor to the platelet membrane GP Ib-IX-V complex. *Blood* 2000;95:903–910.

12. Ware J. Molecular analyses of the platelet glycoprotein Ib-IX-V receptor. *Thromb Haemost* 1998;79:466–78.

13. Carpenter G, Ji Q. Phospholipase C-gamma as a signal-transducing element. *Exp Cell Res* 1999;253:15–24.

14. Coughlin SR. Protease-activated receptors and platelet function. *Thromb Haemost* 1999;82:353–356.

15. Levy-Toledano S. Platelet signal transduction pathways: could we organize them into a 'hierarchy'? *Haemostasis* 1999;29:4–15.

16. Leevers SJ, Vanhaesebroeck B, Waterfield MD. Signaling through phosphoinositide 3-kinases: the lipids take centre stage. *Curr Opin Cell Biol* 1999;11:219–225.

17. Rittenhouse SE. Phosphoinositide 3-kinase activation and platelet function. *Blood* 1996;88:4401–14.

18. Vanhaesebroeck B, Waterfield MD. Signaling by distinct classes of phosphoinositide 3-kinases. *Exp Cell Res* 1999;253:239–254.

19. Plow EF, Byzova T. The biology of glycoprotein IIb-IIIa. *Coron Artery Dis* 1999;10:547–551.

20. Nelsestuen GL, Shah AM, Harvey SB. Vitamin K-dependent proteins. *Vitam Horm* 2000;58:355–389.

21. Vlot AJ, Koppelman SJ, Bouma BN, et al. Factor VIII and von Willebrand factor. *Thromb Haemost* 1998;79:456–465.

22. Bouma BN, von dem Borne PA, Meijers JC. Factor XI and protection of the fibrin clot against lysis—a role for the intrinsic pathway of coagulation in fibrinolysis. *Thromb Haemost* 1998;80:24–27.

23. Mann KG. Biochemistry and physiology of blood coagulation. *Thromb Haemost* 1999;82:165–174.

24. Schmaier AH, Rojkjaer R, Shariat-Madar Z. Activation of the plasma kallikrein/kinin system on cells: a revised hypothesis. *Thromb Haemost* 1999;82:226–233.

25. Walsh PN. Platelets and factor XI bypass the contact system of blood coagulation. *Thromb Haemost* 1999;82:234–242.

26. Lorand L. Sol Sherry Lecture in Thrombosis: research on clot stabilization provides clues for improving thrombolytic therapies. *Arterioscler Thromb Vasc Biol* 2000;20:2–9.

27. Haller H. Endothelial function. General considerations. *Drugs* 1997;53:1–10.

28. Mantovani A, Garlanda C, Introna M, Vecchi A. Regulation of endothelial cell function by pro- and anti-inflammatory cytokines. *Transplant Proc* 1998;30:4239–4243.

29. Rosenberg RD, Aird WC. Vascular-bed—specific hemostasis and hypercoagulable states. *N Engl J Med* 1999;340:1555–1564.

30. Marcus AJ, Broekman MJ, Drosopoulos JH, et al. The endothelial cell ecto-ADPase responsible for inhibition of platelet function is CD39. *J Clin Invest* 1997;99:1351–1360.

31. Riddell DR, Owen JS. Nitric oxide and platelet aggregation. *Vitam Horm* 1999;57:25–48.

32. Lindahl AK. Tissue factor pathway inhibitor: from unknown coagulation inhibitor to major antithrombotic principle. *Cardiovasc Res* 1997;33:286–291.

33. Bjork I, Olson ST. Antithrombin. A bloody important serpin. *Adv Exp Med Biol* 1997;425:17–33.

34. Carrell RW. How serpins are shaping up [comment]. *Science* 1999;285:1861.

35. Crowther DC, Evans DL, Carrell RW. Serpins: implications of a mobile reactive centre. *Curr Opin Biotechnol* 1992;3:399–407.

36. Lawrence DA. The role of reactive-center loop mobility in the serpin inhibitory mechanism. *Adv Exp Med Biol* 1997;425:99–108.

37. Wells MJ, Sheffield WP, Blajchman MA. The clearance of thrombin-antithrombin and related serpin-enzyme complexes from the circulation: role of various hepatocyte receptors. *Thromb Haemost* 1999;81:325–337.

38. Whisstock J, Skinner R, Lesk AM. An atlas of serpin conformations. *Trends Biochem Sci* 1998;23:63–67.

39. Esmon CT, Gu JM, Xu J, et al. Regulation and functions of the protein C anticoagulant pathway. *Haematologica* 1999;84:363–368.

40. Esmon CT, Xu J, Gu JM, et al. Endothelial protein C receptor. *Thromb Haemost* 1999;82:251–258.

41. O'Brien LM, Mastri M, Fay PJ. Regulation of factor VIIIa by human activated protein C and protein S: inactivation of cofactor in the intrinsic factor Xase. *Blood* 2000;95:1714–1720.

42. Bajzar L, Morser J, Nesheim M. TAFI, or plasma procarboxypeptidase B, couples the coagulation and fibrinolytic cascades through the thrombin-thrombomodulin complex. *J Biol Chem* 1996;271:16603–16608.

43. Bajzar L, Nesheim M, Morser J, et al. Both cellular and soluble forms of thrombomodulin inhibit fibrinolysis by potentiating the activation of thrombin-activable fibrinolysis inhibitor. *J Biol Chem* 1998;273:2792–2798.

44. Boffa MB, Reid TS, Joo E, et al. Characterization of the gene encoding human TAFI (thrombin-activable fibrinolysis inhibitor; plasma procarboxypeptidase B). *Biochemistry* 1999;38:6547–6558.

45. Bouma BN, Mosnier LO, Meijers JC, et al. Factor XI dependent and independent activation of thrombin activatable fibrinolysis inhibitor (TAFI) in plasma associated with clot formation. *Thromb Haemost* 1999;82:1703–1708.

46. Collen D. The plasminogen (fibrinolytic) system. *Thromb Haemost* 1999;82:259–270.

47. Dear AE, Medcalf RL. The urokinase-type-plasminogen-activator receptor (CD87) is a pleiotropic molecule. *Eur J Biochem* 1998;252:185–193.

48. Hajjar KA, Krishnan S. Annexin II: a mediator of the plasmin/plasminogen activator system. *Trends Cardiovasc Med* 1999;9:128–138.

49. Menell JS, Cesarman GM, Jacovina AT, et al. Annexin II and bleeding in acute promyelocytic leukemia. *N Engl J Med* 1999;340:994–1004.

50. Cattaneo M, Lecchi A, Agati B, et al. Evaluation of platelet function with the PFA-100 system in patients with congenital defects of platelet secretion. *Thromb Res* 1999;96:213–217.

51. Nosek-Cenkowska B, Cheang MS, Pizzi NJ, et al. Bleeding/bruising symptomatology in children with and without bleeding disorders. *Thromb Haemost* 1991;65:237–241.

52. Harrison P, Robinson MS, Mackie IJ, et al. Performance of the platelet function analyser PFA-100 in testing abnormalities of primary haemostasis. *Blood Coagul Fibrinolysis* 1999;10:25–31.

53. Peyvandi F, Mannucci PM. Rare coagulation disorders. *Thromb Haemost* 1999;82:1207–1214.

54. Rand ML, Carcao MD, Blanchette VS. Use of the PFA-100 in the assessment of primary, platelet-related hemostasis in a pediatric setting. *Semin Thromb Hemost* 1998;24:523–529.

55. Sramek A, Eikenboom JC, Briet E, et al. Usefulness of patient interview in bleeding disorders. *Arch Intern Med* 1995;155:1409–1415.

56. Bertina RM. Molecular risk factors for thrombosis. *Thromb Haemost* 1999;82:601–609.

57. Briet E, van der Meer FJ, Rosendaal FR, et al. The family history and inherited thrombophilia. *Br J Haematol* 1994;87:348–352.

58. Lane DA, Grant PJ. Role of hemostatic gene polymorphisms in venous and arterial thrombotic disease. *Blood* 2000;95:1517–1532.

59. Samama M, Gerotziafas G, Conard J, et al. Clinical aspects and laboratory problems in hereditary thrombophilia. *Haemostasis* 1999;29:76–99.

47

Laboratory Evaluation of Platelet Disorders

John T. Brandt

Platelet disorders are a common clinical cause of excessive bleeding. These disorders most commonly represent acquired problems and the etiology is often apparent from a thorough clinical history and performance of a few straightforward laboratory procedures. Congenital disorders of platelet function comprise an interesting group of disorders that teach much about the key mechanisms of the platelet response in hemostasis. While identification of abnormal platelet function may be relatively straightforward in patients with congenital disorders, pinpointing the actual biochemical defect may involve very specialized studies and is often beyond the scope of the clinical laboratory. The congenital disorders are usually evident early in life, although clinical manifestations may be delayed until adulthood. In contrast, the acquired abnormalities of primary hemostasis are common causes of clinical bleeding problems at any age. This chapter presents a diagnostic approach to congenital and acquired disorders affecting platelet function, starting with an overview of the laboratory procedures that are used in the evaluation of these disorders. The pathogenesis and evaluation of thrombocytopenia, as well as the therapeutic options for various platelet disorders, are briefly reviewed.

LABORATORY METHODS FOR EVALUATION OF PLATELET FUNCTION

Part of "47 - Laboratory Evaluation of Platelet Disorders"

A variety of laboratory techniques may be used to assess platelet function, but a limited number of techniques are necessary for most clinical situations (Table 47.1). Procedures for the evaluation of von Willebrand factor (vWF) are among the most important because von Willebrand disease (vWD) is the most common congenital disorder of platelet function. Other procedures may be used to assess the platelet response to selected agonists, the ability of platelets to secrete the contents of their granules or the activation state of circulating platelets. Selection of which assays to use for a specific patient should be based on clinical history, results of screening assays and a review of platelet morphology on the peripheral blood smear.

TABLE 47.1. LABORATORY TECHNIQUES FOR EVALUATION OF PLATELET FUNCTION

Procedure	Property Evaluated	Clinical Uses
Platelet count	Platelet number in peripheral blood	Platelet production
Mean platelet volume	Average platelet size	Assessing platelet mass, platelet production
Bleeding time	Global platelet function	Screen for platelet function
PFA 100®	Platelet function	Screen for platelet function
Platelet aggregation studies	Platelet function	Delineate type of platelet dysfunction; detection of variant vWD, HIT
vWF antigen (vWF:ag)	Quantity of vWF protein in blood	Evaluation of possible vWD
vWF:Rcof	Measurement of vWF functional activity	Evaluation of possible vWD
Collagen binding of vWF	Measurement of vWF interaction with collagen	Screen for variant vWD
GP Ibα binding of vWF	Measurement of vWF function	Screen for vWD
Factor VIII binding to vWF	Evaluation of factor VIII binding site on vWF	Screen for Type 2N vWD
vWF multimeric analysis	Measure of vWF polymerization	Evaluation of variant vWD
Electron microscopy	Platelet morphology, platelet granules	Evaluation of congenital platelet disorders
Bone marrow exam	Assessment of platelet production	Evaluation of thrombocytopenia, thrombocytosis
Platelet antibody	Detection of antiplatelet antibodies	Evaluation of possible ITP
Platelet survival	Rate of platelet turnover	Mechanism of thrombocytopenia
PF4, βTG	In vivo platelet release	Detection of platelet activation
Flow cytometry	Expression of membrane proteins	Evaluation of platelet activation, function and diagnosis of HIT
Flow cytometry	Measurement of platelet RNA	Assessment of platelet production

Abbreviations used: vWD, von Willebrand disease; vWF, von Willebrand factor; vWF: Rcof, ristocetin cofactor activity of vWF; HIT, heparin-induced thrombocytopenia; ITP, idiopathic/immune thrombocytopenic purpura; PF4, platelet factor 4; βTG, β-thromboglobulin

The Bleeding Time

The bleeding time is a commonly performed test to assess global function of primary hemostasis (1,2 and 3). A number of techniques for determining the bleeding time have been described, but the most commonly performed method is the template bleeding time using a disposable device. Although this test appears to be simple and straightforward, a number of technical variables can confound results. Therefore, each laboratory should attempt to standardize the procedure as completely as possible. Key variables include the accuracy of the pressure in the blood pressure cuff, the pressure with which the template device is applied to the volar surface of the forearm, the time of stasis and the direction of the incision. Following completion of the bleeding time the wound should be closed with a butterfly-type bandage and covered with a bandage as performance of the bleeding time can be associated with significant scar formation in some individuals. The bleeding time is prolonged in conditions associated with significant impairment of platelet function, including thrombocytopenia. However, the bleeding time is not very sensitive to mild platelet disorders and may be normal in patients with significant clinical defects. Therefore, if a patient has a history of significant mucocutaneous bleeding, further workup is indicated even if the bleeding time is normal. The bleeding time is not predictive of the risk of bleeding during surgical procedures and thus there is no role for the bleeding time as a routine preoperative screening procedure. Because of the technical and diagnostic limitations of the bleeding time, other procedures have been developed to assess primary hemostasis.

Analysis of Von Willebrand Factor

Commonly used procedures to evaluate vWF include vWF antigen concentration (vWF:ag), ristocetin cofactor activity (vWF:Rcof), vWF binding to collagen or recombinant glycoprotein Ibα (GP Ibα), the platelet aggregation response to ristocetin, and vWF multimeric analysis (4,5 and 6). Immunoelectrophoresis was once the only practical method available for determination of vWF:ag, but enzyme-linked immunosorbant assays (ELISAs) and automated immunoassays are now available. Quantitation of vWF:ag provides a measure of the total plasma concentration of vWF that is independent of the multimeric composition and function of vWF. In general, the plasma concentration of vWF:ag parallels that of functional assays for vWF, such as factor VIII activity and vWF:Rcof, when the result of each of assay is expressed as a percent of pooled normal plasma.

An alteration in the ratio of vWF:ag to one or more functional assays may indicate the presence of a dysfunctional vWF.

vWF has two major functions in hemostasis; first it mediates platelet adhesion and second, it is an obligate carrier of factor VIII in plasma. Therefore, an assessment of the functional capacity of vWF involves an assessment of both platelet-associated vWF function and factor VIII-vWF interaction. The aggregation response of formalin-fixed platelets to ristocetin and a limiting concentration of plasma as a source of vWF is the classic technique for assessing the platelet-associated activity of vWF; the activity measured by this type of assay is designated vWF:Rcof. Limiting factors of the vWF:Rcof assay include the technical difficulty of the assay, the observation that vWF:Rcof activity does not always correlate with *in vivo* activity in patients with variant (type 2) forms of vWD and the lack of precision in the assay. This has led to the development of alternative assays for assessment of platelet-associated vWF function. The collagen-binding assay provides a measure of the ability of vWF to bind to collagen immobilized in microtiter plate wells. As such binding is dependent on the multimeric composition of vWF, this assay can be used to screen for Type 2 vWD associated with a decrease in HMW multimers. Another assay measures the ability of patient vWF to bind to an immobilized recombinant GP 1bα peptide in the presence of ristocetin. The recombinant peptide includes the binding site for vWF and thus this assay, which measures the interaction between vWF and its platelet receptor, is an alternative to the ristocetin-induced aggregation assay.

A functional assay for factor VIII activity is used to screen for the interaction between vWF and factor VIII. The specific ability of patient vWF to bind normal factor VIII can be determined using a binding assay in which patient vWF is captured on a microtiter plate by an anti-vWF monoclonal antibody. This assay may be useful for identification of Type 2N vWD, which is characterized by an abnormality in vWF that diminishes its ability to bind and carry factor VIII in plasma. Such patients may present with a hemophilialike clinical phenotype.

The multimeric composition of vWF, which is helpful in differentiating type 2 vWD, is most accurately assessed by SDS-agarose gel electrophoresis with immunoblotting. High-resolution electrophoresis and blotting may permit identification of specific abnormalities characteristic of certain sub-types of type 2 vWD. Performance of multimeric analysis is technically complex and subject to a number of analytical variables. This procedure should be performed by experienced, expert laboratories that use appropriate internal controls to verify the adequacy of the procedure. The major goal of multimeric analysis is to determine whether or not the normal high molecular weight multimers of vWF are present.

Platelet Response to Agonists

Platelet aggregation studies are used to assess the response of platelets to a variety of agonists (5, 7,8,9,10,11,12,13,14,15 and 16). Platelet aggregation is most commonly performed by a turbidimetric method using platelet-rich plasma but also can be performed by an impedance method that utilizes either whole blood or platelet-rich plasma. Commonly used platelet agonists for *in vitro* studies include ADP, collagen, epinephrine, arachidonic acid, and ristocetin. The basic principle of turbidimetric aggregometry is that light transmittance increases as platelet aggregates form (Fig. 47.1). ADP and epinephrine at the appropriate concentration characteristically induce a biphasic wave of aggregation in normal individuals. The initial (primary) wave is because of formation of small aggregates in direct response to agonist interaction with the platelet surface. The second wave is caused by the formation of large aggregates and is dependent on the release/secretion of endogenous platelet products. Patients may show a normal biphasic response to these agonists, a primary wave followed by disag

gregation (designated a secondary wave defect) or a total lack of response (designated a primary wave defect). The responses to arachidonic acid and collagen are typically monophasic; abnormalities of response to these agonists may include a partial, delayed, or absent response. Secretion of ATP from the responding platelets, a marker of dense granule release, can be simultaneously measured in some aggregometers using a chemiluminescent procedure.

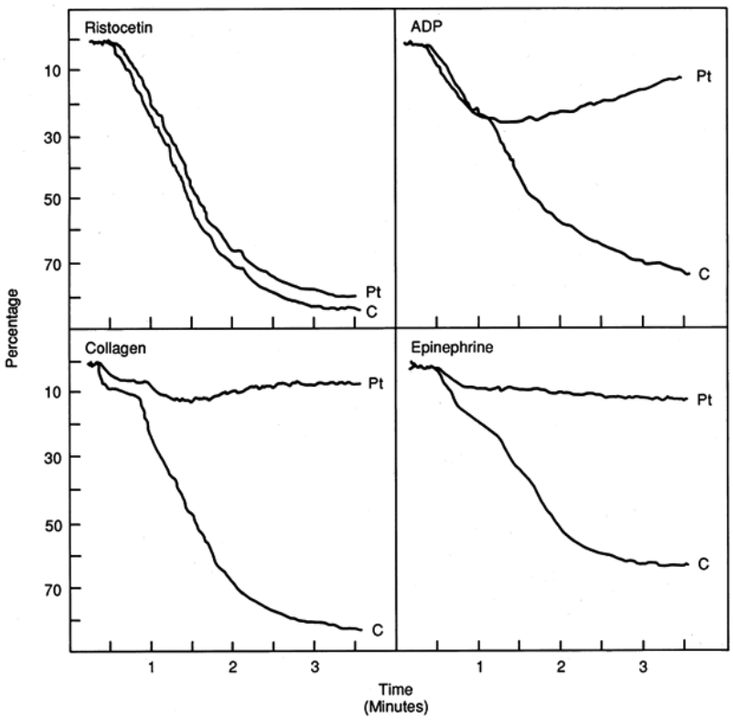


FIGURE 47.1. Normal (C) and abnormal (Pt) aggregation responses for ADP, collagen, epinephrine, and ristocetin. The normal response to ADP at optimal concentrations is a biphasic wave, with the secondary wave dependent on an intact release response by the responding platelets. An inability to undergo release results in a primary wave of aggregation, followed by disaggregation (Pt); this pattern is designated a secondary wave defect. A total lack of response to ADP would be designated a primary wave defect. Abnormal responses to collagen include a delay in onset of aggregation or, as in this case, a decreased response to collagen. Epinephrine usually results in a biphasic wave of aggregation; abnormalities to this agonist vary from total lack of response to a primary wave followed by disaggregation. Diminished response to ristocetin (not shown) is characteristic of abnormalities of platelet adhesion.

Normal platelets tend to aggregate to ristocetin in a monophasic pattern when the ristocetin concentration is greater than 1.0 mg/mL. Patients with vWD or Bernard-Soulier syndrome may show a lack of response at these concentrations. Normal platelets do not aggregate at ristocetin concentrations less than 0.8 mg/mL. However, platelet aggregation may be seen at concentrations less than 0.8 mg/mL in samples from patients with platelet-type vWD and type 2B vWD. Thus, it is important to test with both high (>1.0 mg/mL) and low (<0.8 mg/mL) concentrations of ristocetin when performing ristocetin-induced platelet aggregation assays.

Although thrombin is a key platelet agonist *in vivo*, it is not commonly used in the clinical laboratory. One of the limiting factors is the propensity of platelet-rich plasma to clot upon addition of activating concentrations of thrombin. This can be circumvented by using washed platelets or adding agents that block fibrin polymerization. The development of a peptide agonist that binds to and activates the thrombin receptor, protease activated receptor-1 (PAR-1), has provided a mechanism for evaluation of the thrombin-activated pathways in platelet-rich plasma. However, it is unclear whether the PAR-1 pathway is the only pathway activated by thrombin and thus use of the peptide may not reflect all of thrombin's platelet stimulating activities.

An alternative to turbidometric aggregometry is impedance aggregometry. The principle of impedance aggregation is that formation of platelet aggregates on the surface of a sensitive electrode effectively “insulates” the electrode, altering the resistance to an electric current. The change in resistance over time is used as a measure of the platelet aggregation response. As with turbidometric techniques, a variety of agonists may be used. An

advantage of impedance aggregometry is that it can be performed in whole blood, which is thought to be a more natural environment. However, as the aggregation response is affected by the concentration of red cells, the whole blood sample usually is diluted to give a constant red cell concentration and thus it does not actually represent a completely natural environment.

The PFA 100 (Dade Behring, Deerfield, Illinois) measures the platelet aggregation response in a flow chamber. Whole blood, anticoagulated with citrate, is passed through an aperture coated with collagen and either ADP or epinephrine. As platelets pass through the orifice, they adhere and become activated to form platelet aggregates that eventually occlude the aperture and block flow. The time that it takes for closure of the aperture is dependent on platelet function and relatively independent of fibrin formation. This procedure appears to be very sensitive to a variety of platelet function abnormalities, including vWD, platelet storage pool defects and platelet inhibitors such as aspirin and GP IIb/IIIa inhibitors.

Flow cytometry has emerged as a powerful tool for the analysis of platelet function. Platelet activation is associated with surface expression of proteins not found on quiescent platelets. Analysis of these proteins has been used to evaluate *in vivo* platelet activation, which occurs in a variety of clinical settings. A similar approach can be taken to follow the response of platelets to agonists added *in vitro*. For example, methods have been developed for the diagnosis of heparin-induced thrombocytopenia (HIT) using flow cytometry. Advantages of this type of approach include the ability to study platelets in whole blood and the requirement for only small amounts of blood. Flow cytometry also can be used to evaluate the formation of platelet microparticles *in vivo* or *in vitro*. Activation of platelets is associated with shedding of membrane-rich particles in which procoagulant phospholipids are readily available to support the coagulation cascade. Formation of such procoagulant microparticles is a significant component of the pathogenesis of some platelet disorders, including HIT.

Another use of flow cytometry is quantitation of platelet RNA content. As platelets are shed from megakaryocytes they contain a moderate amount of RNA that is lost over time in the circulation, analogous to reticulocytes. Measurement of platelet RNA content may indicate whether there has been a shift to younger platelets (higher RNA content); this may help differentiate a thrombocytopenia resulting from increased consumption from a thrombocytopenia from decreased production.

Evaluation of vWF and platelet function plays a major role in assessing the possibility of congenital platelet disorders, congenital vWD and acquired vWD. In contrast, these studies have a limited role in the evaluation of most acquired platelet disorders. Thus the laboratory evaluation should be guided by whether the clinical history suggests a congenital or acquired disorder. Congenital disorders of platelet function usually present with a history of mucocutaneous bleeding starting during childhood. Consequently, this is the population that will be evaluated most frequently. Congenital platelet disorders can be classified according to the aspect of platelet function (adhesion, activation or aggregation) that is abnormal. The laboratory features of disorders of adhesion, activation, and aggregation are sufficiently distinct to permit construction of a flow diagram for their laboratory evaluation and diagnosis (Fig. 47.2).

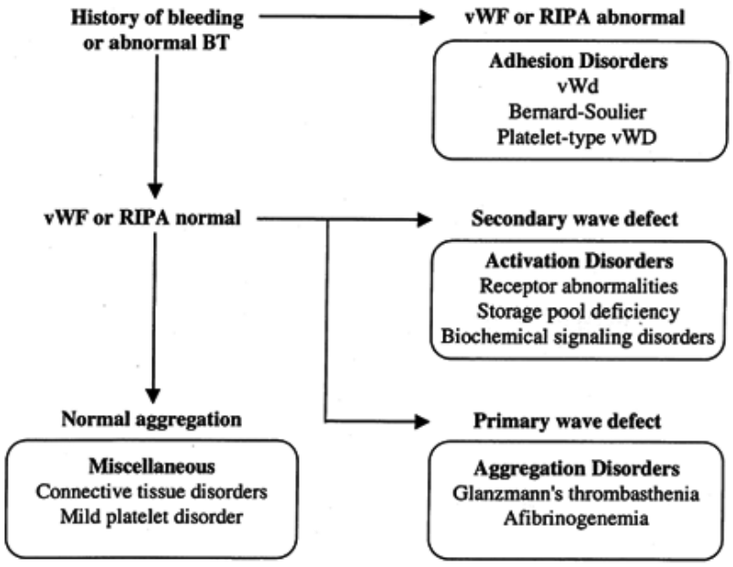


FIGURE 47.2. Flow diagram for diagnosis of abnormalities of platelet function. Abnormalities of platelet adhesion are characterized by an abnormal aggregation response to ristocetin and/or abnormalities of von Willebrand factor. Abnormalities of platelet activation usually are characterized by a secondary wave-type defect with ADP, while abnormalities of platelet aggregation are characterized by a primary wave defect. Occasional patients with mild platelet dysfunction may show a normal response to all the agonists used in routine studies.

CONGENITAL DISORDERS OF PLATELET FUNCTION

Part of "47 - Laboratory Evaluation of Platelet Disorders"

Abnormalities of Platelet Adhesion

A number of congenital disorders affecting platelet adhesion have been described, the most common of which is vWD. The other disorders in this group, although rare, are important for what they teach about the normal process of platelet adhesion and because a correct diagnosis is necessary in order to provide appropriate therapy.

von Willebrand Disease

vWD is the most common congenital bleeding disorder known, but its actual frequency still has not yet been accurately defined (4,5,13,17,18,19 and 20). Some have estimated the frequency to be as high as 1:100, but clinical experience would suggest a lower incidence. It is typically inherited as an autosomal-dominant trait; that is, heterozygous patients are clinically symptomatic. Compound heterozygous and homozygous forms of the disease also have been described. Patients with vWD have clinical manifestations characteristic of platelet disorders, with the most common complaints related to increased bruising, epistaxis, menorrhagia, and bleeding following surgical procedures. Indeed, the development of iron deficiency anemia in a pubescent female may be an indication of an underlying platelet disorder such as vWD. The family history usually is positive, with male and female relatives being affected with equal frequency.

The bleeding in vWD tends to be mild to moderate, and symptoms often spontaneously improve after adolescence. This disorder is relatively well tolerated, unless the hemostatic system is challenged by trauma or surgery, at which time bleeding may

be serious. During childhood, this may occur with oral surgery such as tooth extractions or tonsillectomy. The post-partum period is another high-risk time for bleeding in women with vWD, despite the pregnancy-associated rise in the level of vWF. The level of vWF drops precipitously in the post-partum period, and the patient may begin to bleed seriously at that time. Occasionally, a diagnosis is not established during childhood and the patient remains asymptomatic until later in life (usually after age 50), when recurrent gastrointestinal bleeding may be the initial manifestation of vWD. An extensive evaluation of the gastrointestinal tract is usually unrevealing or demonstrates only angiodysplasia.

Homozygous (type 3) vWD usually is associated with a more severe hemorrhagic diathesis. The clinical manifestations may mimic those of hemophilia A because of the very low or absent factor VIII. These patients may be misdiagnosed as having hemophilia A if appropriate laboratory studies are not performed. It is important to discriminate between type 3 vWD and hemophilia A because the therapy of these two disorders is different.

vWD is caused by a deficiency or structural abnormality of vWF. vWF is synthesized by endothelial cells and megakaryocytes and is found in subendothelial spaces, plasma, and platelet α-granules. Platelet α-granule vWF appears to make a significant contribution to platelet adhesion, as reflected by the relatively shorter bleeding times in those vWD patients who have a low plasma vWF concentration but a normal platelet vWF concentration. vWF has two major functions; first, it mediates platelet adhesion and second, it is an obligate carrier for factor VIII in plasma. The platelet-associated function of vWF is dependent on production of high-molecular-weight multimers (see Chapter 46) whereas the binding of factor VIII is less dependent on polymerization of vWF. This dichotomy leads to dissociation between vWF-mediated platelet function and vWF-dependent factor VIII concentration in some subtypes of vWD.

Many of the functional domains of vWF have now been mapped to structural areas of the molecule (Fig. 47.3). The factor VIII binding site is located in the amino-terminal portion of the monomer, which also contains a heparin binding site. The A1 domain, which is bounded by a disulfide-bonded loop, contains the site that binds to platelet GP Ib/IX/V. The A2 domain contains a protease-sensitive domain that is involved in plasma processing of vWF. The vWF that is secreted from endothelial cells is normally larger than that found in plasma. A plasma protease cleaves the secreted vWF, resulting in removal of the largest multimers. This is an important regulatory step, as such very large multimers have an intrinsic propensity to bind to GP Ib/IX/V and initiate a platelet aggregation response. A collagen-binding domain is located in the A3 domain of the molecule, another loop structure delimited by disulfide bonds. The RGD sequence that binds to GP IIb/IIIa under conditions of high-shear stress is located in the carboxyl-terminal end of the molecule.

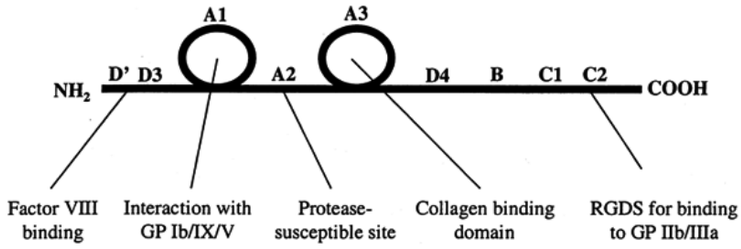


FIGURE 47.3. The domain structure for monomeric von Willbrand factor (vWF) is depicted in this figure. The binding site for factor VIII is located in domains D′ and D3 near the amino terminal end of the monomer. The binding site for platelet GP Ib/IX/V is located in the A1 domain, which is delineated by a disulfide-linked loop. There is a critical protease susceptible region located in the A2 domain. This is the site of extracellular processing that removes the very large multimers secreted by endothelial cells. Mutations that lead to excessive proteolysis cause Type 2A von Willebrand's disease while a deficiency of the protease is associated with thrombotic thrombocytopenic purpura. The A3 domain contains a collagen binding site that is probably involved in vWF interaction with extracellular matrix. There is an RGDS sequence near the carboxyl terminal end of the molecule that mediates the binding of vWF to platelet GP IIb/IIIa.

A number of genetic defects may give rise to vWD. The syndrome is divided into three major types, based on whether the defect results in a simple quantitative deficiency of vWF or a qualitative defect in vWF (Table 47.2). Type 1 vWD is the most common type of vWD and is characterized by a moderate quantitative deficiency of vWF, with vWF:ag usually in the range of 30% to 50% of normal. The vWF that is present has a normal multimeric structure and normal function in terms of mediating platelet adhesion and binding factor VIII. Type 2 vWD is characterized by qualitative abnormalities in vWF. These abnormalities may be associated with loss of high-molecular-weight multimers (type 2A), enhanced affinity for GP Ib/IX/V (type 2B), decreased affinity for GP Ib/IX/V without loss of high-molecular-weight-multimers (type 2M) or loss of affinity for factor VIII (type 2N). Type 3 vWD is characterized by a severe deficiency of vWF and, consequently, a deficiency of factor VIII.

Analysis of the molecular biology of vWD has indicated that mutations in certain regions of the molecule are associated with distinct phenotypes. Type 2A vWD is the second most common type of vWD and represents a heterogeneous group of defects. In some patients, the disorder is associated with missense mutations that result in amino acid substitutions in the A2 domain. Apparently these mutations make this region of the molecule more susceptible to the plasma protease that normally processes vWF. The enhanced degradation of vWF leads to a loss of the high-molecular-weight multimers and diminished vWF-mediated platelet function. In other patients, there appears to be a mutation that affects endothelial cell processing and secretion of vWF such that the high-molecular-weight multimers are not appropriately formed and secreted. Patients with type 2A vWD show a significant loss of high-molecular-weight multimers of vWF.

Type 2B vWD is the third most common type of vWD and is associated with point mutations in or near the A1 domain that result in amino-acid substitutions that increase the affinity of vWF for GP Ib/IX/V so that vWF can bind to GP Ib/IX/V under normal flow conditions. The bound abnormal vWF is re

moved from the circulation, resulting in a mild deficiency of the high-molecular-weight multimers. The enhanced affinity of the abnormal vWF for GP Ib/IX/V also accounts for the enhanced platelet aggregation response seen at low concentrations of ristocetin. A clinical correlate of the interaction between the abnormal vWF and GP Ib/IX/V is the propensity of these patients to develop mild to moderate thrombocytopenia.

TABLE 47.2. CLASSIFICATION OF VON WILLEBRAND DISEASE

VWD Sub-Type	Defect	Inheritance
Type 1	Decreased quantity of all vWF multimers High-molecular-weight multimers present Concordance between vWF:ag and vWF:Rcof	Autosomal dominant
Type 2A	High molecular weight multimers absent Abnormal platelet dependent vWF function Discordance between vWF:ag and vWF:Rcof	Autosomal dominant
Type 2B	Loss of some high-molecular-weight multimers Increased affinity of vWF for GP Ib/IX/V Association with thrombocytopenia	Autosomal dominant
Type 2M	Abnormal platelet-dependent vWF function Discordance between vWF:ag and vWF:Rcof High-molecular-weight multimers present	Autosomal dominant
Type 2N	Decreased affinity for factor VIII Normal platelet-dependent vWF function Decreased factor VIII, prolonged APTT	Autosomal recessive
Type 3	Severe deficiency of vWF Multimeric analysis cannot be performed Factor VIII also markedly decreased	Autosomal recessive
Platelet-Type	Abnormal GPIb/V/IX (Ibα) Increased affinity for normal vWF Association with thrombocytopenia	Autosomal dominant

Type 2M vWD is an uncommon type of vWD and also has been associated with missense mutations in the A1 domain of vWF. In contrast to type 2B vWD, the mutations in the GP Ib/IX/V-binding domain of vWF associated with type 2M result in a loss of affinity for GP Ib/IX/V. Thus, the abnormal vWF is not able to adequately bind to GP Ib/IX/V and mediate platelet adhesion. The vWF multimeric structure usually is normal in such patients. Type 2N vWD is a rare type of vWD and is characterized by mutations in the amino-terminal end of the molecule that affect the affinity of vWF for factor VIII. The relative impact of the mutation on affinity for factor VIII depends on the precise site of the mutation. Consequently, the plasma level of factor VIII can range from very low (˜1%) to moderately low (10% to 15%) in affected patients.

The molecular pathology of type 1 and type 3 vWD is more complex. Although it might seem that type 3 vWD would simply represent a homozygous form of type 1 vWD, this does not appear to be the case. Most heterozygous carriers of mutations that result in type 3 vWD are asymptomatic. The mutations associated with type 1 vWD appear to be distinct from those associated with type 3 vWD, and, in many cases, are still not defined. A variety of defects, including point mutations and gene deletions may give rise to a quantitative deficiency of vWF. The synthesis of vWF is a complicated process and it is possible that defects in other components of the synthetic pathway may affect the plasma level of vWF, giving rise to a type 1 vWD phenotype. In a mouse model of vWD, the plasma deficiency is from abnormal glycosylation of vWF that results in a shorter half-life for vWF. The finding of normal levels of platelet vWF (which would be protected from enhanced clearance) and reduced plasma vWF in some patients with type 1 vWD suggests that a similar mechanism may be operative in humans.

Given the complex nature of the pathogenesis of vWD, it should not be surprising that the laboratory phenotype is quite variable. Therefore, a number of procedures must be utilized to evaluate the possibility of vWD whenever a patient presents with a history of mucocutaneous bleeding of unknown etiology. The bleeding time is often used as a screening test and is typically mildly to moderately prolonged in vWD. However, the bleeding time may fluctuate significantly over time and may be within the normal range at any given point. In particular, patients with normal levels of platelet vWF may have only a modest, if any, prolongation of the bleeding time. Therefore, a normal bleeding time in a patient with a history of mucocutaneous bleeding does not rule out vWD, and further laboratory studies are indicated in such patients. The aspirin tolerance test, in which the bleeding time is repeated 2 to 8 hours after a challenge with aspirin (650 mg), may be helpful; patients with mild vWD (or other mild platelet function disorders) usually show a marked prolongation of the bleeding time after such an aspirin challenge. The PFA 100 analyzer also appears to be a very sensitive procedure for the detection of abnormal vWF-mediated platelet function and is less invasive than the bleeding time.

Factor VIII activity and vWF:ag assays are used to measure the concentration of the components of the vWF-factor VIII complex in plasma. Patients with type 1 vWD also usually show reduced, but concordant results whereas the results in samples from patients with type 2 vWD may show discordance between these parameters; indeed, such a finding should raise the possibility of type 2 vWD. Blood type has a significant impact on the plasma concentration of vWF, with lower levels of vWF associated with blood type O. Of note, the ABO carbohydrate group is present on vWF and it is possible that the blood type O phe

notype is associated with enhanced clearance of vWF. As a consequence, the patient's blood type should be taken into consideration when assessing the possibility of vWD. The platelet-associated function of vWF may be assessed using ristocetin-induced platelet aggregation at low (<0.8 mg/mL) and high (>1.0 mg/mL) concentrations, the ristocetin cofactor assay, the collagen binding assay or the recombinant GP Ibα binding assay. Of these assays, the platelet aggregation assay is perhaps the least sensitive for the diagnosis of type 1 vWD. The multimeric composition of vWF is best assessed using SDS-agarose electrophoresis. This procedure may provide pivotal information because the only abnormalities that may be found in some patients with type 2 vWD are an abnormal screening test (e.g., bleeding time) and an abnormal vWF multimeric composition.

The diagnosis of vWD depends on the combination of an appropriate clinical history, laboratory abnormalities involving vWF and evidence of hereditary transmission (Table 47.3). Patients with type 1 vWD typically have a prolonged bleeding time, a concordant decrease in vWF:ag, vWF:Rcof and factor VIII, a decreased aggregation response to ristocetin (at concentrations >1.0 mg/mL), and a normal multimeric composition. Patients with type 2A vWD tend to have a low to normal vWF:ag concentration, with vWF:Rcof activity even lower than the antigen concentration; multimeric analysis demonstrates the characteristic loss of high-molecular-weight multimers. The results in patients with type 2B are quite variable. Perhaps the most useful parameter for identification of these patients is the increased platelet aggregation at low concentrations (<0.8 mg/mL) of ristocetin. High-molecular-weight multimers also are decreased, but not usually to the extent of the decrease associated with type 2A vWD. Type 2M vWD is associated with a low to normal vWF:ag and a significantly reduced vWF:Rcof activity. Multimeric analysis is normal in patients with type 2M vWD. Type 2N vWD may be easily misdiagnosed as hemophilia unless the ability of patient vWF to bind factor VIII is assessed. All other parameters of vWF may be normal. Type 3 vWD is characterized by very long bleeding times, absent vWF, very low or undetectable factor VIII activity, and absent vWF:Rcof. Multimeric analysis cannot be performed in these patients because of the lack of antigen.

TABLE 47.3. TYPICAL LABORATORY PARAMETERS IN VON WILLEBRAND DISEASE AND RELATED DISORDERS

Parameter	Type 1	Type 2A	Type 2B	Type 2M	Type 2N	Type 3	Platelet-Type	Bernard-Soulier Syndrome
Bleeding time	↑ or N	↑	↑	↑	N	↑	↑	↑
Platelet count	N	N	↓ or N	N	N	N	↓ or N	↓
APTT	↑	↑ or N	N or ↑	N or ↑	↑	↑	N or ↑	N
vWF:Ag	↓	↓ or N	N or ↓	↓ or N	N or slight ↓	↓↓	N or ↓	N
vWF:Rcof	↓	↓	↓, N or ↑	↓	N or slight ↓	↓↓	↓ or N	N
VIII	↓	N or ↓	N or ↓	N or ↓	↓↓	↓↓	N or ↓	N
RIPA	↓	↓	↑	↓	N	↓↓	↑	↓↓
Multimers	N	Abn	Abn	Abn	N	ND	Abn	N

Abbreviations used: APTT, activated partial thromboplastin time; vWF:ag, von Willebrand factor antigen; vWF:Rcof, ristocetin cofactor activity; VIII, factor VIII coagulant activity; N, normal; ND, not done (antigen too low); RIPA, ristocetin-induced platelet aggregation.

Many of the bleeding episodes with type 1 vWD can be treated effectively with desmopressin (1-deamino-8-D-arginine vassopressin, Aventis, Bridgewater, NJ). This agent induces a twofold to tenfold increase in the plasma level of vWF, a response that is often adequate to maintain hemostasis during dental procedures or other minor surgical procedures. The fibrinolytic inhibitor, ε-aminocaproic acid [Amicar (Immunex Corp., Seattle, Washington)], may be a useful adjunct, particularly in procedures involving the oral mucosa. vWF can be replaced using cryoprecipitate, an excellent source of fully functional vWF, but this is no longer recommended as a first line of therapy because of the potential for transmission of viral diseases. Certain highly purified factor VIII concentrates contain high concentrations of vWF and are effective in the therapy of vWD. An advantage of the factor VIII concentrates is that they have undergone one or more viral inactivation steps. Such products are now recommended as a first line of therapy for patients needing replacement of vWF. Generally, this would include patients with type 2 or 3 vWD. In particular, patients with type 2B vWD must be treated with care. Treatment of these patients with desmopressin may induce thrombocytopenia, presumably resulting from release of the abnormal vWF from endothelial cells. The therapy of choice in these patients is a factor VIII concentrate with sufficient concentrations of vWF.

Platelet-Type von Willebrand Disease (Pseudo-vWD)

Platelet-type vWD is a rare disorder that presents with a clinical history and laboratory profile consistent with a diagnosis of type 2B vWD (5,21,22,23,24 and 25). These patients have a history of mucocutaneous bleeding, prolonged bleeding time, intermittent thrombocytopenia, and decreased concentration of the high-molecular-weight multimers of vWF. In contrast to patients with Type 2B vWD, the defect in these patients lies in platelet GP Ib/IX/V rather than vWF. The abnormal GP Ib/IX/V complex has an increased affinity for normal vWF. Normal high-molecular-weight multimers of vWF bind to the abnormal receptor and are cleared from the circulation. The inappropriate interaction between the abnormal platelets and normal vWF also can lead to removal of the platelets from the circulation, causing the thrombocytopenia. Thus the pathogenesis of platelet-type vWD is similar to the pathogenesis of type 2B vWD in that a mutation leads to an inappropriate interaction between vWF and GP Ib/IX/V. Molecular analysis of patients with platelet-type vWD has revealed missense mutations in the gene for GP Ibα that result in amino-acid substitutions in the vWF binding domain of the molecule. Peptide fragments of GP Ibα with these mutations expressed by recombinant technology show increased affinity for vWF in the

presence of low concentrations of ristocetin, consistent with the presumed mechanism for loss of vWF multimers associated with this disorder.

The laboratory picture in these patients resembles type 2B vWD (Table 47.3). The bleeding time usually is prolonged, the concentration of vWF:ag is variable, but often reduced and patient platelets show an enhanced aggregation response to low concentrations of ristocetin. The clearance of high-molecular-weight vWF multimers by the abnormal platelets results in a decrease in plasma high-molecular-weight multimers evident on multimeric analysis. Platelet-type vWD can be distinguished from type 2B vWD by demonstrating that the defect lies in the patient's platelets and not in vWF. Platelets from patients with platelet-type vWD have been shown to bind increased amounts of normal vWF at low concentrations of ristocetin. This allows separation of these two disorders, as platelets from patients with type 2B bind normal amounts of vWF. In addition, plasma from patients with type 2B vWD may support aggregation of normal platelets at low ristocetin concentrations. The clinical management of patients with platelet-type vWD differs from those with type 2B vWD; patients with platelet-type vWD should be treated with platelets for control of bleeding episodes.

Bernard-Soulier Syndrome

The Bernard-Soulier syndrome is a rare, autosomal-recessive disorder characterized by the onset of mucocutaneous bleeding at an early age, a prolonged bleeding time, and thrombocytopenia with giant platelets (5,21,27 and 27). The Bernard-Soulier syndrome is caused by a deficiency or functional abnormality of the GP Ib/IX/V complex, resulting in abnormal platelet adhesion. However, thrombin also binds to the GP Ib/IX/V receptor and thus thrombin-mediated aggregation is abnormal in patients with Bernard-Soulier syndrome. This effect is distinct from the major thrombin receptor, as platelet aggregation induced by peptides capable of activating PAR-1 is normal in platelets from patients with Bernard-Soulier syndrome.

The GP Ib/IX/V complex is composed of at least four proteins — GP Ibα, GP Ibß, GP IX and GP V in a stoichiometry of 2:2:2:1 (Fig. 47.4). Each protein is coded for by a distinct gene, but all belong to the leucine-rich motif superfamily of proteins. GP Ibα and GP Ibß are covalently linked by a disulfide bond in the extracellular region near the transmembrane domains; failure to form this linkage impairs expression and function of the receptor. Mutations in the genes for GP Ibα, GP Ibß and GP IX all have been associated with the Bernard-Soulier syndrome. In most patients, the mutations result in the inability to express the receptor complex on the cell surface. In a few patients, a mutation in GP Ibα results in a loss of affinity for vWF and impaired function even though the receptor is expressed on the cell surface.

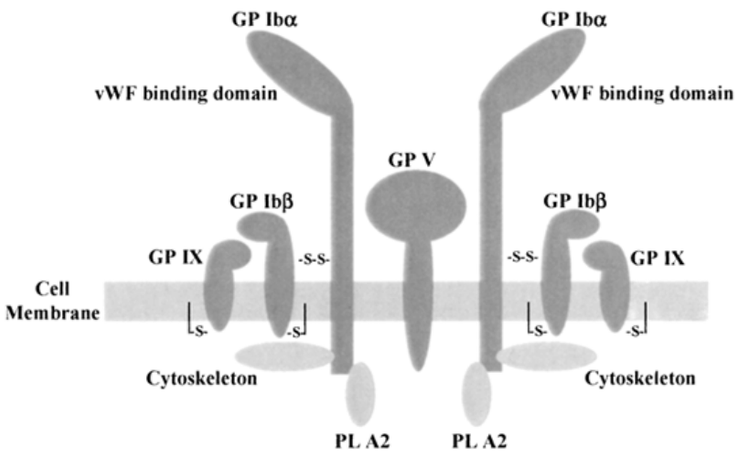


FIGURE 47.4. The GP Ib/IX/V receptor mediates platelet adhesion through interaction with von Willebrand factor (vWF). The receptor complex is composed of four distinct proteins, GP Ibα, GP Ibß, GP IX and GP V in a stoichiometry of 2:2:2:1. All four proteins appear to be necessary for normal expression and function of the receptor. The binding site for vWF is in an acidic extracellular domain of GP Ibα near the amino terminal end of the molecule. The extracellular region of GP Ibα nearest the cell membrane contains a mucinlike domain that is rich in O-linked carbohydrates. The intracellular domain of GP Ibα interacts with elements of the cytoskeleton, including actin binding protein, and with phospholipase A2 (PL A2). Ligand binding initiates the biochemical signaling pathways through these protein-protein interactions. GP Ibß is covalently linked to GP Ibα through a disulfide bond (-S-S-) in the extracellular domain. The short intracellular domain contains a free cysteine (-S), which can be linked to membrane lipids. GP IX is not covalently linked to the complex, but interacts with GP Ibß. The intracytoplasmic tail of GP IX also includes a free cysteine that can bind to membrane lipids. As with each of the proteins in the complex, GP V contains leucine-rich repeats. Glycoprotein V contains a thrombin susceptible site and the molecule may be cleaved by thrombin during platelet activation, releasing a 69 kDa fragment that includes most of the extracellular domain.

Initial laboratory evaluation of patients with the Bernard-Soulier syndrome usually demonstrates a mild to moderate thrombocytopenia with giant platelets, a prolonged bleeding time and a normal prothrombin time (PT) and activated partial thromboplastin time (APTT). The aggregation response to agonists such as ADP, epinephrine, collagen, and arachidonic acid is normal. However, platelets from patients do not aggregate in response to ristocetin. Addition of normal plasma or a source of vWF does not correct the ristocetin-induced aggregation defect, as it does in a patient with vWD. Analysis of vWF shows normal levels of vWF:ag and normal vWF:Rcof activity. The multimeric composition of vWF in plasma from patients with Bernard-Soulier syndrome also is normal (Table 47.3). The absence of the GP Ib/IX/V complex on the patient's platelets can be documented with flow cytometry using monoclonal antibodies directed at the receptor complex. Therapy of patients with Bernard-Soulier syndrome revolves around transfusion of platelets to control bleeding episodes. Alloimmunization is a problem in those patients requiring frequent transfusions.

Collagen Receptor Deficiency

Congenital deficiency of platelet membrane receptors for collagen results in impaired platelet adhesion and aggregation (5,28). Platelet adhesion to collagen appears to augment the process initiated by vWF binding to GP Ib/IX/V and is necessary for adequate function *in vivo*. At least two membrane proteins appear to be involved in collagen binding. The initial interaction appears to be between collagen and glycoprotein Ia/IIa, with subsequent binding to GP VI. Patients with abnormalities of collagen binding usually have a prolonged bleeding time, mild to moderate mucocutaneous bleeding, and defects in the platelet aggregation response to collagen. The platelet response to ristocetin and vWF concentration and function are normal.

Connective Tissue Disorders

Patients who have congenital disorders of connective tissue, such as the Ehlers-Danlos syndrome, may have problems with mucocutaneous bleeding (5,29). Bleeding times in such patients are often mildly prolonged. *In vitro* assessment of platelet function usually shows normal results and normal levels of vWF. The defect in such patients appears to be in the vessel wall components that mediate platelet adhesion.

Disorders of Platelet Activation

A number of congenital disorders that affect platelet activation have been described (5,14,15,30,31,32,33,34,35,36,37 and 38). These patients typically present with a clinical history characteristic of platelet dysfunction, with the onset of mucocutaneous bleeding early in life. The family history is often positive, suggesting an autosomal-dominant pattern of inheritance. However, the precise inheritance pattern for some of these disorders is still poorly defined because only a few kindred have been studied. The bleeding time is typically prolonged, but, as in vWD, it may be within the upper range of normal. The aspirin tolerance test can be used to unmask an underlying platelet defect in this group of patients. The PFA 100 appears to be an alternate, sensitive method for detecting abnormal platelet function in this group of disorders.

Laboratory studies characteristically show normal vWF concentration and function and a normal platelet aggregation response to ristocetin. Platelet aggregation with ADP and epinephrine typically shows a secondary wave defect, characterized by an initial wave of aggregation followed by dissociation of the platelets (Figure 47.1). The response to arachidonic acid and collagen may be blunted or delayed as well. A number of specific disorders, as briefly described below, have been delineated, but platelet transfusion therapy for bleeding episodes is the common therapy for all of these disorders. Although the mechanism is unknown, desmopressin has been shown to be effective in some patients and may be a useful adjunctive therapy.

Disorders of platelet activation can be divided into platelet membrane receptor disorders, storage pool disorders, and biochemical signaling disorders (Table 47.4). Deficiency of GP Ia, a collagen receptor, is associated with an abnormal aggregation response to collagen and defective platelet adhesion. Deficiency of α_2-adrenergic receptors (epinephrine receptors) has been associated with a mild bleeding tendency and failure of platelets to aggregate in response to epinephrine *in vitro*. In many of these patients, the aggregation response to ADP, collagen, and arachidonic acid has been normal. Patients with an abnormality of the ADP receptors show decreased binding of ADP, a decreased aggregation response to ADP and functional changes that mimic the defect seen with platelets exposed to the platelet inhibitor clopidogrel. Abnormal platelet function because of mutations in the receptor for TxA_2 also has been described. As one would expect, the aggregation response to arachidonic acid is abnormal in such patients.

TABLE 47.4. CONGENITAL PLATELET DISORDERS ASSOCIATED WITH ABNORMAL PLATELET ACTIVATION

Membrane Receptor Disorders	Storage Pool Disorders	Biochemical Signaling Disorders
Collagen receptor deficiency	Hermansky-Pudlak syndrome	Cyclooxygenase deficiency
α_2-Adrenergic receptor deficiency	Wiscott-Aldrich syndrome	Thromboxane synthetase deficiency
ADP receptor deficiency	Chediak-Higashi syndrome	Decreased arachidonic-acid release
TxA_2 receptor abnormality	Idiopathic dense granule deficiency	Phospholipase C disorders
	Gray platelet syndrome	G-protein disorders
	Quebec platelet disorder	Abnormal phosphatidylinositol metabolism Idiopathic

Abnormalities of platelet α and dense granules result in what are known as the storage pool disorders. This is a heterogeneous group of disorders that have been associated with clinically significant platelet dysfunction. The abnormalities may affect dense granules, α-granules, or both types of granules. Deficiency of dense granules has been associated with congenital disorders such as the Hemansky-Pudlak syndrome and the Wiskott-Aldrich syndrome, but may occur independent of such disorders (Table 47.5). The Hermansky-Pudlak syndrome is an autosomal-recessive disorder characterized by tyrosinase-positive ocular-cutaneous albinism, mucocutaneous bleeding as a result of absent dense granules and lysosomal accumulation and storage of ceroid-lipofuscin. The clinical course often is complicated by granulomatous colitis and/or progressive pulmonary fibrosis. The disorder is related to genetic abnormalities affecting the formation and function of melanosomes, platelet dense gran-ules and lysosomes. While mutations at distinct genetic loci

have been associated with the syndrome, the mechanism for the organelle abnormalities has not been elucidated. Platelet aggregation studies from patients with Hermansky-Pudlak syndrome characteristically show a secondary wave defect and lack of secretion of serotonin and ATP in response to ADP and epinephrine.

TABLE 47.5. ACQUIRED ABNORMALITIES OF PLATELET FUNCTION

Medications	Multiple myeloma
Antiplatelet antibodies	Monoclonal proteins
Uremia	Macromolecules
Cardiopulmonary bypass	Hypothyroidism
Consumptive coagulopathies	Liver disease
Acquired von Willebrand disease	

The Wiskott-Aldrich syndrome is an X-linked recessive disorder characterized by immunodeficiency, platelet abnormalities, eczema, and an increased risk of autoimmune disorders and malignancy. Platelet abnormalities include thrombocytopenia, small platelets, a decreased number of platelet granules and an abnormal platelet aggregation response to ADP, epinephrine and collagen. Megakaryocytes are present in normal or mildly increased numbers and platelet turnover is decreased, consistent with ineffective thrombopoiesis. The pathogenesis of the Wiskott-Aldrich syndrome has been linked to abnormalities of a protein (WASP) that is involved in transmitting signals from the cell membrane to the actin cytoskeleton.

The Chediak-Higashi syndrome is an autosomal recessive disorder characterized by partial oculocutaneous albinism, predisposition to infection, an immunodeficiency that predisposes to the development of lymphoproliferative disorders and abnormal granules in a variety of cell types, including platelets. Chediak-Higashi syndrome is caused by a defect in the lysosomal trafficking regulator protein (LYST). The precise mechanisms by which defects in this protein lead to the clinical manifestations of the disorder are still unknown.

The gray-platelet syndrome is a rare, autosomal-recessive disorder characterized by mild to moderate thrombocytopenia and pale-appearing platelets on a Wright-Giemsa–stained peripheral blood film. The abnormal platelet morphology is caused by an inability to store soluble proteins in the α-granules. In contrast, platelet α-granule membrane proteins such as P-selectin (CD62) are present in relatively normal concentration and translocate to the surface of the platelets following platelet activation. The aggregation response of platelets from patients with the gray-platelet syndrome usually show a normal response to epinephrine and arachidonic acid but a decreased response to thrombin and low concentrations of collagen. The aggregation response to ADP and ristocetin is variable. A common finding among patients with the gray-platelet syndrome is the early onset of myelofibrosis, perhaps resulting from an inability of megakaryocytes to normally store platelet-derived growth factors.

The Quebec platelet disorder is a rare, autosomal-dominant disorder characterized by moderate to severe mucocutaneous bleeding that is not responsive to platelet transfusion. The platelet count in patients has varied from normal to moderately reduced. Platelet aggregation in response to epinephrine is abnormal, but the response to ADP and collagen may be normal or mildly reduced. There is a deficiency of multimerin, a normal α-granule protein, and proteolysis of many, but not all α-granule proteins. Platelet factor V, fibrinogen, vWF, thrombospondin, P-selectin, osteonectin, and fibronectin are all abnormally degraded. The degraded fibrinogen can be detected by assays for fibrin degradation products if whole blood is allowed to clot before the serum is separated. However, the degraded fibrinogen is not detected with assays for D-dimer or if platelet-poor plasma is allowed to clot and used for the assay, suggesting that the abnormal platelets are able to secrete the degraded fibrinogen *in vitro*. The therapy of choice for bleeding episodes in patients with the Quebec platelet disorder appears to be fibrinolytic inhibitors, such as ε-aminocaproic acid.

The biochemical signaling abnormalities represent a heterogeneous group of disorders that share a common clinical phenotype characterized by mucocutaneous bleeding, a normal platelet count and abnormalities in the platelet aggregation response to one or more agonists. Deficiency of either cyclooxygenase or thromboxane synthetase in the phospholipase A_2 pathway results in deficient synthesis of thromboxane A_2. Such patients typically show no aggregation response to arachidonic acid and a secondary wave-type defect in response to ADP. Defects in the phospholipase C pathway may lead to decreased formation of inositol trisphosphate and diacylglycerol, leading to decreased mobilization of calcium from storage sites. Defects in this pathway include abnormalities of specific isotypes of phospholipase C and abnormalities of the G-proteins that couple receptor activation to the signaling pathway. Abnormalities in downstream signaling pathways also may lead to an abnormal platelet response.

The biochemical signaling disorders can be differentiated from storage pool disorders with specialized studies such as electron microscopic evaluation of platelet granules, measurement of serotonin uptake and release, measurement of α-granule proteins and measurement of ATP release in response to strong agonists such as thrombin or thrombin receptor agonists. Typically, patients with dense granule deficiency show a decreased uptake of radiolabeled serotonin, a decreased number of granules on electron microscopy, and decreased release of ATP in response to thrombin. Distinction between these types of disorders generally does not affect the clinical management of these patients, which, in most situations, involves transfusion of platelets to control significant bleeding episodes.

Disorders of Platelet Aggregation

Glanzmann's thrombasthenia is the major disorder affecting platelet-platelet interaction (5,14,39,40). Glanzmann's thrombasthenia is a rare, autosomal-recessive disorder characterized by a prolonged bleeding time, recurrent mucocutaneous bleeding, and an absent aggregation response to ADP, epinephrine, collagen, and arachidonic acid (primary wave defect). The aggregation response to ristocetin often is normal in these patients. Glanzmann's thrombasthenia is caused by a deficiency or functional abnormality of the GP IIb/IIIa receptor. Because of the defect, the platelets cannot bind fibrinogen or other adhesive proteins and thus cannot be cross-linked to form platelet aggregates. The diagnosis is established by the clinical history, a long

bleeding time, and the characteristic lack of platelet aggregation with normal fibrinogen levels. The platelet surface concentration of GP IIb/IIIa measured by flow cytometry is significantly reduced in most patients. The concentration of the receptor may be normal in a few patients, but it is unable to function normally.

Clot retraction usually is diminished in samples from patients with Glanzmann's thrombasthenia because of the interaction of GP IIb/IIIa with the platelet cytoskeleton. The α-granules of patients with Glanzmann's thrombasthenia are deficient in fibrinogen, indicating that the GP IIb/IIIa receptor is involved in the normal uptake of fibrinogen from plasma. As the human platelet antigens PL^{A1}, Bak^{a}, and Lek^{a} are located on the glycoprotein IIb/IIIa complex, patients with Glanzmann's thrombasthenia usually are negative for these antigens and are thus susceptible to alloimmunization. Alloimmunization can be a major clinical problem in the management of patients requiring frequent transfusions.

A variety of mutations involving the genes for GP IIb and IIIa have been identified in patients with Glanzmann's thrombasthenia. Affected patients may be homozygous for a single type of mutation or a compound heterozygote for two mutations. Although the incidence of Glanzmann's thrombasthenia is rare in the general population, there are selected populations in which it may occur as commonly as hemophilia, including French gypsies, Iraqi Jews, Jordanian nomads and South Indian Hindus. A relatively high prevalence of a characteristic mutation is found in each of these populations.

Patients with afibrinogenemia also may have problems with mucocutaneous bleeding and may lack a response to agonists such as ADP, collagen, arachidonic acid, and epinephrine. The defect in platelet aggregation in these cases is from the absence of the fibrinogen as a cross-linking material. The clinical picture in such patients usually is dominated by the major defect in fibrin clot formation. As one would expect, the PT and APTT are basically not measurable in such patients.

Disorders Affecting Platelet Procoagulant Activity

Activation of the platelet by its various physiologic agonists leads to calcium dependent, transient expression of negatively charged aminophospholipids, particularly phosphatidylserine, on the external platelet surface (5,14,41). The aminophospholipids provide the requisite surface for many of the coagulation activation complexes described in Chapter 48. Scott syndrome is a rare, autosomal-recessive disorder characterized by impaired formation of factor Xa and thrombin in the presence of platelets. The defect is related to a lack of expression of aminophospholipids on the platelet surface following platelet activation. A similar defect is present in red cells and lymphocytes from patients with Scott syndrome. The pathogenesis of the disorder is still uncertain, as the cells from patients with Scott syndrome have a normal, functional concentration of the “scramblase” enzyme that mediates the translocation of the phospholipid. Other parameters of platelet function, such as platelet aggregation to ADP and collagen, dense and α-granule content and platelet secretion are normal in patients with Scott syndrome. However, platelets from patients with Scott syndrome do not shed microparticles following activation. The binding of annexin V to platelets is dependent on phospholipid translocation and thus annexin V does not bind to platelets from patients with Scott syndrome following platelet activation. This provides a method for screening for the disorder using flow cytometry and measurement of the binding of annexin V labeled with fluorescent dye following *in vitro* platelet activation.

Constitutive expression of phosphatidylserine on nonactivated platelets has been described in three members of a family (Stormorken syndrome). As one would expect, platelets from these patients bound annexin V without platelet activation. In addition, circulating platelet microparticles were increased. Platelet aggregation in response to collagen was abnormal and borderline in response to ADP. Other findings in the affected family members included asplenia, reduced platelet survival, miosis, dyslexia, muscle fatigue, and icthyosis. Because of the continuous expression of procoagulant phospholipids, one might expect these patients to suffer from thromboembolic problems. However, the principal hemostatic problem has been recurrent bleeding. Despite the continuous expression of procoagulant phospholipids, the defects in platelet function appear to dominate the clinical picture.

ACQUIRED DISORDERS OF PLATELET FUNCTION

Part of "47 - Laboratory Evaluation of Platelet Disorders"

Acquired abnormalities of platelet function are a more common cause of clinical bleeding than congenital disorders. A number of conditions have been associated with abnormal platelet function (Table 47.5). Among these, drug-induced abnormalities are the most common cause of abnormal platelet function. Indeed, intentional inhibition of platelet function has become an important component in the management of patients with cardiovascular disease. In contrast to congenital disorders, in which platelet aggregation studies play a pivotal role in delineating the nature of the platelet defect, platelet aggregation studies have only a minor role in evaluating acquired platelet dysfunction. In most patients, the combination of the clinical history and a prolonged bleeding time or other screening test of platelet function are sufficient to establish the platelet abnormality; platelet aggregation studies do not add any specific or clinically predictive information in most patients.

Medications That May Affect Platelet Function

Numerous drugs and foods have been found to have an inhibitory effect on platelet function either *in vivo* or *in vitro* (Table 47.6) (10,42,43,44,45,46,47,48,49,50,51,52 and 53). Major groups of drugs that affect platelet function include nonsteroidal antiinflammatory agents, of which aspirin is the most common offender; ß-lactam-containing antibiotics; a variety of cardiovascular drugs, particularly ß-blockers; psychotropic drugs; anesthetics; antihistamines; and some chemotherapeutic agents. Aspirin is taken by many patients on a routine basis and is also a common ingredient in many over-the-counter medications. Therefore, a history of aspirin ingestion may be difficult to elicit unless detailed questioning is performed. Antibiotics are often an overlooked source of

impaired platelet function in hospitalized patients and can have an adverse effect on *in vivo* bleeding tendencies.

TABLE 47.6. SUBSTANCES ASSOCIATED WITH ABNORMAL PLATELET FUNCTION IN VIVO OR IN VITRO

Antiinflammatory agents	*ß-Blockers*	*Platelet Inhibitors*
Aspirin	Propranolol	Ticlopidine
Indomethacin	Labetalol	Clopidogrel
Ibuprofen		Dipyridamole
Phenylbutazone	*Calcium-channel blockers*	Abciximab
Sulfinpyrazone	Verapamil	Tirofiban
Naproxen	Nifedipine	Eptifibatide
Diclofenac	Diltiazem	
Zompirac		*Miscellaneous*
Tolmetin	*Lipid-lowering drugs*	Ethanol
	Clofibrate	Sodium valproate
Antibiotics	Halofenate	Hydralazine
Ampicillin	Cyproheptadine	Dextran
Penicillin		Papaverine
Carbenicillin	*Antihistamines*	Suloctidil
Ticarcillin	Theophylline	Furosemide
Mezlocillin	Aminophylline	Ethacrynic acid
Piperacillin	Chlorpheniramine	Acetazolamide
Cephalosporins	Diphenhydramine	Hydroxychloroquine
		Hydrocortisone
Tricyclic antidepressants	*Foods*	Methylprednisolone
Imipramine	Garlic	Daunorubicin
Desipramine	N-3 fatty acids	Mithramycin
Amitriptyline	Ginger	Chlorpromazine
Nortriptyline	Onion	Triprolidine
	Cumin	Reserpine
	Turmeric	Nitrofurantoin
		Heparin

In contrast to the “inadvertent” and generally undesirable effect of these drugs on platelet function, inhibitors of platelet function now are used frequently as part of the therapy of patients with cardiovascular disease. This group of drugs includes aspirin, an inhibitor of cyclooxygenase; ticlopidine, and clopidogrel, ADP receptor antagonists; dipyridamole, a phosphodiesterase inhibitor; and abciximab, eptifibatide and tirofiban, inhibitors of GP IIb/IIIa. The most potent of these agents are the GP IIb/IIIa inhibitors, which block the “final common pathway” of platelet aggregation.

Ticlopidine has been associated with a variety of toxic side effects, including gastrointestinal disturbance, skin rash, severe neutropenia, aplastic anemia, and thrombotic thrombocytopenic purpura (TTP). In view of the seriousness of the side effects, therapy with ticlopidine must be monitored closely. Clopidogrel has a safer profile in terms of bone marrow toxicity but it also has been associated with the development of TTP in a few patients.

Inhibition of greater than 80% of the platelet GP IIb/IIIa receptors is necessary for the GP IIb/IIIa inhibitors to have a significant clinical effect while inhibition of 100% of the receptors may be associated with clinical bleeding. Thus, there is a relatively narrow therapeutic window with the GP IIb/IIIa inhibitors. Flow cytometric assays have been developed to determine the percent of receptors inhibited, but it is unclear whether these assays provide a significant benefit in the management of patients being treated with these agents. Eptifibatide and tirofiban have a short half-life, both in terms of plasma concentration and inhibitor activity. Consequently, their antiplatelet effect is short-lived and essentially limited to the period of infusion. In contrast, abciximab, which is a chimeric monoclonal antibody fragment, has a short plasma half-life but a prolonged platelet-associated half-life. This appears to be from persistent binding of the antibody fragment to platelet GP IIb/IIIa. Thus, the antiplatelet effect of abciximab persists well beyond the period of active infusion. An uncommon, but clinically significant, side effect of the GP IIb/IIIa inhibitors is the sudden onset of severe thrombocytopenia, with platelet counts decreasing to less than 20×10^9/l. The mechanism of thrombocytopenia is uncertain but may be related to naturally occurring antibodies directed at neoepitopes on GP IIb/IIIa exposed when the drug binds to the receptor.

Uremia

Uremia is associated with a complex coagulopathy dominated by abnormal platelet function (46, 54, 55). A number of biochemical changes in platelets have been described in patients with uremia, but the precise lesion(s) responsible for the clinical manifestations is still unclear. Anemia contributes significantly to the bleeding tendency associated with uremia, particularly when the hemoglobin is less than 90 to 100 g/L. Studies have documented that the bleeding time and bleeding tendency both decrease if the hemoglobin concentration is raised by transfusion of red blood cells or administration of erythropoietin. Other defects that have been reported in uremia include: decreased adhesion of platelets to subendothelium, decreased availability of platelet factor 3 (platelet procoagulant activity), increased cytoplasmic c-AMP

levels, alterations in cytoplasmic calcium and a diminished rise in calcium following stimulation by a variety of platelet agonists, decreased storage products in both dense and α-granules, and alterations in vessel wall prostacyclin synthesis. Platelet aggregation studies in patients with uremia have been variable, with some groups reporting essentially normal aggregation responses and others reporting markedly impaired aggregation responses. In most studies, the aggregation responses observed *in vitro* have not correlated with clinical manifestations.

In addition to the platelet abnormalities listed above, there is evidence of ongoing activation of the coagulation cascade and fibrinolysis in patients with renal failure, even in the absence of hemodialysis. For example, the levels of thrombin-antithrombin complexes and prothrombin fragment F1.2 are increased, consistent with activation of the coagulation cascade. Fibrin degradation products and plasmin-antiplasmin complexes are also increased, consistent with ongoing activation of fibrinolysis. Thus, there may be a component of a low-grade consumptive coagulopathy in patients with uremia. The low-grade hemostatic activation could lead to platelet secretion without consumption or removal from the circulation and to modification of membrane receptors by plasmin. Such “exhausted” platelets would be expected to be functionally impaired.

Dialysis improves platelet function in patients with uremia and is an important part of the therapeutic approach to these patients. In addition to dialysis and correction of the anemia, a number of other therapeutic options for managing clinical bleeding in these patients have been developed. Cryoprecipitate has been reported to decrease the bleeding time and bleeding severity. The mechanism of action of cryoprecipitate is unclear, but it is relatively rich in vWF, which could promote platelet adhesion. Desmopressin has also been associated with a decrease in the bleeding time and correction of the bleeding tendency in patients with uremia. More recently, conjugated estrogens have been used in these patients with a beneficial effect.

Antiplatelet Antibodies

The most common clinical manifestation of antiplatelet antibodies is thrombocytopenia (56,57 and 58). However, antibodies that bind to the platelet surface also can cause abnormalities of platelet function. Decreased platelet function resulting from antiplatelet antibodies is most commonly associated with chronic autoimmune thrombocytopenic purpura (chronic ITP). The precise mechanism of abnormal platelet function in these patients is still unclear, but probably is related to interaction of the antibody with platelet surface membrane glycoproteins, as GP Ib/IX/V and IIb/IIIa are frequent epitopes for antibodies in chronic ITP. It has been suggested that antibody binding to these critical membrane receptors may impair their physiological function.

Binding of platelet antibodies to the surface of the platelet also may lead to platelet activation and thrombosis. This has been most commonly associated with HIT, but has also been noted with heparin-independent antiplatelet antibodies. The platelet activation is often a result of binding of immune complexes to the platelet receptor for IgG, FcγRIIa. Binding of immune complexes clusters the FcγRIIa receptors, initiating an intracytoplasmic signaling cascade that leads to platelet aggregation and secretion. Complement activation on the platelet surface can also lead to platelet activation and aggregation.

Myeloproliferative Disorders

A variety of platelet defects have been described in patients with underlying acute or chronic myeloproliferative disorders (46, 59). In some patients, there is evidence of increased platelet reactivity, as indicated by an increase in circulating platelet aggregates, plasma platelet factor 4, β-thromboglobulin and increased *in vitro* aggregation responses to platelet agonists. In other patients, platelet function appears to be impaired. A variety of changes in platelet structure and function may be observed in platelets from patients with myeloproliferative disorders, including a decrease in the α and dense granules; abnormalities of granule morphology; a decrease in platelet vWF and fibrinogen; an alteration in platelet membrane glycoproteins; decreased uptake and release of serotonin; decreased aggregation in response to a variety of agonists; and decreased procoagulant activity. Impaired platelet function may also be suggested by a discrepancy between the platelet count and the bleeding time.

Clinical manifestations of abnormal platelet function in these patients have included both thromboembolic and hemorrhagic complications. Unfortunately, there are no laboratory tests that can determine with accuracy which patient is likely to have thromboembolic or hemorrhagic complications. Therefore, there is little use for platelet aggregation studies in evaluating these patients. The one setting in which platelet aggregation studies may be helpful is in assessing a patient with marked thrombocytosis. Many patients with essential thrombocythemia have an abnormal *in vitro* aggregation response, particularly when epinephrine is used as the agonist, whereas patients with reactive thrombocytosis often have normal aggregation studies, unless there is a medication or other medical condition that could affect platelet function.

Fibrinolysis

Enhanced fibrinolysis can impair platelet function through a variety of mechanisms, including fibrin degradation products (FDPs) inhibiting fibrinogen binding to GP IIb/IIIa and plasmin-mediated degradation of membrane GP Ib/IX/V and GP IIb/IIIa (46). Abnormal platelet function resulting from enhanced fibrinolysis is perhaps most commonly seen in patients undergoing cardiopulmonary bypass procedures. The resulting impairment in platelet function probably plays a significant role in post-operative bleeding in such patients. A similar defect may be seen in patients treated with a systemic fibrinolytic agents such as streptokinase or tissue plasminogen activator and in patients with disseminated intravascular coagulation (DIC).

Acquired von Willebrand Disease

Acquired vWD is an uncommon clinical disorder characterized by the sudden onset of a bleeding diathesis in a previously asymptomatic patient (60, 61). As with other platelet disorders, the bleeding tends to be mucocutaneous in nature, with gas

trointestinal bleeding a frequent manifestation in this disorder. Acquired vWD has been seen in patients with severe congenital vWD following replacement therapy, lymphoproliferative disorders, myeloproliferative disorders, nonhematologic malignancies, autoimmune disorders, hypothyroidism, certain medications and spontaneously in the elderly. The pathogenesis of acquired vWD may involve one or more of several mechanisms. A common mechanism is the development of an antibody to vWF that may lead to increased clearance of vWF from the circulation and/or inhibition of vWF function. Another mechanism is adsorption of circulating vWF by binding of vWF to tumor cells. In some cases, this abnormal adsorption of vWF has been linked to aberrant expression of GP Ibα on the tumor cells. Other mechanisms that have been proposed include enhanced proteolysis of vWF and decreased synthesis of vWF.

The laboratory findings in acquired vWD are quite variable and somewhat dependent on the pathogenesis; most commonly the bleeding time is prolonged, and vWF, as measured by vWF:ag and vWF:Rcof, is decreased. The factor VIII level is usually concordant with the vWF:ag level in these patients. Some patients with acquired vWD may show a type 2 pattern of vWD; in these patients there may be a discordance between vWF:ag and vWF:Rcof. Multimeric analysis of such patients demonstrates the characteristic decrease in high-molecular-weight multimers of vWF. An inhibitor of vWF function can be demonstrated in some patients with immune-mediated acquired vWD. However, inhibitor activity cannot be demonstrated in patients with acquired vWD from nonimmune mechanisms or nonneutralizing antibodies.

In patients with an underlying malignancy, therapy of the tumor often is associated with regression of the acquired vWD. In those who have a documented inhibitor of vWF, immunosuppression forms the cornerstone of therapy. Various agents have been used to achieve this, including steroids, cytotoxic drugs, and intravenous immunoglobulin. Replacement of vWF with cryoprecipitate is occasionally helpful in such patients during periods of bleeding.

Other Disorders

Abnormal platelet function has been associated with macromolecules, including high concentrations of monoclonal IgM and high-molecular-weight dextrans used for volume expansion. Patients with myeloma and high concentrations of monoclonal IgG occasionally have abnormalities of platelet function as well. On occasion, patients with hypothyroidism or end-stage liver disease may have impaired platelet function.

QUANTITATIVE PLATELET DISORDERS

Part of "47 - Laboratory Evaluation of Platelet Disorders"

Evaluation of Quantitative Platelet Disorders

Thrombocytopenia and thrombocytosis are common clinical findings. Thrombocytopenia generally is of greater clinical concern because of its association with an increased risk of bleeding. The major concern with thrombocytosis is distinguishing reactive thrombocytosis from thrombocytosis secondary to a myeloproliferative disorder (thrombocythemia), which may be associated with either bleeding or thromboembolic manifestations. Thrombocytopenia usually is classified by the underlying mechanism, whereas thrombocytosis usually is divided into reactive processes and hematologic malignancies.

Thrombocytopenia may be from increased destruction, decreased production, spurious platelet counting, or splenic sequestration (Table 47.7). Thrombocytopenia is most commonly an acquired problem, but there are congenital disorders associated with either decreased production or increased clearance of platelets. Patients with congenital thrombocytopenia may be misdiagnosed (and mistreated) as having autoimmune thrombocytopenic purpura (ITP).

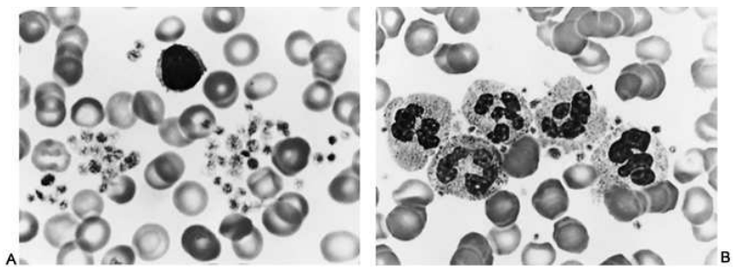


FIGURE 47.5. Pseudothrombocytopenia may be caused by *in vitro* agglutination of platelets (**A**) or platelet satellitosis (**B**). *In vitro* platelet agglutination may be EDTA dependent or independent. Both platelet agglutination and platelet satellitosis may affect the accuracy of the white blood cell count as well as the platelet count (original magnification × 500). It is thought that these findings are related to antibodies to platelet epitopes exposed by chelation of calcium by EDTA.

TABLE 47.7. MECHANISMS OF THROMBOCYTOPENIA

Increased Destruction	**Decreased Production**
Immune-mediated destruction	*Congenital disorders*
Autoimmune	Amegakaryocytic thrombocytopenia
Alloimmune (posttransfusion purpura, neonatal)	Fanconi's anemia
Drug-dependent antiplatelet antibodies	Wiskott-Aldrich
Immune complex disease	Alport syndrome and variants
	May-Hegglin anomaly
Microangiopathic processes	Familial thrombocytopenia
Thrombotic thrombocytopenic purpura	Bernard-Soulier syndrome
Hemolytic-uremic syndrome	Gray platelet syndrome
Disseminated intravascular coagulation	Quebec platelet syndrome
Sepsis	Neonatal rubella
Preeclampsia/eclampsia	
Metastatic carcinoma	*Bone marrow replacement*
Malignant hypertension	Metastatic tumor
Vascular tumors and lesions (including artificial valves)	Leukemia
	Myeloma
Cardiopulmonary bypass	Storage disorders (e.g., Gaucher's)
	Granulomatous disease
Major burns	
	Bone marrow failure
Massive transfusion	Aplastic anemia
	Radiation
Abnormalities of von Willebrand factor or GPIba	Chemotherapy
	Drugs
Hypersplenism	Immune suppression
	Viral infections
Spurious thrombocytopenia	
EDTA-dependent antibodies	*Ineffective thrombopoiesis*
Temperature-dependent platelet agglutinins	Megaloblastic anemia
Platelet satellitosis	Alcoholism
	Myeloproliferative syndromes
	Myelodysplastic syndromes

The laboratory approach to the evaluation of alterations in the platelet count begins with a careful review of the clinical history and the peripheral blood morphology, with particular emphasis on platelet morphology. Analysis of the mean platelet volume (MPV), platelet volume distribution, and examination of the bone marrow are additional useful steps (Table 47.1). Measurement of platelet RNA content by flow cytometry may be helpful in identifying a shift to “younger” platelets with a higher RNA concentration.

The initial goal in the evaluation of thrombocytopenia is to determine if the cause is related to increased peripheral destruction or decreased bone marrow production. The peripheral blood film should be carefully reviewed for evidence of platelet clumps or platelet satellitosis, giant platelets, and the presence of abnormal, hypogranular platelets (Fig. 47.5). Morphologic evidence of other hematologic abnormalities such as myelodysplasia, acute leukemia, megaloblastic anemia, autoimmune hemolytic anemia, TTP, or the May-Hegglin anomaly, should also be sought.

The MPV is another parameter that can provide useful information in the initial evaluation of thrombocytopenia. The normal bone marrow response to thrombocytopenia is an increase in megakaryocyte cytoplasmic size, associated with an increase in the size of zones of demarcation, resulting in the production of larger platelets. Accordingly, an increase in the MPV is often a hallmark of active thrombopoiesis and may be used as an indicator of the marrow response to thrombocytopenia associated with increased peripheral destruction. On the other hand, thrombocytopenia associated with a low MPV suggests a decreased marrow response to the thrombocytopenia (decreased production). Often the review of the smear combined with the clinical history and MPV will suggest a specific etiology.

If the etiology of the altered platelet count is not evident from the clinical history and peripheral blood findings, a bone marrow aspirate and biopsy should be examined to evaluate hematopoiesis in general and megakaryocyte number and morphology in detail. Megakaryocyte number and growth pattern are assessed most accurately by bone-marrow biopsy. In reactive conditions associated with an increase in megakaryocytes, the megakaryocytes tend to be randomly spaced throughout the marrow and do not form tight clusters or colonies. In contrast, clusters of megakaryocytes that may have a normal or strikingly abnormal morphology are commonly seen in bone marrow samples from patients with myeloproliferative disorders and abnormal megakaryopoiesis (Fig. 47.6). Both the bone marrow biopsy and the bone marrow aspirate can yield useful information re

garding megakaryocyte morphology, particularly if a myeloproliferative disorder is suspected. The bone marrow aspirate is particularly useful for assessing whether megakaryocytes are normally shedding platelets. Finally, the bone marrow aspirate and biopsy should be reviewed carefully for evidence of other diseases that may affect platelet production.

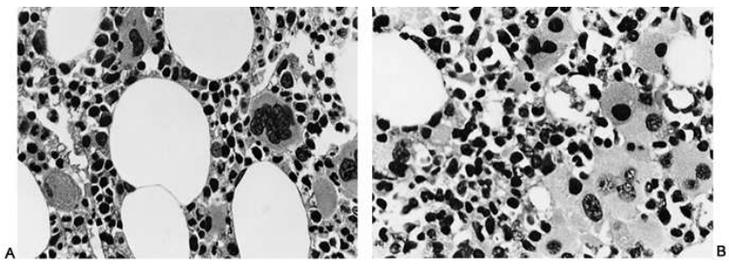


FIGURE 47.6. Megakaryocytic hyperplasia in reactive conditions is usually characterized by an increase in megakaryocytes throughout the marrow without formation of large clusters (**A**). In contrast, megakaryocytic hyperplasia associated with myeloproliferative disorders is often characterized by formation of clusters of closely approximated megakaryocytes (**B**). In addition, megakaryocyte morphology is frequently abnormal in patients with myeloproliferative disorders.

Detection of platelet-associated antibodies has been frequently used to assess the possibility of immune-mediated platelet destruction. However, a consensus panel of experts has indicated that measurement of platelet-associated IgG is not helpful in the routine evaluation of adult patients with autoimmune thrombocytopenic purpura (ITP). This recommendation was based on the lack of sensitivity and specificity of assays for measurement of platelet-associated antibody. Informative assays, such as those that measure surface-bound immunoglobulin through detection of the “free” Fc component of the antibody and those that measure an interaction with a specific platelet membrane protein, may be helpful in selected cases.

Additional procedures may be helpful on occasion, depending on the suspected underlying disorder. Such procedures may include electron-microscopy, cytogenetics, bone-marrow culture assays, and enzyme assays (e.g., Gaucher's disease) to evaluate platelet production and hematopoietic clonality. Other specialized assays may be used to confirm the diagnosis of disorders such as HIT, TTP, or paroxysmal nocturnal hemoglobinuria.

Thrombocytopenia Due to Increased Destruction

A number of disease processes may lead to increased platelet destruction (Table 47.7). In clinical practice, antibody-mediated thrombocytopenia is one of the more common mechanisms. Antibodies may induce thrombocytopenia through one or more of several mechanisms, including autoimmune, drug-dependent, alloimmune, and immune complex pathways. Microangiopathic processes, including TTP, the hemolytic-uremic syndrome (HUS) and disseminated intravascular coagulation (DIC) are characterized by accelerated platelet consumption, microvascular thrombosis and red cell fragmentation (schistocytes). A major side effect of cardiopulmonary bypass is transient thrombocytopenia resulting from both hemodilution and consumption.

Autoimmune Thrombocytopenia

ITP is the prototypic example of immune-mediated destruction of platelets (57, 62,63,64,65 and 66). Clinically, ITP is often divided into two forms, acute/childhood and chronic/adult (Table 47.8). Acute ITP often is associated with an antecedent viral infection and it has been suggested that the platelet is an innocent casualty of a cross-reacting antibody that recognizes both the pathogen and platelet membrane. In contrast, chronic ITP appears to be more of a classic autoimmune disorder and is often associated with other immunologic disorders such as systemic lupus erythematosus, rheumatoid arthritis, and autoimmune thyroid disease. As indicated in Table 47.8, the clinical course of these two forms of ITP is quite different, with acute ITP having an excellent prognosis and a high rate of spontaneous remission. In contrast, chronic ITP tends to be long-lasting, often requiring long-term immunosuppressive therapy.

TABLE 47.8. COMPARISON OF ACUTE AND CHRONIC AUTOIMMUNE THROMBOCYTOPENIA

Parameter	Acute ITP	Chronic ITP
Age	All ages; most frequent in childhood	All ages; most frequent in young adults
Gender	F = M	F > M
Onset	Abrupt	Gradual
Platelet count	Often <50 × 10^9/L	Often >50 × 10^9/L
Other autoimmune disorders	Usually absent	Often present
Clinical remission	Frequent	Infrequent

In recent years, the target antigens for the antiplatelet antibodies have been identified in many patients with ITP. The most frequent target antigens are GP Ib/IX/V and GP IIb/IIIa. In some patients the antibodies react with only one of these recep

tor complexes, while in other patients the antibodies react with both complexes. Although antibodies commonly bind to these proteins, they only occasionally cause a clinical picture of acquired Bernard-Soulier (GP 1b/V/IX) or acquired Glanzmann's thrombasthenia (GP IIb/IIIa).

Typical laboratory features of ITP include moderate to marked thrombocytopenia with a normal to increased MPV. Platelet morphology on peripheral blood smear tends to be normal, except for the shift to slightly larger platelets. Bone marrow examination usually demonstrates normal to increased numbers of megakaryocytes, which have a normal morphology. Bone marrow examination is most useful in patients who are greater than 60 years of age, in whom there is a higher incidence of myeloproliferative disorders, or in patients who are being considered for splenectomy. The bleeding time in patients with acute ITP is often proportional to the decrease in platelet count or is even shorter than expected. In contrast, in patients with chronic ITP, the bleeding time may be more prolonged than predicted by the decrease in platelet count, reflecting interference with platelet function by the platelet-associated antibody. However, changes in the bleeding time are not diagnostic for ITP, nor predictive of the clinical manifestations of the disorder. Therefore, performance of the bleeding time is not recommended as part of the evaluation of patients for ITP. Assays for platelet-associated IgG are not sufficiently sensitive or specific to aid in the diagnosis of ITP. The most important diagnostic step is to rule out other causes of thrombocytopenia. This may include specific evaluation for other disorders that may present with thrombocytopenia, such as human immunodeficiency virus (HIV) infection or Type 2B vWD. The diagnosis of ITP is thus based on the clinical history, physical examination, complete blood count and careful review of the peripheral blood film. In the absence of atypical findings or results that suggest another diagnosis, no further evaluation is usually needed.

Alloimmune Thrombocytopenia

Immune-mediated platelet destruction also is found in post-transfusion purpura, neonatal thrombocytopenia, and thrombocytopenia refractory to random platelet transfusions (57, 67). Post-transfusion purpura is an uncommon disorder characterized by the sudden onset of thrombocytopenia 7 to 10 days following a blood transfusion. The thrombocytopenia is from immune-mediated destruction of the patient's own platelets. The disorder usually occurs in patients with the PLA2 (HPA-1b) phenotype who receive blood products from PLA1 (HPA-1a) positive donors. The affected patients frequently have antibodies to PLA1, as well as antibodies to other platelet membrane proteins. Development of HLA-related antibodies and specific antiplatelet antibodies underlies the development of refractoriness to platelet transfusion. Therapeutic options in such patients include utilization of HLA-matched platelets and actual cross-matching of potential donor platelets with patient serum. Neonatal thrombocytopenia is usually a consequence of a maternal antibody that crosses the placenta and recognizes the neonatal platelets as foreign. Neonatal thrombocytopenia also may occur in the setting of maternal chronic ITP. Nonimmune neonatal thrombocytopenia is commonly seen in patients with hemolytic disease of the newborn and should be differentiated from these disorders.

Drug Induced Thrombocytopenia

Over 100 drugs have been implicated in the development of thrombocytopenia (58, 68,69,70, 71 and 72). In most cases, the platelet destruction is a result of the development of a drug-dependent antibody. At least three potential mechanisms for immunologic destruction in this setting have been described. Quinidine-related thrombocytopenia is an example of the mechanism in which drug binds to and alters platelet surface glycoproteins (GP Ib/IX/V), leading to immunologic destruction. Penicillin is an example of a second mechanism, in which the drug serves as a hapten that binds to the platelet surface. The hapten induces an antibody response and the IgG then binds to the membrane-associated hapten, leading to premature removal of the platelet from the circulation. A third mechanism is through immune complex interaction with the Fc receptor on the platelet surface (FcγRIIa); HIT is an example of this third mechanism.

Heparin-induced thrombocytopenia is one of the major side effects of heparin therapy. Two forms of heparin-related thrombocytopenia have been identified; the first is characterized by a mild drop in the platelet count (rarely less than 100×10^9/L), while the second is associated with more severe thrombocytopenia. The drop in the platelet count in the mild form usually occurs shortly after heparin therapy is instituted and the platelet count returns to normal whenever heparin therapy is discontinued. The platelet count may even rise toward the preheparin platelet count even if heparin therapy is continued. The mechanism for this form of heparin-related thrombocytopenia is unclear, but no immunologic component has been identified. Patients with this form of heparin-related thrombocytopenia do not appear to be at increased risk of either bleeding or thromboembolic complications.

The more severe type of heparin-related thrombocytopenia, known as HIT, is caused by an antibody that binds to a complex composed of heparin and a protein, usually PF4. When heparin binds to soluble PF4 a neoepitope is exposed on PF4. Antibodies, usually IgG, form to this eiptope and bind to the complex when heparin and PF4 are in the correct relative concentration. The heparin/PF4/IgG immune complex binds to platelet FcγRIIa receptors, clustering the receptors by the multivalent binding of the immune complex. The FcγRIIa cytoplasmic domain has a signaling motif that is activated when the receptor is clustered. This triggers activation of biochemical signaling pathways, platelet activation and platelet aggregation. In addition, the activated platelets shed membrane phosphoplipid particles (platelet microparticles) that support thrombin generation. Consequently, platelet activation in HIT also is associated with enhanced thrombin formation. These pathways are thought to account for the thromboembolic complications that commonly occur with this syndrome. Alternate terms that have been used to describe HIT include heparin-induced thrombocytopenia-thrombosis (HITT), white clot syndrome, and heparin-associated thrombocytopenia (HAT). The diagnostic and therapeutic approaches to this disorder are discussed in detail in Chapter 48.

Thrombocytopenia Associated with Human Immunodeficiency Virus

Thrombocytopenia is a common complication in patients with HIV infection (65). The mechanism of thrombocytopenia in these patients appears to be multifactorial. In many patients antiplatelet antibodies form, and the thrombocytopenia in these patients is similar to the chronic form of ITP. It has been hypothesized that formation of immune complexes, particularly between viral proteins and antibodies, also may promote thrombocytopenia in these patients. Finally, there is some evidence that the virus may infect megakaryocytes and that thrombopoiesis may be decreased in patients with HIV infection. Substantiating this viewpoint is the increase in platelet count that occurs in many patients following the introduction of antiviral therapy.

Thrombotic Thrombocytopenic Purpura and Hemolytic-Uremic Syndrome

Thrombotic thrombocytopenic purpura and hemolytic-uremic syndrome (HUS) are distinct but related disorders characterized by thrombocytopenia, microangiopathic hemolytic anemia with frequent schistocytes on the peripheral blood smear, and microvascular thrombosis (43, 73,74,75,76,77,78,79 and 80). The pathogenesis of TTP has been linked to abnormalities of the protease that normally modifies plasma vWF, whereas development of HUS has been commonly linked to certain infections, particularly specific serotypes of *Escherichia coli*. A TTP/HUS-like syndrome also may be seen with certain drugs, including mitomycin, cisplatin, cyclosporin, ticlopidine, and clopidogrel. A similar microangiopathic hemolytic anemia with microvascular thrombosis also has been associated with HIV infection.

The pathogenesis of TTP appears to be related to the presence of abnormally large vWF multimers. Plasma vWF is derived from endothelial cells, where it is stored in Weibel-Palade bodies. Newly secreted vWF includes very large multimers that can mediate platelet adhesion, activation and aggregation, particularly in areas of high shear. A plasma protease normally cleaves these very large multimers into smaller forms once the vWF is released into the circulation. Acquired TTP appears to be related to the development of an IgG that inhibits the function of this protease, permitting persistence of the abnormally large vWF multimers. A familial form of TTP has been described and it is associated with a congenital deficiency of the plasma protease. Thus, congenital or acquired loss of protease function appears to result in TTP.

In contrast to TTP, the plasma activity of the vWF-cleaving protease is normal in patients with HUS. Hemolytic-uremic syndrome is seen most commonly in the pediatric age range and usually is preceded by a phase of diarrhea, which is often bloody. The prodromal gastroenteritis usually is caused by a verocytotoxin-producing strain of *E. coli* (VTEC), most commonly strain O157:H7. Verocytotoxin produced by the VTEC in the gastrointestinal track can enter the systemic circulation and bind to endothelial cells, especially after the endothelial cells have been activated by inflammatory mediators, such as tumor necrosis factor-α, that increase the expression of a binding site for verocytotoxin. The bound verocytotoxin can damage endothelial cells, initiating the procoagulant response mechanisms. Thus, microvascular endothelial-cell injury is thought to be an underlying mechanism for HUS.

The classic pentad of clinical findings in TTP consists of fever, fluctuating neurological deficits, thrombocytopenia, red-cell fragmentation, and renal failure. One of the major clinical differences between HUS and TTP is that neurological symptoms tend to be absent and renal failure more pronounced in HUS. Laboratory parameters typical of both HUS and TTP include marked thrombocytopenia, often with giant platelets, frequent schistocytes and nucleated red blood cells in the peripheral blood, increased reticulocyte count and lactate dehydrogenase (LD), particularly LD1, and a normal PT, APTT, fibrinogen, and antithrombin. Assays for fibrin degradation products often are negative but may show a mild increase. Bone marrow examination, when performed, shows normal to increased numbers of megakaryocytes. Marrow vessels occasionally show the characteristic microthrombi, however, this does not occur frequently enough to warrant performing a bone marrow biopsy for the purpose of identifying microthrombi. As the disease progresses, evidence of a consumptive coagulopathy may emerge. The LD level is useful in monitoring the response to therapy, with a decreasing LD being a favorable sign. A rebound increase in the LD may signify an early relapse of the disorder.

Plasma exchange (apheresis) with fresh frozen plasma appears to be the therapy of choice in these patients. In the absence of prompt therapeutic intervention, the mortality rate for TTP approaches 80%; with prompt institution of plasma exchange, the mortality rate probably is less than 20%. Vincristine has been used as an adjunct in some patients who have not responded promptly to plasma exchange. In small series of patients, this appears to be an effective additional agent. Steroids, aspirin therapy, and splenectomy remain controversial in these disorders. Relapses occur in up to 25% to 30% of patients who respond to the initial therapy, necessitating continued monitoring of these patients. Relapses in TTP correspond to a persistent lack or inhibition of vWF-cleaving protease activity.

Other Mechanisms of Increased Platelet Destruction

Increased platelet destruction may be seen in association with a number of other clinical disorders. Sepsis, particularly with Gram-negative bacteria, is one of the more common causes of increased platelet destruction in hospitalized patients. Many of these patients also have a component of acute DIC as well. Diagnosis in these patients depends on recognition of the infectious or consumptive process and the characteristic signs and symptoms of the systemic inflammatory response syndrome.

Cardiopulmonary bypass is associated with an acute drop in the platelet count during the procedure. The post-cardiopulmonary bypass platelet count is often 50% to 60% of the baseline count because of the combination of hemodilution and consumption. The platelet count tends to return to the preoperative level over the first 1 to 2 days following surgery. In most patients, the platelet count following bypass is sufficient to support hemostasis; platelet-type bleeding in these patients usually is a result of a qualitative defect rather than a lack of platelets. However, for patients who start the procedure with borderline platelet

counts, the decrease in platelet count may compromise hemostatic plug formation in the immediate post-operative period.

A number of snake venoms and other biological toxins may lead to a decrease in the platelet count. In these clinical settings, there is often a component of a DIC associated with the envenomation. This may be assessed by measuring fibrinogen, fibrin degradation products, prothrombin time, activated partial thromboplastin time, as well as the platelet count.

A variety of vascular malformations and tumors have been associated with increased platelet consumption due to the abnormal blood flow through the deformed vessels. A classic example of this is the Kasabach-Merritt syndrome, which usually presents in the pediatric age range. Cardiac valve disease and artificial cardiac valves also may lead to a shortened platelet half-life and overt thrombocytopenia in some patients.

Functional hypersplenism also is commonly associated with low peripheral platelet count. The problem in these patients is one of sequestration rather than true thrombocytopenia. Administration of epinephrine or other vasoactive compounds can lead to splenic contraction and release of platelets from the spleen. When this is done, the platelet count may transiently rise into the normal range. Therefore, the total peripheral platelet mass tends to be normal but with an abnormal distribution.

Thrombocytopenia Due to Decreased Production

Decreased platelet production occurs in a variety of clinical settings (Table 47.7). The characteristic findings in these patients are thrombocytopenia with a normal to decreased MPV and abnormal bone marrow findings. Additional abnormalities evident on the peripheral blood smear may indicate the nature of the underlying disorder.

Congenital Disorders Associated With Decreased Platelet Production

Thrombocytopenia has been associated with a number of congenital disorders, some of which also demonstrate abnormal platelet function (Table 47.9) (5, 26, 36, 38, 81,82,83,84,85,86,87,88 and 89). In some patients, the accompanying clinical manifestations suggest the diagnosis while in other patients the congenital nature of the thrombocytopenia may be missed. Such patients are often misdiagnosed as having ITP and may receive inappropriate immunosuppressive therapy. Therefore, the possibility of a congenital cause of thrombocytopenia should be considered before establishing the diagnosis of ITP.

Amegakaryocytic thrombocytopenia is an autosomal-recessive disorder characterized by nearly total absence of megakaryocytes from the marrow. Amegakaryocytic thrombocytopenia may be associated with frequent episodes of leukemoid reactions and skeletal abnormalities, particularly involving the radii. This later form is designated thrombocytopenia with absent radii or TAR syndrome. Platelet size (MPV) tends to be normal to mildly increased; platelet survival is normal; and the bleeding time is proportional to the platelet count. There may be multiple genetic defects that give rise to amegakaryocytic thrombocytopenia; the TAR syndrome has not been associated with any abnormalities of the receptor for thrombopoietin (c-Mpl) whereas amegakaryocytic thrombocytopenia without skeletal malformations has been associated with mutation of the c-Mpl gene.

Fanconi's anemia (congenital aplastic anemia) is an autosomal-recessive disorder characterized by severe marrow hypoplasia. Thrombocytopenia may dominate the clinical picture, and death from hemorrhage is a frequent complication of this disorder. The platelets in Fanconi's anemia have a normal MPV and half-life; the bleeding time is proportional to the platelet count. Megakaryocytes and other hematopoietic elements in the bone marrow are markedly decreased. Fanconi's anemia has been associated with a variety of cytogenetic abnormalities and may terminate as a hematopoietic neoplasm. Thrombocytopenia can also be a component of other congenital hematopoietic stem-cell defects.

Congenital thrombocytopenia has been described in several families with familial nephritis and high-frequency deafness (Alport syndrome). The clinical manifestations have varied between families, some showing the classic association between nephritis and deafness, and others showing only thrombocytopenia and nephritis. This syndrome is inherited as an autosomal-dominant trait and usually has been associated with macrothrombocytes (increased MPV). The bone marrow usually shows a normal numbers of megakaryocytes. There is little information concerning platelet survival in these patients. Platelet function is quite variable between families. Some have a bleeding time proportional to the thrombocytopenia, while others have a relative prolongation of the bleeding time for the degree of thrombocytopenia. Variants of Alport syndrome include Epstein syndrome, characterized by severe thrombocytopenia in addition to the other features of Alport syndrome; Fechtner syndrome, characterized by neutrophil inclusions resembling Dohle bodies (and the May-Hegglin anomaly) in addition to other features of Alport syndrome; and Sebastian syndrome, characterized by thrombocytopenia, neutrophil inclusions, and lack of the other features of Alport syndrome.

The May-Hegglin anomaly is an autosomal-dominant trait characterized by macrothrombocytes (increased MPV) and the presence of Döhle body-like inclusions within neutrophils. The bone marrow in patients with the May-Hegglin anomaly shows a normal number of megakaryocytes, and platelet survival in these patients is also normal. The bleeding time is variable and may be disproportionately prolonged for the degree of thrombocytopenia, suggesting additional qualitative platelet defects in some patients. It has been suggested that the total circulating platelet mass is normal in patients with May-Hegglin anomaly from the increase in the MPV.

Familial thrombocytopenia is an autosomal-dominant trait that has been occasionally reported in the literature. The clinical and laboratory features in these families have been quite variable. Platelet size has varied from normal to increased. Bone marrow examination has shown normal to decreased numbers of megakaryocytes. Platelet survival, when examined, has been quite variable. Usually the bleeding time in affected patients is proportional to the platelet count, suggesting that this is primarily a problem of platelet production. These cases usually have an

isolated thrombocytopenia without other significant clinical or laboratory findings.

TABLE 47.9. CONGENITAL THROMBOCYTOPENIAS

Disorder	MPV	Inheritance	Platelet Function	Other Laboratory Findings	Other Clinical Findings
Bernard-Soulier syndrome	↑	Autosomal dominant	Abnormal	Decreased GP Ib/IX/V, Decreased RIPA	Mucocutaneous bleeding
Type 2B vWD	N or ↑	Autosomal dominant	Abnormal	Decreased HMW vWF multimers, increased RIPA	Mucocutaneous bleeding
Platelet-type vWD	N or ↑	Autosomal dominant	Abnormal	Decreased HMW vWF multimers, increased RIPA, abnormal GP Ibα	Mucocutaneous bleeding
Gray platelet syndrome	↑	Autosomal dominant	Abnormal	Decreased platelet α-granule proteins	Mucocutaneous bleeding
Quebec platelet disorder	N or ↑	Autosomal dominant	Abnormal	Decreased aggregation to epinephrine, proteolysis of α-granule proteins	Mucocutaneous bleeding
May-Hegglin anomaly	↑	Autosomal dominant	Normal	Neutrophil inclusions (Dohle bodies)	Nephritis, familial spastic paralysis
Alport syndrome	↑	Autosomal dominant	Normal or abnormal	Abnormal secondary wave aggregation	Nephritis, high-frequency hearing loss
Epstein syndrome	↑	Autosomal dominant	Normal or abnormal	Alport syndrome with severe thrombocytopenia	Nephritis, high-frequency hearing loss
Sebastian syndrome	↑	Autosomal dominant	Normal	Neutrophil inclusions	Variant of Alport syndrome
Fechtner syndrome	↑	Autosomal dominant	Normal	Neutrophil inclusions	Nephritis, high-frequency hearing loss
Montreal platelet syndrome	↑	Autosomal dominant	Abnormal	Spontaneous aggregation, decreased aggregation response to thrombin, decreased calpain activity	Mucocutaneous bleeding
Wiskott-Aldrich syndrome	↓	X-linked recessive	Abnormal	Decreased aggregation to ADP, epinephrine, collagen; abnormal B and T cell function	Immunodeficiency, eczema, autoimmune diseases, malignancies
X-linked micro-thrombocytopenia	↓	X-linked recessive	Abnormal	Mild storage pool defect	Variant of Wiskott-Aldrich, without nonplatelet manifestations
Paris-Trousseau platelet disorder		Autosomal dominant	Abnormal	Megakaryocytes increased, micromegakaryocytes, giant α-granules in some platelets, partial deletion of chromosome 11 (del(11)q23.3→qter)	Mild bleeding tendency
Fanconi's anemia	N	Autosomal recessive		Decrease megakaryocytes, bone marrow aplasia	Propensity to develop hematologic malignancies, mucocutaneous bleeding

Abbreviations used: MPV, mean platelet volume; RIPA, ristocetin-induced platelet aggregation; N, normal; HWM, high-molecular-weight

Drug Induced Suppression of Platelet Production

Perhaps the most common cause of decreased platelet production among hospitalized patients today is chemotherapy. Transient thrombocytopenia is a common complication of intensive chemotherapy. Thrombocytopenia also is commonly seen in alcoholics. The etiology of the decreased platelet count in this setting is often multi-factorial, including splenic sequestration, folate deficiency, and a direct toxic effect of ethanol on thrombopoiesis. Ionizing radiation can also affect platelet production as well as overall hematopoiesis production. A number of nonchemotherapeutic drugs, including tetracycline, ganciclovir, and chloramphenicol, may affect the marrow, leading to hematopoietic hypoplasia and thrombocytopenia.

Other Disorders Associated With Decreased Platelet Production

A number of viral illnesses have been associated with thrombocytopenia, including hepatitis, parvovirus, and HIV. It has long been suspected that decreased platelet production can be caused by autoimmune mechanisms, either humoral or cellular. However, documentation of the mechanism in such patients has proven to be quite difficult. The bone marrow in these patients usually shows a marked decrease in megakaryocytes, similar to congenital amegakaryocytic thrombocytopenia. Nutritional deficiencies such as folate or vitamin B12 deficiency also may be as

sociated with significant thrombocytopenia. Infiltrative processes in the bone marrow may result in replacement of normal hematopoietic elements and thrombocytopenia. Such infiltrative processes may be non-neoplastic, (e.g., Gaucher's disease, granulomatous disease), or neoplastic (e.g., acute leukemia). Metastatic tumor also may replace hematopoietic bone marrow, leading to thrombocytopenia; common tumors associated with this problem include breast and prostate carcinoma.

Thrombocytosis

An increase in the peripheral platelet count may occur in a variety of clinical settings (Table 47.10). The mechanism of thrombocytosis in many of these cases is not known. In some cases (e.g., alcohol abuse) the thrombocytosis appears to be a rebound phenomenon following recovery from toxic suppression of thrombopoiesis. The clinical and laboratory evaluation of thrombocytosis revolves around the determination of the associated condition and separation of a reactive thrombocytosis from a neoplastic thrombocythemia.

TABLE 47.10. CONDITIONS ASSOCIATED WITH THROMBOCYTOSIS

Reactive Thrombocytosis
Acute inflammatory disorders
Chronic inflammation or infections
Hemolytic anemia
Acute hemorrhage
Rebound following suppression (e.g., ethanol chemotherapy, etc.)
Postsplenectomy
Exercise
Iron deficiency
Trauma
Drugs (vincristine, epinephrine)
Malignancy
Neoplastic Thrombocytosis (thrombocythemia)
Essential thrombocythemia
Chronic myelogenous leukemia
Polycythemia vera
Myelofibrosis with myeloid metaplasia

Thrombocytosis is a common finding in myeloproliferative disorders. If a myeloproliferative disorder is suspected, a bone marrow examination, cytogenetics, and red blood cell mass determination may be helpful in delineating the specific disorder. The presence of the Philadelphia chromosome is indicative of chronic myelogenous leukemia, even though the peripheral blood may have the appearance of essential thrombocythemia. Platelet function may be abnormal in association with myeloproliferative disorders, leading to either bleeding or thromboembolic complications.

REFERENCES

1. Carr ME, Jr. In vitro assessment of platelet function. *Transfus Med Rev* 1997;11:106–115.

2. Peterson P, Hayes TE, Arkin CF, et al. The preoperative bleeding time test lacks clinical benefit: College of American Pathologists' and American Society of Clinical Pathologists' position article. *Arch Surg* 1998;133:134–139.

3. Rodgers RP, Levin J. A critical reappraisal of the bleeding time. *Semin Thromb Hemost* 1990;16:1–20.

4. Batlle J, Torea J, Rendal E, et al. The problem of diagnosing von Willebrand's disease. *J Intern Med Suppl* 1997;740:121–128.

5. Nurden AT. Inherited abnormalities of platelets. *Thromb Haemost* 1999;82:468–480.

6. Vanhoorelbeke K, Cauwenberghs N, Vauterin S, et al. A reliable and reproducible ELISA method to measure ristocetin cofactor activity of von Willebrand factor. *Thromb Haemost* 2000;83:107–113.

7. Cattaneo M, Lecchi A, Agati B, Lombardi R, Zighetti ML. Evaluation of platelet function with the PFA-100 system in patients with congenital defects of platelet secretion. *Thromb Res* 1999;96:213–217.

8. Harrison P, Robinson MS, Mackie IJ, et al. Performance of the platelet function analyser PFA-100 in testing abnormalities of primary haemostasis. *Blood Coagul Fibrinolysis* 1999;10:25–31.

9. Kerenyi A, Schlammadinger A, Ajzner E, et al. Comparison of PFA-100 closure time and template bleeding time of patients with inherited disorders causing defective platelet function. *Thromb Res* 1999;96:487–92.

10. Kottke-Marchant K, Powers JB, Brooks L, et al. The effect of antiplatelet drugs, heparin, and preanalytical variables on platelet function detected by the platelet function analyzer (PFA-100). *Clin Appl Thromb Hemost* 1999;5:122–130.

11. Michelson AD. Flow cytometry: a clinical test of platelet function. *Blood* 1996;87:4925–4936.

12. Michelson AD, Furman MI. Laboratory markers of platelet activation and their clinical significance. *Curr Opin Hematol* 1999;6:342–348.

13. Rand ML, Carcao MD, Blanchette VS. Use of the PFA-100 in the assessment of primary, platelet-related hemostasis in a pediatric setting. *Semin Thromb Hemost* 1998;24:523–529.

14. Rao AK. Congenital disorders of platelet function: disorders of signal transduction and secretion. *Am J Med Sci* 1998;316:69–76.

15. Rao AK, Gabbeta J. Congenital disorders of platelet signal transduction. *Arterioscler Thromb Vasc Biol* 2000;20:285–289.

16. Schmitz G, Rothe G, Ruf A, et al. European Working Group on Clinical Cell Analysis: Consensus protocol for the flow cytometric characterisation of platelet function. *Thromb Haemost* 1998;79:885–896.

17. Ginsburg D. Molecular genetics of von Willebrand disease. *Thromb Haemost* 1999;82:585–591.

18. Mannucci PM. Platelet von Willebrand factor in inherited and acquired bleeding disorders. *Proc Natl Acad Sci USA* 1995;92:2428–2432.

19. Mannucci PM. Treatment of von Willebrand disease. *Haemophilia* 1998;4:661–664.

20. Phillips MD, Santhouse A. von Willebrand disease: recent advances in pathophysiology and treatment. *Am J Med Sci* 1998;316:77–86.

21. Andrews RK, Shen Y, Gardiner EE, et al. The glycoprotein Ib-IX-V complex in platelet adhesion and signaling. *Thromb Haemost* 1999;82:357–364.

22. Kunishima S, Heaton DC, Naoe T, et al. De novo mutation of the platelet glycoprotein Ib alpha gene in a patient with pseudo-von Willebrand disease. *Blood Coagul Fibrinolysis* 1997;8:311–315.

23. Miller JL. Platelet-type von Willebrand disease. *Thromb Haemost* 1996;75:865–869.

24. Moriki T, Murata M, Kitaguchi T, et al. Expression and functional characterization of an abnormal platelet membrane glycoprotein Ib alpha (Met239 → Val) reported in patients with platelet-type von Willebrand disease. *Blood* 1997;90:698–705.

25. Ware J. Molecular analyses of the platelet glycoprotein Ib-IX-V receptor. *Thromb Haemost* 1998;79:466–478.

26. Lopez JA, Andrews RK, Afshar-Kharghan V, et al. Bernard-Soulier syndrome. *Blood* 1998;91:4397–418.

27. Mhawech P, Saleem A. Inherited giant platelet disorders. Classification and literature review. *Am J Clin Pathol* 2000;113:176–190.

28. Arai M, Yamamoto N, Moroi M, et al. Platelets with 10% of the normal amount of glycoprotein VI have an impaired response to collagen that results in a mild bleeding tendency. *Br J Haematol* 1995;89:124–130.

29. Grenko RT, Burns SL, Golden EA, et al. Type IV Ehlers-Danlos syndrome with aspirin sensitivity. A family study. *Arch Pathol Lab Med* 1993;117:989–992.

30. Brantly M, Avila NA, Shotelersuk V, et al. Pulmonary function and high-resolution CT findings in patients with an inherited form of pulmonary fibrosis, Hermansky-Pudlak syndrome, due to mutations in HPS-1. *Chest* 2000;117:129–136.

31. Certain S, Barrat F, Pastural E, et al. Protein truncation test of LYST reveals heterogenous mutations in patients with Chediak-Higashi syndrome. *Blood* 2000;95:979–983.

32. Dell'Angelica EC, Aguilar RC, Wolins N, et al. Molecular characterization of the protein encoded by the Hermansky-Pudlak syndrome type 1 gene. *J Biol Chem* 2000;275:1300–1306.

33. Hayward CP, Rivard GE, Kane WH, et al. An autosomal dominant, qualitative platelet disorder associated with multimerin deficiency, abnormalities in platelet factor V, thrombospondin, von Willebrand factor, and fibrinogen and an epinephrine aggregation defect. *Blood* 1996;87:4967–4978.

34. Hayward CP, Cramer EM, Kane WH, et al. Studies of a second family with the Quebec platelet disorder: evidence that the degradation of the alpha-granule membrane and its soluble contents are not secondary to a defect in targeting proteins to alpha-granules. *Blood* 1997;89:1243–1253.

35. McNicol A, Israels SJ. Platelet dense granules: structure, function and implications for haemostasis. *Thromb Res* 1999;95:1–18.

36. Nonoyama S, Ochs HD. Characterization of the Wiskott-Aldrich syndrome protein and its role in the disease. *Curr Opin Immunol* 1998;10:407–412.

37. Smith MP, Cramer EM, Savidge GF. Megakaryocytes and platelets in alpha-granule disorders. *Baillieres Clin Haematol* 1997;10:125–148.

38. Thrasher AJ, Burns S. Wiskott-Aldrich syndrome: a disorder of haematopoietic cytoskeletal regulation. *Microsc Res Tech* 1999;47:107–113.

39. Kato A. The biologic and clinical spectrum of Glanzmann's thrombasthenia: implications of integrin alpha IIb beta 3 for its pathogenesis. *Crit Rev Oncol Hematol* 1997;26:1–23.

40. Markovitch O, Ellis M, Holzinger M, et al. Severe juvenile vaginal bleeding due to Glanzmann's thrombasthenia: case report and review of the literature. *Am J Hematol* 1998;57:225–227.

41. Solum NO. Procoagulant expression in platelets and defects leading to clinical disorders. *Arterioscler Thromb Vasc Biol* 1999;19:2841–2846.

42. Baker RI, Hankey GJ. Antiplatelet drugs. *Med J Aust* 1999;170:379–382.

43. Bennett CL, Weinberg PD, Rozenberg-Ben-Dror K, et al. Thrombotic thrombocytopenic purpura associated with ticlopidine. A review of 60 cases. *Ann Intern Med* 1998;128:541–544.

44. Bennett CL, Davidson CJ, Raisch DW, et al. Thrombotic thrombocytopenic purpura associated with ticlopidine in the setting of coronary artery stents and stroke prevention. *Arch Intern Med* 1999;159:2524–2528.

45. Coukell AJ, Markham A. Clopidogrel. *Drugs* 1997;54:745–750.

46. George JN, Shattil SJ. The clinical importance of acquired abnormalities of platelet function. *N Engl J Med* 1991;324:27–39.

47. Kereiakes DJ. Oral platelet glycoprotein IIb/IIIa inhibitors. *Coron Artery Dis* 1999;10:581–594.

48. Nurden AT, Poujol C, Durrieu-Jais C, et al. Platelet glycoprotein IIb/IIIa inhibitors: basic and clinical aspects. *Arterioscler Thromb Vasc Biol* 1999;19:2835–2840.

49. Quinn MJ, Fitzgerald DJ. Ticlopidine and clopidogrel. *Circulation* 1999;100:1667–1672.

50. Schafer AI. Effects of nonsteroidal antiinflammatory drugs on platelet function and systemic hemostasis. *J Clin Pharmacol* 1995;35:209–219.

51. Steinhubl SR, Tan WA, Foody JM, et al. Incidence and clinical course of thrombotic thrombocytopenic purpura due to ticlopidine following coronary stenting. EPISTENT Investigators. Evaluation of Platelet IIb/IIIa Inhibitor for Stenting. *JAMA* 1999;281:806–810.

52. Tcheng JE. Clinical challenges of platelet glycoprotein IIb/IIIa receptor inhibitor therapy: bleeding, reversal, thrombocytopenia, and retreatment. *Am Heart J* 2000;139:S38–45.

53. Zusman RM, Chesebro JH, Comerota A, et al. Antiplatelet therapy in the prevention of ischemic vascular events: literature review and evidence-based guidelines for drug selection. *Clin Cardiol* 1999;22:559–573.

54. Mezzano D, Tagle R, Panes O, et al. Hemostatic disorder of uremia: the platelet defect, main determinant of the prolonged bleeding time, is correlated with indices of activation of coagulation and fibrinolysis. *Thromb Haemost* 1996;76:312–321.

55. Weigert AL, Schafer AI. Uremic bleeding: pathogenesis and therapy. *Am J Med Sci* 1998;316:94–104.

56. Deckmyn H, Vanhoorelbeke K, Peerlinck K. Inhibitory and activating human antiplatelet antibodies. *Baillieres Clin Haematol* 1998;11:343–359.

57. Porcelijn L, von dem Borne AE. Immune-mediated thrombocytopenias: basic and immunological aspects. *Baillieres Clin Haematol* 1998;11:331–341.

58. Warkentin TE, Chong BH, Greinacher A. Heparin-induced thrombocytopenia: towards consensus. *Thromb Haemost* 1998;79:1–7.

59. Manoharan A, Gemmell R, Brighton T, et al. Thrombosis and bleeding in myeloproliferative disorders: identification of at-risk patients with whole blood platelet aggregation studies. *Br J Haematol* 1999;105:618–25.

60. Tefferi A, Nichols WL. Acquired von Willebrand disease: concise review of occurrence, diagnosis, pathogenesis, and treatment. *Am J Med* 1997;103:536–540.

61. van Genderen PJ, Michiels JJ. Acquired von Willebrand disease. *Baillieres Clin Haematol* 1998;11:319–330.

62. Anonymous. Diagnosis and treatment of idiopathic thrombocytopenic purpura: recommendations of the American Society of Hematology. The American Society of Hematology ITP Practice Guideline Panel. *Ann Intern Med* 1997;126:319–326.

63. George JN, Raskob GE. Idiopathic thrombocytopenic purpura: diagnosis and management. *Am J Med Sci* 1998;316:87–93.

64. George JN, Raskob GE. Idiopathic thrombocytopenic purpura: A concise summary of the pathophysiology and diagnosis in children and adults. *Semin Hematol* 1998;35:5–8.

65. Gillis S, Eldor A. Immune thrombocytopenic purpura in adults: clinical aspects. *Baillieres Clin Haematol* 1998;11:361–372.

66. Sutor AH, Gaedicke G. Acute autoimmune thrombocytopenia. *Baillieres Clin Haematol* 1998;11:381–389.

67. Bussel J, Kaplan C. The fetal and neonatal consequences of maternal alloimmune thrombocytopenia. *Baillieres Clin Haematol* 1998;11:391–408.

68. Amiral J, Meyer D. Heparin-induced thrombocytopenia: diagnostic tests and biological mechanisms. *Baillieres Clin Haematol* 1998;11:447–460.

69. George JN, Raskob GE, Shah SR, et al. Drug-induced thrombocytopenia: a systematic review of published case reports. *Ann Intern Med* 1998;129:886–890.

70. Rizvi MA, Shah SR, Raskob GE, George JN. Drug-induced thrombocytopenia. *Curr Opin Hematol* 1999;6:349–353.

71. Visentin GP, Malik M, Cyganiak KA, et al. Patients treated with unfractionated heparin during open heart surgery are at high risk to form antibodies reactive with heparin:platelet factor 4 complexes. *J Lab Clin Med* 1996;128:376–383.

72. Visentin GP. Heparin-induced thrombocytopenia: molecular pathogenesis. *Thromb Haemost* 1999;82:448–456.

73. Durand JM, Lefevre P. Mitomycin-induced thrombotic thrombocytopenic purpura: possible successful treatment with vincristine and cyclophosphamide. *Haematologica* 1991;76:421–423.

74. Furlan M, Robles R, Galbusera M, et al. von Willebrand factor-cleaving protease in thrombotic thrombocytopenic purpura and the hemolytic-uremic syndrome. *N Engl J Med* 1998;339:1578–1584.

75. Gerritsen HE, Turecek PL, Schwarz HP, et al. Assay of von Willebrand factor (vWF)-cleaving protease based on decreased collagen binding affinity of degraded vWF: a tool for the diagnosis of thrombotic thrombocytopenic purpura (TTP). *Thromb Haemost* 1999;82:1386–1389.

76. Mannucci PM. Thrombotic thrombocytopenic purpura: a simpler diagnosis at last? *Thromb Haemost* 1999;82:1380–1381.

77. Obert B, Tout H, Veyradier A, et al. Estimation of the von Willebrand factor-cleaving protease in plasma using monoclonal antibodies to vWF. *Thromb Haemost* 1999;82:1382–1385.

78. Ruggenenti P, Remuzzi G. Pathophysiology and management of thrombotic microangiopathies. *J Nephrol* 1998;11:300–310.

79. Tsai HM, Lian EC. Antibodies to von Willebrand factor-cleaving protease in acute thrombotic thrombocytopenic purpura. *N Engl J Med* 1998;339:1585–94.

80. van de Kar NC, Monnens LA. The haemolytic-uraemic syndrome in childhood. *Baillieres Clin Haematol* 1998;11:497–507.

81. Bellucci S. Megakaryocytes and inherited thrombocytopenias. *Baillieres Clin Haematol* 1997;10:149–162.

82. Ihara K, Ishii E, Eguchi M, et al. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. *Proc Natl Acad Sci USA* 1999;96:3132–3136.

83. Jantunen E. Inherited giant platelet disorders. *Eur J Haematol* 1994;53:191–196.

84. Kunishima S, Kojima T, Tanaka T, et al. Mapping of a gene for May-Hegglin anomaly to chromosome 22q. *Hum Genet* 1999;105:379–383.

85. Letestu R, Vitrat N, Masse A, et al. Existence of a differentiation blockage at the stage of a megakaryocyte precursor in the thrombocytopenia and absent radii (TAR) syndrome. *Blood* 2000;95:1633–1641.

86. Nichols KE, Crispino JD, Poncz M, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet* 2000;24:266–270.

87. Noris P, Spedini P, Belletti S, et al. Thrombocytopenia, giant platelets, and leukocyte inclusion bodies (May-Hegglin anomaly): clinical and laboratory findings. *Am J Med* 1998;104:355–360.

88. Strippoli P, Savoia A, Iolascon A, et al. Mutational screening of thrombopoietin receptor gene (c-mpl) in patients with congenital thrombocytopenia and absent radii (TAR). *Br J Haematol* 1998;103:311–314.

89. Zhu Q, Zhang M, Blaese RM, et al. The Wiskott-Aldrich syndrome and X-linked congenital thrombocytopenia are caused by mutations of the same gene. *Blood* 1995;86:3797–3804.

48

Coagulation Abnormalities

Douglas A. Triplett

Abnormalities of secondary hemostasis (i.e., coagulation proteins) typically present with a different clinical picture than abnormalities of primary hemostasis. In classical severe hemophilia the pattern of bleeding will be musculoskeletal and often “spontaneous.” However, in acquired disorders of secondary hemostasis, the pattern of bleeding may often present as mucocutaneous bleeding. A mixed pattern of bleeding reflects the complex pathophysiology of many acquired coagulopathies.

This chapter will address both hereditary and acquired disorders of coagulation proteins. The hereditary abnormalities will be discussed initially followed by the acquired disorders. The laboratory tests used to evaluate secondary hemostasis will be reviewed with special emphasis on interpretation.

HEREDITARY DISORDERS OF COAGULATION PROTEINS

Part of "48 - Coagulation Abnormalities"

Hemophilia A (Factor VIII Deficiency)

Hemophilia A is an X-linked inherited disorder of factor VIII (Antihemophiliac Globulin) (1). The incidence is approximately 1:5,000 males (1). As previously discussed, factor VIII is a critical cofactor in the intrinsic coagulation pathway at the level of factor X activation.

There are variable degrees of factor VIII deficiency that have arbitrarily been divided into severe, moderate, and mild (Table 48.1). Severe hemophilia A is characterized by “spontaneous” musculoskeletal bleeding. This pattern of bleeding is often first evident when the child begins to walk at 9 to 12 months of age. These patients also may present with bleeding following circumcision; however, it is not uncommon for a severely affected child to have minimal bleeding with this procedure. Mild and moderately affected patients may not present until much later in life. Mild hemophilia is often diagnosed when a patient presents with unexplained bleeding following surgery or in association with trauma. Occasionally, very mild hemophilia A patients may only complain of easy bruising or small superficial hematomata.

With the advent of molecular biological techniques, the genetics of hemophilia A have been extensively studied. In 1984, investigators first reported the successful cloning of the human factor VIII gene. The entire gene spans 186 Kb (103 base-pairs), which comprises approximately 0.1% of the X chromosome. There are a total of 26 exons and 25 introns. The coding DNA is 9 Kb, which codes for 2,351 amino acids. The first 19 amino acids are the secretory leader peptide of the nascent factor VIII. The complete amino-acid sequence now is established and the calculated molecular weight is 267,039 daltons. There are three homologous A domains, and two C domains as well as an intervening B domain, which is rich in carbohydrate (Fig. 48.1). Factor VIII is represented by the following formula: Al-A2-B-A3-Cl-C2. The A domains show significant homology with similar domains in ceruloplasmin.

Both deletions and point mutations have been found in patients with hemophilia A. Prior to the availability of molecular biology procedures, it was evident that a significant number of cases of hemophilia A occurred in families with a negative history. Consequently, it was anticipated that “hotspots” would be identified to account for *de novo* mutations. The “hotspots” appear to be related to CpG dinucleotides. The C can be methylated in the 5′ position of the pyrimidine ring and then deaminated to thymine. This leads to the CG → TG and CG → CA mutations. All of these mutations are recognized by TaqI recognition sites.

With the introduction of polymerase chain reaction (PCR), investigators were able to significantly increase their efforts to characterize the factor VIII gene in hemophilia patients. In 1992, Naylor et al. found factor VIII mRNA of all patients with hemophilia A had mutations of the factor VIII gene (2, 3). However, 45% of the patients with severe disease had an unusual mRNA defect. This defect prevented amplification of any segment that crossed the boundary between exons 22 and 23. This mutation results in inversions which affect patient mRNA. Both “proximal” and “distal” inversions have been described.

The laboratory diagnosis of hemophilia A is relatively straightforward. In the majority of cases, the activated partial thromboplastin time (APTT) will be prolonged while the prothrombin time (PT) and bleeding time (BT) will be within normal limits (Table 48.2). Factor VIII:C assays are necessary to establish the diagnosis. Patients with severe hemophilia A will have factor VIII:C levels, which are less than 1% (0.01 unit/mL). Factor assays are relatively easy to perform. However, the laboratory must have a strict quality assurance program to ensure accuracy and precision.

The treatment of hemophilia A has undergone rather dramatic changes in the last 30 years. Prior to the use of cryoprecipitate, the hemorrhaging patient was transfused with either whole blood or fresh frozen plasma. Following the introduction of cryoprecipitate, commercially prepared factor VIII concentrates

were introduced in the early 1970s. These concentrates were processed from large pools of plasma, which were obtained from thousands of donors. As a complication of replacement therapy, many hemophiliacs were found to have chronic persistent or chronic active hepatitis. In some cases frank cirrhosis was also diagnosed. With the introduction of specific testing for hepatitis B surface antigen and subsequently surrogate and specific testing for non-A non-B hepatitis virus (hepatitis C) the incidence of post-transfusion hepatitis decreased. Unfortunately, coincident with the advances of hepatitis screening, transfusion associated human immunodeficiency virus (HIV) was recognized in the early 1980s. This was a particularly devastating finding to all hemophilia A patients who had found greater social and employment opportunities with the use of self-administered factor VIII concentrates.

TABLE 48.1. CLINICAL CLASSIFICATION OF HEMOPHILIA A

Severity	Factor VIII Activity	Clinical Picture	Activated Partial Thromboplastin Time	Incidence
Severe	<1%	Severe hemarthrosis and spontaneous bleeding	Prolonged	48%
Moderate	2% to 15%	Spontaneous bleeding uncommon; serious bleeding from minimal trauma	Prolonged	31%
Mild	>15% to <30%	Spontaneous bleeding rare; unsuspected bleeding after surgical intervention or trauma	Variable	21%
Subclinical	>30% to <50%	Often asymptomatic, will bleed with trauma or surgery	Prolonged/Normal	

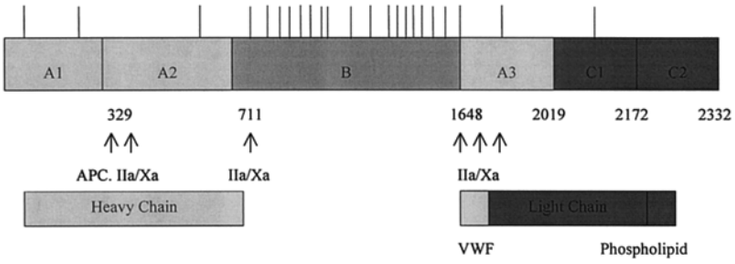


FIGURE 48.1. Diagram of factor VIII molecule.

TABLE 48.2. LABORATORY RESULTS IN HEMOPHILIA A

PT	Normal
APTT	Abnormal*
Bleeding Time	Normal
Platelet Count	Normal
Factor VIII:C Assay	Abnormal

APTT, activated partial thromboplastin time.
PT, prothrombin time
The APTT may be borderline or within the reference interval in patients with mild or moderate hemophilia A. This variability is from both choice of APTT reagent and biological patient variables. For instance, during times of stress, the factor VIII levels may increase, resulting in a normal APTT.

The late 1980s and early 1990s were characterized by greater emphasis on increased purity of factor VIII concentrates with enhanced viral attenuation or inactivation. A variety of viral attenuation techniques were used including: dry heat pasteurization, solvent/detergent treatment, steam treatment and heating in an organic solvent. Also attempts to increase factor VIII purity were introduced using monoclonal murine antibodies to either human VIII or human von Willebrand factor (vWF). The monoclonal antibodies were used with immunoaffinity chromatography.

In 1984, successful cloning and expression of factor VIII was reported. Recombinant factor VIII products are prepared by transfecting cDNA encoding human factor VIII into either baby hamster kidney or Chinese hamster ovary cell lines (4). Also, recombinant factor VIII that lacks the B domain has been manufactured. The availability of recombinant factor VIII eliminates problems of virally transmitted diseases associated with earlier factor VIII concentrates. Recently, transgenic animals (cows and goats) have been used to produce human proteins, which can be harvested from their milk. Transgenic pigs that produce human factor VIII also have been reported.

Recombinant factor VIII concentrates should be used to treat an acute bleeding episode or to prepare the patient for surgery. One unit of factor VIII/Kg of body weight will raise the factor VIII level approximately 2% (0.02 u/mL). The half-life of infused factor VIII is approximately 12 hours. Typically, a bleeding episode is treated with 40 to 50 u/Kg with subsequent infusion of 3 to 4 u/Kg/hr. In order to monitor response, factor VIII:C assays are recommended on a daily basis. Prophylactic use of factor VIII concentrates may be indicated following a CNS hemorrhage.

In mild/moderate hemophilia A patients, the initial treatment of choice is DDAVP (1-deamino-8-D arginine vasopressin) (Stimate, Ferning Pharmaceuticals, Malmo, Sweden). This synthetic analogue of vasopressin will cause a two- to 10-fold increase in factor VIII:C following an IV infusion of 0.3 to 0.4 ug/Kg of body weight. The postulated mechanism of action is release of vWF from endothelial storage sites (Weibel Palade Bodies). Repeated doses may lead to diminished response of factor VIII. Side effects of DDAVP include: facial flushing, fluid retention resulting from the antidiuretic effect and recent reports of thrombosis. The latter complication has been most often

identified in older individuals with significant atherosclerotic vascular disease.

One of the most common complications of hemophilia A is the appearance of an alloantibody to factor VIII:C (factor VIII:C inhibitors) (5). Factor VIII:C inhibitors are most commonly seen in severe hemophilia A patients. In the majority of cases they appear prior to the age of 20 years. As part of the annual evaluation of hemophilia A patients it is necessary to screen each patient for factor VIII: C inhibitors. These inhibitors are quantitated utilizing the Bethesda Assay system. Because most factor VIII:C inhibitors are time and temperature dependent, the Bethesda assay involves mixing patient plasma and a source of normal plasma. The mixture is incubated at 37°C for 2 hours. APTTs are performed at 30, 60, and 120 minutes. By definition one Bethesda Unit (BU) is defined as the amount of inhibitor that destroys half the factor VIII:C activity in an equal mixture of patient and normal plasma utilizing the above conditions. It is imperative to quantitate a factor VIII:C inhibitor so proper therapy can be instituted. Patients with high titer inhibitors (e.g., >30 BU) frequently are very difficult to manage. In addition to the Bethesda titer, it is also necessary to obtain a clinical history regarding the patient's anamnestic response. Arbitrarily, the patients have been divided into “low” and “high” responders. “High” responders will demonstrate a marked increase in factor VIII:C inhibitor titer upon exposure to factor VIII concentrates.

A variety of modalities have been used to treat factor VIII:C inhibitors (Table 48.3). Porcine factor VIII is of value particularly in patients whose factor VIII inhibitor has little cross reactivity with porcine VIII. Both “activated” and “nonactivated” prothrombin complex concentrates (PCC) (II, VII, IX, X) have been successfully used to treat acute hemorrhage. Protein A columns have been used to remove IgG from factor VIII:C inhibitor patients. In most cases factor VIII:C inhibitors are IgG (4). Consequently, the protein A will remove the majority of the intravascular inhibitor.

TABLE 48.3. THERAPEUTIC OPTIONS FOR MANAGEMENT OF FACTOR VIII:C INHIBITORS

High-dose factor VIII concentrates
Plasmapheresis
Prothrombin complex concentrates
Activated prothrombin complex concentrates
Immunosuppression
Heterologous factor VIII (Porcine)
Factor VIIa (recombinant)
Platelet transfusions
DDAVP
IV immunoglobulin

Recombinant factor VII (NovoSeven, Novo Nordisk, Copenhagen, Denmark) has been used effectively to treat patients with factor VIII inhibitors (6). Recombinant factor VIIa results in increased thrombin generation and more effective hemostasis. Recombinant factor VIIa has also been used in other clinical states including severe liver disease.

Also, experimental efforts are ongoing to develop gene therapy to “cure” hemophilia. Unfortunately, there have been many technical barriers that have delayed introduction of gene therapy. Selecting the appropriate vector to carry normal genes for either factor VIII or factor IX has been an issue particularly with respect to factor VIII. The factor VIII gene is very large; consequently, viral vectors with a large DNA component are required. Currently, experimental studies using the gene for factor IX have shown promise in animal models (dogs and mice).

Hemophilia B (Factor IX Deficiency)

Hemophilia B is also a sex-linked disorder. Christmas disease and factor IX deficiency are frequently used synonyms. The incidence of hemophilia B is approximately 1:40,000 to 1:50,000 population. However, in certain groups such as the Amish and East Indians, hemophilia B occurs as often as hemophilia A. Varying degrees of hemophilia B have been recognized with severe <1% (0.01 u/mL), moderate <5% (0.01 to 0.05 u/mL) and mild >5% (>0.05 u/mL) phenotypes analogous to factor VIII deficiency. The clinical presentation and genetics of hemophilia B are similar to hemophilia A. Consequently, to establish the diagnosis a specific factor IX assay is necessary.

With severe and moderate factors IX deficiency the APTT will be prolonged. However, there is marked variability among the commercial APTT reagents with respect: to factor IX sensitivity. Consequently, mild and moderate hemophilia B patients may have a normal screening APTT. In this situation, the family and personal history are extremely important in identifying a bleeding disorder. Typically, the PT is normal, however, there is a variant (hemophilia B_m) in which the PT is prolonged if bovine brain thromboplastin is used (Table 48.4). Tests of primary hemostasis (BT and platelet count) typically are normal.

TABLE 48.4. LABORATORY RESULTS IN HEMOPHILIA B

PT	Normal*
APTT	Abnormal*
Bleeding Time	Normal
Platelet Count	Normal
Factor IX Assay	Abnormal

APTT, activated partial thromboplastin time
PT, prothrombin time
In some instances, the prothrombin time may be prolonged if a bovine brain thromboplastin is used. This has been referred to as the Hemophilia B_m variant. The APTT is variable and may be normal in mild hemophilia B patients. This is due to the varying sensitivities of commercial APTT reagents.

Factor IX is a vitamin-K–dependent protein that is produced in the liver. The existence of a second coagulation protein, a deficiency of which was associated with hemophilia A, was first identified by Pavlovsky. He demonstrated mutual correction of

prolonged clotting times when plasmas from two hemophiliacs were mixed *in vitro*. Thus, conclusively demonstrating different coagulation factor deficiencies. The subsequent study of the various phenotypes of factor IX deficiency was instrumental in emphasizing the marked laboratory heterogeneity in factor IX deficient patients despite apparent clinical homogeneity. Utilizing antibodies to factor IX, early investigators found occasional hemophilia B patients whose plasma neutralized these antibodies. This established the existence of cross-reacting material (CRM) in patient plasmas with no apparent functional activity. Based on reactivity of patient plasma with heterologous antisera to factor IX, three phenotypic groups have been identified (Table 48.5). The CRM^+ and CRM^R groups have been further characterized by their reactions with XIa, VIIa, ability to activate factor X, electrophoretic mobility, etc. Full discussion of these variants is beyond the scope of this text.

TABLE 48.5. PHENOTYPIC CLASSIFICATION OF HEMOPHILIA B

	Factor IX Assays		
Type	Activity	Antigen	PT with Bovine Thromboplastin
I CRM^+	Low	Normal	Normal
CRM^+ (Bm)	Low	Normal	Abnormal
CRM^+	Low	Normal	Variable
II CRM^R	Low	Intermediate	Normal
III CRM^-	Low	Low	Normal

Factor IX deficiency demonstrates a wide variety of phenotypes. Patients may have concordant decrease in both IX activity as measured by a coagulation assay and IX antigen, which may be quantitated utilizing various immunologic assays (e.g., ELISA, RIA, Laurell). The variants of factor IX deficiency have variable levels of IX antigen (ranging from normal to reduced), which exceed the activity levels. Some of the variant IX deficiencies also have an abnormal prothrombin time when a bovine thromboplast in is employed. (Modified from Giddings, J.C. *Molecular Genetics and Immunoanalysis in Blood Coagulation*. Ellis Horwood, Chichester, England, 1988, p: 55)
CRM, cross-reacting material

The spectrum of clinical symptoms in hemophilia B is similar to those of hemophilia A. With severe factor IX deficiency, patients have recurrent hemarthroses as well as intramuscular hemorrhages. Mild and moderate factor IX deficiency may not be detected until the patient is an adult. Complications are similar to hemophilia A. Recurrent intraarticular hemorrhages may lead to damage of the cartilaginous surfaces of the affected joints and, ultimately, destruction of the joint. Muscle atrophy, nerve injury, contractures, and pseudotumors also may occur.

Alloantibodies to factor IX has been reported in approximately 10% of patients with hemophilia B. In most instances, these inhibitors arise in patients with severe hemophilia B. Also the infectious complications associated with chronic administration of blood products are similar to those of hemophilia A. Treatment of the severe hemophilia B patient in the recent past required the frequent use of PCC. These concentrates are made from pools of plasma and initially were prepared by adsorption of plasma by tricalcium phosphate. Subsequently, alcohol fractionation of plasma and DEAE-cellulose adsorption were utilized to prepare more potent concentrates. Heat treatment of PCC has significantly improved the safety of these products with respect to post-transfusion hepatitis and HIV complications. Recombinant human factor IX is now available and is the preferred treatment of choice. Recipients of recombinant factor IX are not exposed to the potential viral infections previously associated with PCC. Also, the risk of thrombosis that was reported with PCCs is not encountered with recombinant factor IX. On occasion, patients with severe factor IX deficiency may develop anaphylaxis following factor IX replacement. Although this is a rare phenomenon, physicians responsible for the care of such patients need to be vigilant. In most of the cases described, underlying genetic abnormalities involve significant deletions or other mutations of the factor IX gene. These patients develop antibodies to factor IX and, thus, may have an anaphylactic response when given replacement therapy.

Once transfused, factor IX has an *in vivo* half-life of approximately 18 to 40 hours. However, the recovery of factor IX following transfusion is only 20% to 40%. This is because of the ready distribution of factor IX to the extravascular compartment. Consequently, it is extremely difficult to achieve factor IX levels of greater than 50% following transfusion. In most instances of severe hemorrhage, treatment should be aimed at achieving a factor IX level of 20% to 50%. This is particularly important during the acute episode and the first few days thereafter. Minor hemorrhages usually respond to factor IX levels of approximately 20%.

HEMOPHILIA C (FACTOR XI DEFICIENCY)

Part of "48 - Coagulation Abnormalities"

Factor XI deficiency is an incompletely autosomal-recessive hereditary disorder. The majority of factor XI deficiency occurs in individuals of Jewish descent. In an extensive study by Seligsohn of Ashkenazi Jews, the frequency of homozygotes for factor XI deficiency was 0.1% to 0.3% whereas the heterozygotes were found in 5.5% to 11.0% (7). Factor XI is a glycoprotein with a molecular weight of 160,000 daltons. Its concentration in plasma is approximately 5 ug/mL. Native factor XI exists as a dimer of two identical subunits linked by disulfide bonds. Factor XI is activated by surface bound factor XIIa (Hageman factor). High-molecular-weight kininogen (Fitzgerald factor) serves as a cofactor in the activation step of factor XI. The gene for human factor XI has been isolated from a human liver cDNA library. The gene is 25 kb in length and consists of 15 exons and 14 introns. The gene has been located on the distal end of the long arm of chromosome 4. Factor XI is 607 amino acids in length and has significant homology with prekallikrein (Fletcher factor). The major physiologic substrate for factor XIa is factor IX, which is cleaved at two sites to yield factor IXa. In contrast to the marked heterogeneity of factor IX deficiency, factor XI deficient patients typically have levels of factor XI antigen, which parallel the factor XI coagulant activity.

Two major types of factor XI deficiency have been described. The homozygous patients have factor XI activity below 20% (as related to normal plasma). These patients have a variable history of clinical bleeding and are at risk for significant surgical bleeding. Heterozygous patients have levels of factor XI activity of 30% to 65% and typically have a negative bleeding history. Perhaps the most perplexing aspect of diagnosing factor XI deficiency is the lack of correlation between factor XI assay results and clinical bleeding.

The spectrum of clinical symptoms includes mild mucocutaneous bleeding (epistaxis, menorrhagia) and bleeding following minor surgical procedures such as dental extraction and tonsillectomy. A particularly difficult aspect of management of factor XI deficiency is the variability within an individual patient in response to different surgical procedures. Consequently, it is not possible to accurately predict how an individual patient may respond to a subsequent surgery. Bleeding from organs with excess plasminogen activator activity (t·PA) has been observed frequently (e.g., post-tonsillectomy, dental extractions, and transurethral prostate resection). In these instances, the use of Amicar often is helpful.

The management of hemorrhage following surgery or significant trauma involves the use of fresh frozen plasma. The half-life of infused factor XI is approximately 40 to 80 hours. It is recommended the factor XI level be maintained above 50%. Arbitrarily, 1 mL of fresh frozen plasma will produce a 2% rise in factor XI activity. A factor XI concentrate is available in Europe (8).

FACTOR VII DEFICIENCY

Part of "48 - Coagulation Abnormalities"

Deficiency of factor VII is inherited in an autosomal-recessive manner. Severe factor VII deficient patients (homozygotes) may present with a spectrum of clinical bleeding ranging from central nervous system hemorrhage in the neonatal state, hemarthroses, to mild mucocutaneous bleeding. The heterozygous patients typically are asymptomatic.

Factor VII is a single-chain glycoprotein with an estimated molecular weight of 50,000 d. The mature protein contains 406 amino acids and there are 10 gamma carboxyglutamic acid residues (GLA) located on the amino-terminal end. Beta hydroxyaspartic acid also is found in the factor VII molecule. The activation of factor VII requires cleavage of an arginine isoleucine peptide bond. Factor VII can be activated by several enzymes including: thrombin, factor Xa, factor lXa, and factor XIIa.

The gene for factor VII is located on chromosome 13 (Q 34-qter). The complete sequence for the gene spans approximately 12.8 kb of DNA. There are eight exons and seven introns. The gene for factor X also is located on chromosome 13, very close (2.8 Kb) to the factor VII gene.

As noted, the spectrum of clinical bleeding in factor VII deficiency is probably the most diverse of any of the hereditary coagulopathies. Paradoxically, some patients with factor VII deficiency have been reported to have thromboembolic episodes. Factor VII deficiency also has been described in a variety of hereditary disorders including Dubin-Johnson's, Rotor's, and Gilbert's syndromes. Combined factor VII and factor X deficiency has been described with a hereditary form of carotid body tumors.

The diagnosis of congenital factor VII deficiency is established by specific factor VII assays. Typically, the screening coagulation test results include a normal APTT and thrombin time together with a prolonged PT. However, PT results may vary depending on the choice of thromboplastin. Thus, when indicated, it may be appropriate to utilize a panel of thromboplastins to evaluate patients in whom there is a possible factor VII deficiency (human, bovine, and rabbit tissue). Many variant molecules have been identified based on their reactivity with thromboplastins of various species. The best predictor of clinical bleeding is the level of factor VII activity obtained with a thromboplastin of human origin.

The half-life of factor VII is approximately 6 hours. As a result, it is very difficult to obtain levels of factor VII activity of greater than 50%. Transfusion therapy includes the use of fresh-frozen plasma or PCC. If PCCs are used, heat-treated preparations are indicated. Although not available in the United States, a factor VII concentrate is available in Europe. Interestingly, recombinant factor VIIa has been utilized to treated hemophilia A patients with inhibitors. Factor VIIa leads to direct activation of factor X, thus bypassing the requirement for factor VIII in the coagulation sequence.

HEREDITARY FACTOR X DEFICIENCY (STUART-PROWER FACTOR)

Part of "48 - Coagulation Abnormalities"

Hereditary factor X deficiency is an autosomal-recessive disorder with variable degrees of penetrance. It was originally independently described by two groups who noted there was apparent heterogeneity in patients who were thought to be factor VII deficient on the basis of a prolonged PT. Factor X deficient patients were first clearly separated from factor VII deficiency on the basis of an abnormal thromboplastin generation time. Subsequently, it was appreciated that the Russell Viper Venom Time (RVVT) and APTT also were abnormal in factor X deficiency (Table 48.6).

TABLE 48.6. LABORATORY RESULTS IN FACTOR X DEFICIENCY

Case	PT	APTT	RVVT	Factor X Antigen	X Assay
Stuart	Abnormal	Abnormal	Abnormal	Low	Low
Prower	Abnormal	Abnormal	Abnormal	Normal	Low
Friuli	Abnormal	Abnormal	Normal	Normal	Low

The table summarizes the more common forms of factor X deficiency. There is a relatively high incidence of CRM+ variant deficiencies among all of the vitamin K dependent proteins. Consequently, careful and thorough evaluation is suggested for every case.

APTT, activated partial thromboplastin time

CRM,

PT, platelet transfusions

Factor X is a glycoprotein with a molecular weight of approximately 59,000 d. It is composed of two chains (light-chain molecular weight 16,900 and heavy-chain molecular weight 42,100). Factor X is activated by either the complex of factor VIIIa, factor IXa, phospholipid, and calcium (intrinsic pathway) or tissue factor/VIIa (extrinsic pathway). Activation requires cleavage of a peptide bond in the amino-terminal end of the heavy chain.

The gene for human factor X is located on chromosome 13 in the region Q 34-qter.321. It contains an estimated 25 kb of DNA. There are seven introns and eight exons. Location of the introns is analogous to the introns in factor VII, factor IX, and protein C genes.

The spectrum of hereditary factor X deficiency is analogous to that of other hereditary deficiencies of vitamin K dependent proteins. Patients have been reported in whom there are discrepancies between the RVVT and PT results (Friuli defect). In addition, discrepancies between the level of factor X antigen as measured with either radioimmunoassays or antibody neutralization techniques and factor X activity based on clotting assays have been observed. Thus, there are CRM^{+}, CRM^{R}, and CRM^{-} variants of hereditary X deficiency (Table 48.6). Hereditary deficiency of factor X may be caused by mutations throughout the gene resulting in either reduced or absent synthesis and secretion of factor X or the synthesis of an abnormal molecule.

Clinically, patients may present with severe bleeding; however, the most common complaint is easy bruising and hematoma formation. Hemarthroses, exsanguinating post-operative hemorrhage and central nervous system hemorrhage have been reported in the most severely affected patients. Systemic amyloidosis is an important differential diagnostic consideration when isolated factor X deficiency is encountered. It appears amyloid deposits in these particular cases have an increased affinity for factor X resulting in adsorption of factor X from circulating plasma.

The half-life of transfused factor X is approximately 24 to 40 hours with a bimodal survival curve. Both fresh-frozen plasma and PCC are effective. If concentrates are used, heat treated products are indicated.

FACTOR V DEFICIENCY (LABILE FACTOR)

Part of "48 - Coagulation Abnormalities"

Hereditary factor V deficiency is inherited as an autosomal-recessive trait. Both CRM+ and CRM^{-} variants have been described. In addition, a disorder referred to as factor V Quebec has been identified. In this disorder, there is autosomal-dominant inheritance with severe clinical bleeding. However, plasma factor V activity levels are borderline normal with corresponding factor V antigen levels. Paradoxically, platelet factor V activity levels are very low. Thus, somewhat analogous to von Willebrand disease and afibrinogenemia, it may be necessary to evaluate both plasma and platelet factor V in order to accurately diagnose and characterize factor V deficiency. Several investigators have emphasized the correlation of platelet factor V with clinical bleeding as opposed to a lack of correlation with plasma levels.

The clinical picture in patients with hereditary deficiency of factor V is characterized by a pattern of mucocutaneous bleeding. Hemarthroses are relatively uncommon. The differential diagnosis includes combined deficiency of factor V and VIII and acquired factor V deficiency. Acquired deficiency of factor V usually is a result of specific antibodies. Inhibitors to factor V are most commonly seen in the post-operative setting or in association with administration of certain antibiotics. Acquired factor V antibodies in the postoperative setting are most frequently associated with exposure to topical bovine thrombin used in conjunction with a source of fibrinogen to produce “fibrin-glue.” In most cases, the thrombin preparations are of animal origin (bovine); consequently, particularly in situations where a patient is having a second surgery, reexposure to topical thrombin may result in clinical bleeding due to decreased levels of factor V. Patients at highest risk are those who have cardiovascular surgery, neurosurgery, orthopedic surgery, or plastic surgery. On occasion, patients may have significant postoperative bleeding.

Laboratory diagnosis of factor V deficiency is characterized by an abnormal PT and APTT. The diagnosis is confirmed with a factor V assay. Approximately one third of patients will have a prolonged bleeding time. As noted previously, factor V is carried within the platelet alpha granules and also may be associated with the platelet membrane. Activated factor V on the surface of the platelet serves as a receptor for factor Xa, thus facilitating the assembly of the prothrombinase complex.

Factor V deficiency may be treated with fresh-frozen plasma or fresh plasma. The half-life of transfused factor V is approximately 12 hours. Platelet concentrates have been used successfully to treat patients with inhibitors to factor V.

COMBINED FACTOR V AND FACTOR VIII DEFICIENCY

Part of "48 - Coagulation Abnormalities"

A combined deficiency of factor V and factor VIII has been reported. In most cases this appears to be an autosomal-recessive disorder. Thus, there is a history of consanguinity in many of these cases. The plasma levels of factors V and VIII usually are in the range of 15% to 20%. The genetic basis for this disorder is currently unknown.

PROTHROMBIN DEFICIENCY

Part of "48 - Coagulation Abnormalities"

Hereditary hypoprothrombinemia is an extremely rare autosomal-recessive disorder. Homozygous patients have levels of prothrombin activity varying from 2% to 25%. Prothrombin is a single-chain glycoprotein with a molecular weight of approximately 71,600. In addition to the gamma carboxyglutamic acid residues, which characterize all vitamin-K–dependent proteins, there are two kringle structures present. These kringles are analogous to similar structures seen in plasminogen, t·PA, urokinase, and factor XII.

The gene for human prothrombin is located on chromosome 11. It contains approximately 21 kb of DNA and has 14 exons with 13 introns.

A number of variants of abnormal prothrombin have been described. Defects in factor Xa cleavage site have been identified as well as abnormal calcium binding sites and abnormalities of the thrombin region.

The clinical findings are variable depending on the level of prothrombin activity. Spontaneous hemorrhage and hemarthroses are uncommon although they have been described in patients with severe deficiency states.

Laboratory findings are characterized by the presence of an abnormal PT and APTT. The PT results may be highly variable depending on the nature of the underlying defect. In order to establish the diagnosis, a two-stage functional assay of prothrombin is recommended as well as an immunologic assay to detect the presence of prothrombin antigen. In cases of true hypoprothrombinemia, the functional and antigenic assays will be decreased in parallel. In the dysprothrombinemic states, the antigen level will exceed the functional activity.

Treatment consists of the use of PCC.

AFIBRINOGENEMIA AND HYPOFIBRINOGENEMIA

Part of "48 - Coagulation Abnormalities"

Afibrinogenemia is a rare disorder inherited in an autosomal-recessive pattern. There often is a variable history of clinical bleeding. In many cases, there is bleeding at the time of separation of the umbilical cord or following circumcision. Also, patients may present with intracerebral hemorrhage, spontaneous hematoma formation, or mucocutaneous bleeding. Paradoxically, menorrhagia is uncommon with afibrinogenemia.

Laboratory abnormalities of all of the standard screening tests are present. No end point is detected with the PT, APTT, or thrombin time. The bleeding time may be slightly prolonged and platelet aggregation studies typically show a lack of response to the usual agonists (ADP, epinephrine, and collagen). Although fibrinogen may not be demonstrable in the patient plasma, most patients will have variable amounts of fibrinogen within the platelet alpha granules. Discrepancies between platelet and plasma concentrations of coagulation proteins also have been reported with factor V deficiency and von Willebrand's disease.

Replacement therapy is required to achieve a hemostatic level of fibrinogen between 50 and 100 mg/dL. Cryoprecipitate or fresh-frozen plasma is effective. A few cases of antibody formation to fibrinogen have occurred following replacement therapy.

Hypofibrinogenemic patients are characterized by a fibrinogen level, which is less than 100 mg/dL. Often these patients have a mild bleeding tendency with an autosomal dominant or, in some cases, an autosomal-recessive pattern of inheritance. In many cases, hypofibrinogenemic patients will have normal PT and APTT results. However, typically the thrombin time will be abnormal. Clottable and immunologic fibrinogen assay results are decreased thus establishing the diagnosis.

DYSFIBRINOGENEMIA

Part of "48 - Coagulation Abnormalities"

Hereditary dysfibrinogenemia is inherited in an autosomal-dominant pattern. These abnormal fibrinogen molecules are caused by mutations, which result in single amino-acid alteration in one of the three fibrinogen chains (a-alpha, b-beta, or gamma chains). Many cases of dysfibrinogenemia have been fully characterized. The mutations may affect any one of the functional properties of fibrinogen including delayed or absent release of fibrinopeptides A and B, abnormal polymerization, abnormal cross-linking, decreased thrombin binding, abnormal plasmin digestion, and defective secretion by hepatocytes. Clinically, patients may present with hemorrhage, spontaneous abortion, or thromboembolic complications. However, the vast majority of patients are asymptomatic. Laboratory findings include normal or only minimally prolonged values for PT and APTT. Typically, the thrombin time is prolonged as is the reptilase time. The abnormal thrombin time usually cannot be corrected by the addition of protamine or calcium. Laboratory diagnosis relies on the demonstration of a discrepancy between the level of clottable fibrinogen and antigenic fibrinogen. Specialized studies can be performed including immunoelectrophoresis and SDS polyacrylamide gel electrophoresis to evaluate molecular weights, polypeptide chains, formation of cross-linked gamma-chain dimers and alpha-chain polymers. In addition, polymerization of fibrin monomers may be evaluated by measuring turbidity changes at 305 nm.

Management of patients with dysfibrinogenemia is dependent upon the associated clinical findings. In the case of bleeding, the use of cryoprecipitate is the treatment of choice.

FACTOR XIII DEFICIENCY (FIBRIN STABILIZING FACTOR)

Part of "48 - Coagulation Abnormalities"

Hereditary deficiency of factor XIII is inherited in an autosomal-recessive pattern. Factor XIII is a plasma glycoprotein that is involved in stabilizing the fibrin clot by introducing covalent bonds between fibrin multimers. In addition, factor XIII also cross-links alpha-2 antiplasmin to the fibrin clot and fibronectin to fibrin or collagen.

Factor XIII circulates in plasma as a tetramer composed of two a-subunits and two b-subunits. The a-subunits are also found in other tissues such as platelets and megakaryocytes, placenta, uterus, and macrophages. The a-subunit is responsible for the functional activity of factor XIII.

The gene for the a-subunit is located on the distal end of the short arm of chromosome 6. While the gene for subunit-b is located on the long arm of chromosome I (1 q3l). In affected individuals, the first evidence of clinical bleeding is often manifested at the time of umbilical cord separation (9). Patients also may have intracranial hemorrhage as well as intramuscular hematoma formation and easy bruising. Many patients have abnormal wound healing and a history of habitual abortion.

The usual screening tests are normal. When factor XIII deficiency is suspected, clot solubility in five molar urea or monochloroacetic acid will be abnormal. In addition, a thromboelastogram also will have a markedly abnormal tracing. Subunit-a can be measured using the Laurell technique. In a few selected laboratories, factor XIII activity is quantitated using amine incorporation assays.

Cryoprecipitate and fresh-frozen plasma have been used to treat hereditary factor XIII deficiency. In addition, there is a concentrate of factor XIII available in Europe. Because factor XIII has a rather long half life and the majority of patients do not become

symptomatic until levels of activity are less than 1%, replacement therapy on an intermittent basis is remarkably effective.

HEREDITARY COAGULATION FACTOR DEFICIENCIES NOT ASSOCIATED WITH CLINICAL BLEEDING

Part of "48 - Coagulation Abnormalities"

Hageman factor (factor XII), Fitzgerald factor (high-molecular-weight kininogen), and Fletcher factor (prekallikrein) are proteins that are involved in the contact phase of coagulation (Table 48.7). Deficiencies of these proteins are not associated with clinical bleeding. Typically these patients are detected when a screening APTT is found to be markedly prolonged in the presence of a negative clinical history of bleeding. The diagnosis is established by specific factor assays. In the case of Fletcher factor deficiency, prolonged incubation (10 minutes) of the patient plasma with an activator (e.g., Kaolin, micronized silica) may appreciably shorten the prolonged APTT obtained with routine activation (3 to 5 minutes). Similar shortening is not seen with Hageman factor or Fitzgerald deficiencies.

TABLE 48.7. HEREDITARY DEFICIENCIES OF CONTACT FACTORS

	Factor XI	Factor XII	Prekallikrein	HMW-Kininogen
Year described	1953	1955	1965	1975
Mode of transmission	Autosomal Recessive	Autosomal Recessive	Autosomal Recessive	Autosomal Recessive
Ethnic background	Predominantly in Jews	Variable	Variable	Variable
Bleeding symptoms	Mild	None	None	None
Prothrombin time	Normal	Normal	Normal	Normal
APTT	Prolonged	Prolonged	Prolonged	Prolonged
Intrinsic fibrinolysis	Normal	Impaired	Impaired	Impaired
Kinin formation	Normal	Impaired	Impaired	Impaired
CRM^+ cases	Present	Present	Present	Unknown

APTT, activated partial thromboplastin time
CRM, cross reacting material

Paradoxically, a number of patients with Hageman factor deficiency and prekallikrein deficiency have been reported with thromboembolic events.

ACQUIRED DISORDERS

Part of "48 - Coagulation Abnormalities"

Anticoagulant Therapy (Heparin, Coumadin)

The role of the clinical laboratory in monitoring anticoagulant therapy is a very important contribution to patient care. In fact, monitoring anticoagulant therapy with the PT was the first form of therapeutic drug monitoring.

Heparin is a natural mammalian glycosaminoglycan. The manufacture of heparin involves extraction from bovine lung tissue and porcine or bovine intestinal mucosa. Two different preparations of heparin are available in the United States. The first of these is unfractionated heparin, which is a biological product obtained from either bovine lung or porcine mucosa. Unfractionated heparin has a distribution of molecular weights varying from 5,000 to 30,000+ d. Unfractionated heparin interacts with antithrombin resulting in inhibition of the two most important serine proteases (factor Xa and thrombin). Heparin is used in many areas of the hospital including cardiopulmonary bypass surgery, hemodialysis, keep open venous access, and to anticoagulate patients with a thromboembolic event. Recently, low-molecular-weight heparin (LMWH) has gained considerable momentum in replacing unfractionated heparin. LMWH preparations are manufactured by enzymatic or chemical degradation of unfractionated heparin. In general, LMWHs have significant differences in pharmacologic properties as well as their effect on various coagulation assays (Table 48.8).

Heparin requires a naturally occurring plasma constituent in order to express its anticoagulant activities. This plasma protein is antithrombin (formerly called antithrombin III), an alpha-2 globulin produced in the liver. Antithrombin (AT) is a member of the serpin family of proteins. Thus, the name antithrombin is something of a misnomer. AT inhibits all of the serine proteases involved in hemostasis. The inhibitory properties of AT are time dependent in the absence of heparin. With the addition of heparin, the ability of AT heparin complex to inhibit serine proteases is markedly accelerated.

The half-life of injected heparin is in the range of 1 to 2 hours. However, the disappearance and clearance times of heparin are dose related. Higher doses are associated with decreased clearance and longer disappearance times. Heparin may be cleared from the circulation both by the kidney as well as the liver. Endothelial cells also play a role in heparin clearance. Heparin may be neutralized in the circulation by platelet factor IV, which is a platelet specific protein found in the alpha granule. Heparin also interacts with other plasma proteins including vWF, fibronectin, and vitronectin.

Although historically heparin was given subcutaneously as well as intravenously (either intermittent or continuous infusion), most patients today receive heparin by continuous infusion. Continuous infusion appears to be the preferred mode of administration, not only because of the ease of obtaining samples for monitoring the heparin effect, but there also appears to be less clinical bleeding. The decrease in clinical bleeding is primarily related to a decrease in the cumulative dose a patient receives over the course of treatment. In the treatment of acute deep-vein thrombosis (DVT), usually a loading dose of 5,000 to 10,000

units of heparin is administered immediately followed by therapeutic doses of heparin based on body weight. Dosage may be affected by a variety of factors including the acute nature of the thrombotic process, body weight, and other complicating disease states (e.g., renal and liver disease).

TABLE 48.8. COMPARISON OF UNFRACTIONATED AND LMW HEPARINS

Property	Low-Molecular Weight Heparin	Unfractionated Heparin
Molecular weight	3,000-5,000	10,000-30,000
Activated partial thromboplastin time	++	++++
Thrombin time	++	++++
Anti-Xa assay	++++	++
Platelet aggregation	+	++++
Binding to endothelium	+	++
Protamine neutralization	+	++++
Lipoprotein lipase	+	++++
Adsorption (sub Q)	+++	++
Clinical bleeding (?)	++	+++

The laboratory monitoring of heparin has undergone an evolution beginning with the use of whole blood clotting times (Lee and White) to the recent introduction of synthetic substrate assays which allow quantitation of heparin in units per milliliter. Based on the College of American Pathology (CAP) Survey Program, the vast majority of laboratories (greater than 95%) utilize the APTT as a means of monitoring heparin. However, the activated clotting time (ACT) is utilized frequently in settings where immediate answers are required. This is primarily in the surgical suite for bypass procedures. Also, in dialysis units the ACT is a popular means of evaluating the presence of heparin and subsequent neutralization with protamine sulfate.

The APTT is readily available in virtually any laboratory. In addition, because it is part of an ongoing quality control program, it is the preferred test for monitoring heparin therapy (10). Commercially available, partial thromboplastins vary significantly in their sensitivity to heparin. When defining an APTT system, it is important for the laboratory to consider both the choice of reagent as well as instrumentation. In general, instruments and reagents from the same manufacturer perform better than situations in which there is manufacturer disparity. In addition to sensitivity to the presence of heparin, responsiveness is an important property of the APTT system. Responsiveness reflects the degree of prolongation present at a given level of heparin or factor deficiency. Thus, more responsive reagents will yield longer APTT results. It is important to separate the concept of sensitivity and responsiveness in choosing laboratory reagents. An optimal reagent will be both sensitive and responsive.

Historically, the therapeutic target for prolongation of the APTT was 1.5 to 2.5 times the upper limit of the normal reference interval. However, recent work has significantly changed the way laboratories and their anticoagulation committees set therapeutic ranges for heparin therapy. It is critical for the laboratory to define their test system based on an APTT reagent, coagulation instrumentation, collection tubes, centrifugation, etc. In order to develop a reference interval for therapeutic heparin, the best approach is to collect aliquots of plasma from patients receiving heparin and freeze these aliquots until approximately 30 to 40 samples are available. These samples then are thawed and evaluated with an APTT as well as a heparin assay. The heparin assay may be a chromogenic Xa inhibition assay or, alternatively, a protamine neutralization assay. The vast majority of laboratories utilize the factor Xa inhibition chromogenic assay. Once the APTT values and the factor Xa chromogenic inhibition results are available, one plots each point for an individual sample [APTT on vertical axis in seconds and Xa inhibition on the horizontal axis expressed in units of heparin (0.0 to 1.5 units)]. A line of best-fit regression analysis defines the slope. The therapeutic range for the Xa inhibition assay is 0.3 units to 0.7 units. It is a CAP requirement for laboratories to provide physicians with therapeutic ranges for heparin monitoring as well as INR values for monitoring oral anticoagulants.

Heparin-associated thrombocytopenia (HAT) is one of the most important complications of heparin therapy. HAT has been reported as varying from <1% of patients to as high as 30%. It has become evident that there are at least two types of HAT. Type I is seen frequently after the initiation of heparin therapy and consists of a drop in the platelet count into the range of 100,000 to 150,000/mL. This decrease in platelets may persist for several days; however, even though the patient continues to receive heparin, the platelet count subsequently returns to the original baseline value prior to the initiation of heparin therapy. The exact mechanism for this type of HAT is still speculative although many suggest it is the result of *in vivo* heparin-induced platelet aggregation with clearance of the aggregates in the liver and spleen.

Type II HAT is associated with antibodies directed against a complex of heparin and platelet factor 4 (H-PF4). The antibodies involved are usually IgG (predominantly IgG_4). The immune complex of IgG-H-PF4 then binds to FcγIIa receptors on the platelet membrane. This results in activation of the platelets with release of platelet microparticles as well as granular contents (α granules — dense bodies). The platelet microparticles provide a “prothrombotic environment” leading to generation of thrombin and further activation of platelets as well as endothelial cells. As a result of this process, there is consumption of platelets and, in some cases, profound thrombocytopenia (less than 100,000 per mL) and associated thromboembolic complications. Thrombi may occur in any of the vascular beds but principally involves arterial and venous sites.

Type II HAT usually occurs in patients who have received administration of heparin for greater than 6 days or, alternatively, patients who have recently been exposed to heparin and are re-exposed after a period of days or weeks of being off heparin. Type II HAT can be seen in situations where the patient is only minimally exposed to heparin (use of heparin to keep open an IV line).

When type II HAT is suspected, heparin should be discontinued immediately. Typically, in the patient receiving heparin for DVT, oral anticoagulants already have been initiated and the transition to this form of anticoagulant therapy is relatively easy. In the setting where HAT develops in the patient who requires bypass surgery or renal dialysis, management is considerably more difficult. Alternative forms of anticoagulation that appear to be safe include danaparoid and recombinant hirudin. In the

past, it was hoped LMWH preparations would not interact with antibodies to the complex of PF4-unfrationated heparin. Unfortunately, this was not the case. Patients with antibodies to PF4 UFH often will cross react with LMWH.

The laboratory confirmation of the presence of type II HAT is based on screening assays that are rather crude. Platelet aggregation studies should be performed in the presence of the heparin being utilized for the patient management as well as at least two sources of platelet rich plasma obtained from normal donors whose platelets are known to aggregate in the presence of antibodies to PF4-H. The aggregation studies consist of heparin (which the patient is receiving), normal platelet rich plasma and the patient platelet poor plasma mixed in the aggregation cuvette. The aggregation response typically is monitored over a 15-minute interval. The presence of aggregation suggests the possibility of HAT. It is important to appreciate the absence of a positive result does not rule out HAT. Other test systems that are more specific include the use of the serotonin platelet release assay. This assay is thought to be more sensitive and specific than the standard platelet aggregation assay described above. However, it is considerably more expensive and time-consuming because it requires the use of radiolabeled serotonin.

In addition to the functional assays for HAT, there are ELISA assays that utilize microtiter plates coated by the complex of heparin-PF4. The patient's serum or plasma is then added to the microtiter plate, appropriately incubated, washed and treated with a labeled antibody to IgG. This assay is very sensitive to the presence of antibodies to H-PF4. However, the specificity with respect to clinical complications is somewhat problematic.

ORAL ANTICOAGULANT THERAPY

Part of "48 - Coagulation Abnormalities"

Oral anticoagulant therapy is based on administration of coumarin or its derivatives. This class of drugs blocks the reductase enzyme in the vitamin K pathway (Fig. 48.2) resulting in increased levels of nonfunctional vitamin K epoxide. Vitamin K acts as a cofactor in a post-translational step converting various vitamin K-dependent protein precursors into functional procoagulant proteins. This vitamin K driven reaction involves the carboxylation of glutamic acid to give rise to gamma carboxy-glutamic acid (GLA). These GLA residues function to localize the vitamin K-dependent proteins to phospholipid membrane surfaces. Typically, activated platelets provide this surface; however, other cells such as endothelial cells, monocytes, and tumor cells may function in a similar fashion.

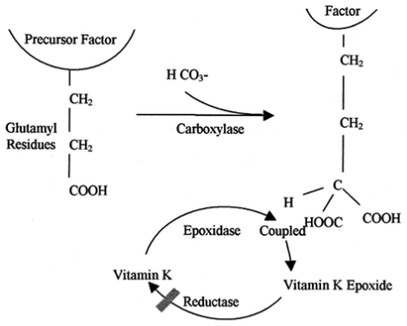


FIGURE 48.2. Vitamin K pathway.

In the presence of oral anticoagulants, formation of GLA residues is impeded. As a result, there is diminished surface localization of the vitamin K-dependent complexes resulting in an anticoagulant effect. Early reports also suggested the presence of PIVKAs (protein induced by vitamin K antagonist) serve an anticoagulant role as well. The PIVKA proteins represent the precursors of the functional procoagulants. Their role in the therapeutic effect of oral anticoagulants remains unresolved.

TABLE 48.9. CLINICAL CONDITIONS ASSOCIATED WITH LUPUS ANTICOAGULANTS

I. Autoimmune Diseases
- A. Systemic lupus erythematosus
- B. Rheumatoid arthritis
- C. Others, including “overlap” syndromes

II. Drug Exposure
- A. Chlorpromazine
- B. Procainamide
- C. Hydralazine
- D. Quinidine
- E. Antibiotics
- F. Phenytoin

III. Infections
- A. Bacterial
- B. Protozoan (Pneumocystis carinii)
- C. Viral

IV. Lymphoproliferative Disorders
- A. Hairy cell leukemia
- B. Malignant lymphoma
- C. Waldenstrom's macroglobulinemia

V. Miscellaneous Disorders
- A. Epithelial malignancies
- B. No underlying disease

The monitoring of oral anticoagulant therapy has relied upon the PT and variants of the PT. For many years there has been an international controversy regarding the optimal test system to monitor oral anticoagulants. For many years, virtually all of the thromboplastins marketed in the United States were of rabbit tissue origin (rabbit brain or rabbit brain/lung). However, in Europe a number of different types of thromboplastins have been utilized including bovine brain, monkey brain, and human brain. Dr. Leon Poller championed the human-brain thromboplastin as a sensitive reagent which more accurately reflected the anticoagulant status of a patient. His efforts resulted in the concept of an International Normalized Ratio (INR) to express PT results. With the introduction of recombinant thromboplastins, many laboratories have converted their PT reagents. The use of recombinant thromboplastin together with synthetic phospholipids results in a reagent with greater consistently between lots.

Furthermore, the ISI value for these reagents approaches 1.0. The vast majority of laboratories in the United States today utilize reagents with ISI values of 1.7 or lower.

In order to report an INR value, it is necessary to know the International Sensitivity Index (ISI) for the thromboplastin being used as well as the geometric mean of the PT range. The INR expresses the ratio of the patient PT to the mean of the normal range raised to the power of the ISI (Fig. 48.3). The INR, therefore, represents the PT, which would have been obtained on the patient plasma if the International Reference Preparation of thromboplastin (human brain) had been utilized. Thus, it provides a means of comparing PT results from laboratory to laboratory correcting for the differences in thromboplastin reagents. The concept of the INR has been well received in Europe and now is widely accepted in the United States. It is important to emphasize the INR is appropriate for patients who are *stably* anticoagulated. Patients who are in the early phases of oral anticoagulant therapy or patients in whom the PT is being obtained for other diagnostic purposes may have misleading INR results.

$$INR = \left[\frac{\text{Pat. PT}}{\text{Mean Normal PT}} \right]^{ISI}$$

FIGURE 48.3. Derivation of the international normalized ratio.

The introduction of the INR has emphasized the need for lower doses of oral anticoagulants. In the past when PT results were reported as ratios, the therapeutic range was typically quoted as PT ratios of 1.5 to 2.5. With the use of the INR, it is now evident the majority of patients are satisfactorily anticoagulated with INRs of 2.0 to 3.0. As a consequence, it is anticipated the incidence of bleeding in patients receiving oral anticoagulant therapy will significantly decrease.

FIBRINOLYTIC THERAPY

Part of "48 - Coagulation Abnormalities"

The use of fibrinolytic agents in the treatment of acute myocardial infarction (AMI) is now well accepted. Fibrinolytic agents such as streptokinase (SK) and urokinase (UK) have been available for many years. However, they were only approved for treatment of pulmonary embolism and massive DVT. Many clinicians were reluctant to use these agents because of their unfamiliarity with dose and the perceived high incidence of hemorrhagic complications. The introduction of tissue plasminogen activator (t·PA) renewed interest in fibrinolytic therapy and with the ready availability of acute coronary angiography led to the introduction of lytic agents in the treatment of AMI.

One of the primary complications associated with the use of fibrinolytic agents in the treatment of AMI is bleeding. Most commonly the bleeding is from arterial or venous access sites. The most catastrophic complication is cerebral bleeding.

The nature of the hemorrhagic defect in patients receiving lytic therapy has only recently been studied in detail. Among the alterations induced by lytic agents are a drop in the fibrinogen level together with a concomitant rise in fibrin/fibrinogen degradation products. It is well known that fibrin/fibrinogen degradation products will inhibit certain aspects of platelet function and, in addition, impede the assembly of fibrin multimers to form a fibrin clot. Also, plasmin will attack other plasma procoagulant proteins including factors V and VIII. Because plasmin is an indiscriminate serine protease, it will digest certain glycoprotein receptors on both platelets and endothelial cells (glycoprotein IIb/IIIa; glycoprotein Ib). As a result, the administration of lytic agents affects both primary and secondary hemostasis. Furthermore, many patients who are receiving lytic agents are also concomitantly receiving heparin and antiplatelet medications such as aspirin.

Therefore, the bleeding seen in association with fibrinolytic therapy is multifactorial. Appropriate management of these patients depends upon the individual case. The use of fresh-frozen plasma and/or platelets may be indicated in severe circumstances.

ACQUIRED CIRCULATING ANTICOAGULANTS

Part of "48 - Coagulation Abnormalities"

Circulating anticoagulants may be defined as endogenous substances that interfere with one or more of the *in vitro* tests of hemostasis. In most cases, they are immunoglobulins, however, there are situations in which other substances produce an anticoagulant effect including the presence of heparin or heparinlike compounds and fibrin/fibrinogen degradation products.

The most commonly encountered circulating anticoagulant is the so-called lupus anticoagulant (LA). A working definition of the LA is a circulating immunoglobulin (IgG or IgM or both) that interferes with one or more of the *in vitro* phospholipid dependent tests of coagulation (e.g., PT, APTT, dRVVT, etc.). The name LA is a misnomer because the vast majority of these patients *do not* have systemic lupus erythematosus. In fact, in most hospital settings, this anticoagulant is seen in a variety of other clinical circumstances including the convalescent phase of acute infections, as a result of drugs (chlorpromazine, hydralazine, hydantoin, procainamide, quinidine and various antibiotics), and in patients in whom there is no apparent underlying disease (Table 48.9). Although most inhibitors of blood coagulation are associated with clinical bleeding, paradoxically, LA is often associated with an increased risk of thrombosis. Clinical bleeding is distinctly uncommon unless there is an additional hemostatic abnormality. Most commonly, the hemostatic abnormalities which may be associated with clinical bleeding in LA patients include: thrombocytopenia, qualitative platelet functional defects, and an isolated deficiency of prothrombin. The latter appears to be directly related to the LA immunoglobulin, which binds to prothrombin creating an antigen antibody complex that is cleared rapidly from the circulation. The laboratory diagnosis of LA requires four distinct steps (11): (i) identification of an abnormal screening test (e.g., APTT, dRVVT); (ii) proof the abnormal test is because of an inhibitor (mixing studies); (iii) demonstrate the inhibitor is phospholipid dependent (Fig. 48.4); (iv) rule out other coagulopathies (e.g., factor VIII inhibitors, etc.).

In the vast majority of cases, the abnormal screening test is an unexplained prolonged APTT. In the past this was usually a serendipitous finding which then lead to further laboratory tests

to explain the abnormality. Recently clinicians are asking the laboratory: “Does my patient have an LA?” The latter scenario is distinctly different and requires a more extensive evaluation. The increased clinical awareness of LA is related to its association with a variety of complications including both arterial and venous thromboembolic events, neurologic complications, recurrent spontaneous abortions and intrauterine fetal death.

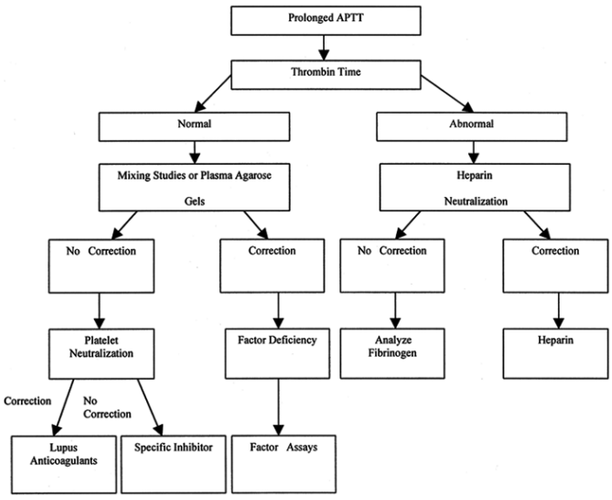


FIGURE 48.4. Workup of LA.

Following the demonstration of an abnormal APTT, mixing studies are indicated to identify the presence of an inhibitor. Although many texts recommend a one-part patient, one-part normal mixture, my experience suggests the use of four parts patient, one part normal plasma is a more sensitive means of identifying the presence of an inhibitor. This is particularly true in the cases where the patient's baseline APTT is only minimally prolonged. LA was initially thought to be an immediate inhibitor in contrast to the time dependency seen with factor VIII:C inhibitors.

TABLE 48.10. CLASSIFICATION OF CIRCULATING ANTICOAGULANTS

Neutralizing	Nonneutralizing
Specific	
Factor V	Prothrombin
Factor VIII:C	Factor VIII:C
Factor IX	Factor X
Factor XI	vWF
vWF	
Factor XIII	
Nonspecific	
Lupus anticoagulants	
Paraproteins	
FSP	
Global	
Heparinlike activity	

FSP, fibrin split product
vWF, von Willebrand factor

However, recently a number of studies have suggested as many as 30% of LAs are time dependent. Conversely, factor VIII:C inhibitors may be immediate. Thus, the presence or absence of time dependency is not a useful differential diagnostic point to separate factor VIII:C inhibitors from LA. An alternate method to identify the presence of a circulating inhibitor is the use of plasma agarose gels.

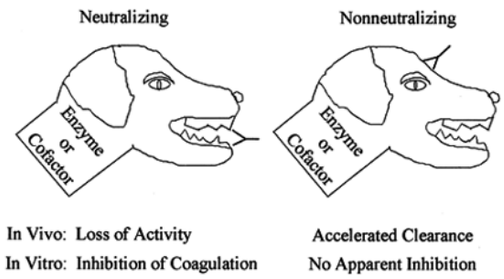


FIGURE 48.5. Neutralizing and nonneutralizing anticoagulants.

A variety of test systems have been utilized to identify phospholipid dependence of circulating anticoagulants. Two basic principles have been employed: (i) decreasing the phospholipid concentration in the test system to enhance the inhibitor effect, and (ii) increasing the amount or changing the configuration of phospholipid in the test system to either “neutralize” or “bypass” the LA.

Among the former tests, the tissue thromboplastin inhibition procedure (TTI) is most commonly employed. Unfortunately, this test system is not particularly sensitive or specific. Perhaps the most sensitive test system employing the concept of dilute phospholipid is the dilute Russell Viper Venom Time (dRVVT). The Platelet Neutralization Procedure (PNP) is the most commonly utilized test employing increased amounts of phospholipid. In this test, freeze-thawed platelets are added to the patient's platelet-poor plasma and the APTT repeated. In the presence of an LA, there will be significant shortening of the baseline APTT value (12). As a control, an appropriate buffer is added to the patient plasma.

Occasionally, the confirmatory tests may be misleading or uninterpretable. In this situation it is appropriate to perform factor assays. Typically in the presence of a LA there is nonparallelism of the factor assay curves with an apparent increase in factor activity as the patient plasma is progressively diluted.

In addition to the fibrin-based assays to demonstrate LA, many laboratories also utilize an ELISA system to demonstrate the presence of anticardiolipin antibodies (ACA). LA and ACA are often present in the same patient; however, in approximately 30% of cases, there will be a lack of concordance between these two test systems. Consequently, when the clinician requests an evaluation of the patient plasma for phospholipid antibodies, it is necessary to perform both fibrin based coagulation assays and the ELISA test system. Recent studies have indicated the antigenic targets for immunoglobulins responsible for LA activity include various plasma proteins, which bind to phospholipids. Among these, β_2 Glycoprotein I (apolipoprotein H) and prothrombin are the two most commonly recognized proteins (12). Consequently, specific ELISA tests for antibodies to β_2 glycoprotein I and prothrombin now are commercially available. It appears antibodies to β_2 Glycoprotein I have a higher frequency of thrombotic complications. Approximately 20% of patients with negative standard ACA results will have positive results utilizing a high sensitive microtiter plate coated with β_2 GPI.

Other inhibitors are rarely encountered (Table 48.10). The recent recognition of “nonneutralizing” inhibitors has further compounded the laboratories' difficulties in demonstrating the presence of an inhibitor (Fig. 48.5). With a nonneutralizing inhibitor the laboratory assays may suggest a factor deficiency since addition of normal plasma to the patient plasma results in correction of abnormal coagulation studies. In these patients it is necessary to demonstrate the presence of an antigen antibody complex utilizing a variety of techniques such as crossed immunoelectrophoresis. The most commonly encountered nonneutralizing inhibitor is the antiprothrombin antibodies encountered in patients with LA.

DISSEMINATED INTRAVASCULAR COAGULATION (DIC)

Part of "48 - Coagulation Abnormalities"

DIC may be encountered in a variety of different clinical situations (Table 48.11). In most cases it is associated with the exposure of tissue factor or thromboplastinlike material with resulting activation of coagulation pathways at the level of factor X. Clinical examples would include amnionic fluid embolus, head injuries, neurosurgery, and certain malignancies. Some tumor cells appear to have a specific enzyme (cysteine protease), which will directly activate factor X. Alternatively with endothelial injury, there may be activation of the extrinsic pathway leading to DIC.

TABLE 48.11. CLASSIFICATION OF CLINICAL STATES ASSOCIATED WITH DISSEMINATED INTRAVASCULAR COAGULATION

Obstetric Complications	
Abruptio placentae	
Placenta previa	
Dead fetus	
Amniotic fluid infusion	
Placenta accreta	
Toxemia of pregnancy	
Cesarean section	
Abortion	
Hydatid mole	
Extrauterine pregnancy	
Forceps delivery	
Normal delivery	
Tissue Trauma	
Major surgery (especially extracorporeal circulation)	
Severe trauma and burns	
Fat embolism	
Rejection of transplant	
Heatstroke	
Hemolytic Processes	
Transfusion of mismatched blood	
Drowning	
Acute hemolysis secondary to infection	
Immune mechanisms	
Ingestion of acid and other causes	
Neoplastic Diseases	
Solid tumor	
Leukemia (Promyelocytic M-3, and Monocytic M-5)	
Snakebites	
Infections (Especially Acute Forms)	gram-negative, meningococcal, and pneumococcal septicemia
Bacterial:	
Rickettsial:	Rocky Mountain spotted fever
Viral:	hemorrhagic smallpox and hemorrhagic fever (Thai, Korean, and others)
Mycotic:	acute histoplasmosis
Parasitic:	malaria (blackwater fever)
Miscellaneous	
Cirrhosis of the liver	
Glomerulonephritis	
Acute pancreatitis	
Purpura fulminans	
Thrombotic thrombocytopenic purpura	
Hemolytic uremic syndrome	
Shock	
Severe progressive stroke	
Severe heart failure	
Giant hemangioma (Kasabach-Merritt)	
Large aortic aneurysm	

It is important to appreciate that DIC is a rapidly evolving pathophysiologic response to an underlying disease. Consequently, laboratory testing to establish the diagnosis often is difficult because of the constantly changing balance of underlying mechanisms. In acute DIC with a predominance of thrombin generation with consumption of coagulation factors and activation of fibrinolytic activity, bleeding is the most common clinical manifestation. Typically, acute DIC presents as systemic bleeding with oozing from the sites of venipunctures or recent surgery. However, in the subacute and chronic states such as retained dead fetus or an underlying disseminated malignancy, the patient may actually present with thrombosis. In this setting,

there may be actual compensatory increase of coagulation factors such as fibrinogen and factor VIII, which may exceed the range found in normal subjects.

In acute DIC, the constellation of laboratory findings reflect both the consumption of coagulation proteins as well as the presence of plasmin. Typically, the APTT, PT, and TT will be prolonged. In addition, there is consumption of platelets as well as the regulatory proteins AT and protein C. Fibrinogen/fibrin degradation products will be increased and the level of fibrinogen will be decreased. The availability of immunologic assays based on monoclonal antibodies to neoantigens has resulted in test systems that are more specific for fibrin degradation products. The D-dimer represents the use of an antibody to such a neoantigen.

In chronic DIC, the platelet count will typically be slightly decreased whereas the coagulation assays (APTT, PT) may be within normal limits. Fibrin/fibrinogen degradation products and D-dimer will be elevated and there may also be consumption of AT.

The treatment of choice for DIC is management of the underlying clinical condition. For instance, in a patient with chronic DIC and retained dead fetus, evacuation of the uterus is the treatment of choice. In a patient with acute DIC secondary to endotoxic shock, appropriate support and antibiotic therapy are treatments of choice.

THROMBOTIC THROMBOCYTOPENIC PURPURA/HEMOLYTIC UREMIC SYNDROME

Part of "48 - Coagulation Abnormalities"

TTP and hemolytic uremic syndrome (HUS) are examples of complex disease states that are closely related. In TTP there is systemic deposition of intraluminal platelets in the smaller arterioles. In contrast, HUS is characterized by involvement of the glomeruli with platelet aggregates. Both diseases present as a microangiopathic hemolytic anemia. Typically, there are schistocytes on the peripheral smear together with thrombocytopenia. Recent work would suggest the etiology of these conditions is variable from case to case. Moreover, the mode of presentation, particularly with TTP, may be characterized by either a single clinical episode, a chronic relapsing form, or an intermittent form. Approximately two thirds of patients with acute TTP appear to have abnormal processing of vWF, which results in the presence of large vWF multimers during periods of remission. It may be difficult to demonstrate these abnormal multimers during the acute episodes. The persistence of large vWF multimers after recovery is sensitive for predicting a relapse. The treatment of choice in TTP is plasma exchange or plasmapheresis plus fresh-frozen plasma. Recent work has suggested abnormalities of a plasma metalloprotease explain both chronic relapsing and acquired TTP (13, 14). In some instances, the deficiencies of metalloprotease are hereditary whereas in others it appears patients form antibodies to the metalloprotease inhibiting its enzymatic activity. As a consequence of reduced metalloprotease activity, there is an accumulation of high molecular weight vWF in the patient plasma predisposing to TTP.

LIVER DISEASE

Part of "48 - Coagulation Abnormalities"

Hemostatic problems encountered in patients with liver disease usually are complex. Because the liver plays a key role in the synthesis of coagulation proteins, regulatory proteins and factors involved in the fibrinolytic system, the balance between these various components may be altered to varying degrees in any patient with significant compromise of liver function. Also the liver plays a role in removing activated clotting factors from the circulation

as well as fibrinolytic breakdown products. In end-stage liver disease often the plasma level of fibrinogen may be decreased. However, in uncomplicated, cirrhosis, obstructive jaundice, chronic hepatitis, and biliary cirrhosis, a high level of fibrinogen may be found. Often in these latter cases, the thrombin time will be prolonged because of an acquired dysfibrinogenemia. This is thought to be because of an increased sialic acid content of fibrinogen.

Levels of the vitamin K dependent proteins are also decreased in liver disease associated with obstructive jaundice or malabsorption of vitamin K. Since factor VII has the shortest half-life of the vitamin K-dependent proteins, the first laboratory abnormality encountered may be an isolated prolongation of the PT.

Patients with liver disease often show a hypofibrinogenemia together with elevation of fibrin/fibrinogen degradation products and thrombocytopenia. This constellation of laboratory findings may be difficult to differentiate from DIC. In the past, many clinicians have relied on the factor VIII:C assay as a means of separating DIC from primary liver disease. Conventional wisdom suggested DIC will have low levels of factor VIII:C in contrast to the elevated levels seen in liver disease. However, this is not a reliable criterion. Often in advanced liver disease there is low-grade intravascular coagulation making the differential diagnosis of DIC and liver disease an exercise in futility.

Management of hemostatic abnormalities in the setting of liver disease is often frustrating. In acute liver failure, replacement of coagulation factors by infusion of fresh-frozen plasma may be helpful. In more chronic states, management of the underlying liver disease is the primary objective. The use of fresh-frozen plasma and platelets may be helpful to treat acute bleeding episodes. The treatment of bleeding esophageal varices may require balloon tamponade or surgery.

BLEEDING ASSOCIATED WITH CARDIOPULMONARY BYPASS

Part of "48 - Coagulation Abnormalities"

Approximately one quarter of a million patients undergo cardiopulmonary bypass (CPB) procedures each year in the United States. This group of patients utilized approximately 25% of blood products in many institutions. However, there is marked institutional variability in the incidence of bleeding associated with CPB. A variety of risk factors have been identified including time on the extracorporeal system, history of previous cardiac surgery, the nature of the surgery, preoperative medications, and the use and reversal of heparin. As a result of exposure of the circulating blood to a foreign surface, there is platelet activation with the release reaction occurring. In addition, there is also activation of the fibrinolytic system, which may lead to consumption of fibrinogen, factor V, factor VIII, and other plasma proteins. Also with activation of the coagulation cascade, there may be consumption of the regulatory proteins AT, protein C, and protein S.

Consequently, bleeding associated with CPB is multifactorial and requires a systematic approach to the bleeding patient. The first step is to rule out an anatomic defect as a potential explanation of post-operative bleeding. Also, coagulation tests should include an APTT, thrombin time, heparin assay, and fibrinogen. The thrombin time and heparin assay, will provide information as to the possibility of excess heparin or heparin rebound.

Management may require the use of platelets. Occasionally fresh-frozen plasma is indicated for serious coagulopathies. Recently, the use of desmopressin (DDAVP) and aprotinin [Trasylol (Bayer Corp., West Haven, Connecticut)] has been advocated as a means of controlling bleeding without the use of blood products and their potential complications.

LABORATORY ANALYSIS OF COAGULATION

Part of "48 - Coagulation Abnormalities"

The most important step in the evaluation of a bleeding patient is a good medical history. Unfortunately, in the clinical laboratory the history is not available unless the laboratory physician takes the initiative to review the patient's chart and talk with the patient. In the majority of cases, the information provided will direct the subsequent laboratory evaluation. The clinical history of patients with abnormalities of primary and secondary hemostasis differ significantly. Thus, mucocutaneous bleeding should lead to an evaluation of primary hemostasis while musculoskeletal bleeding would initiate an evaluation of secondary hemostasis.

In most laboratories, six simple procedures can be performed to evaluate a patient with a potential hemostatic abnormality. These six procedures include: a review of the peripheral smear, bleeding time, platelet count, APTT, PT, and thrombin time. In virtually all cases, the peripheral smear and platelet count are a byproduct of a CBC and are readily available. The bleeding time is useful for evaluating primary hemostasis. In conjunction with the platelet count, it allows a differentiation of qualitative and quantitative platelet disorders. Recently, a new *in vitro* method for evaluating primary hemostasis has become available. The PFA-100 (Dade Behring, Deerfield, Illinois) allows one to evaluate platelet function utilizing a flow system *in vitro*. The patient's blood sample is obtained and is perfused through two tiny apertures that have collagen/epinephrine or collagen/ADP. The end point of the test is when there is no longer perfusion of blood through the aperture (clot formation). This test system is extremely sensitive to aspirin. A patient exposed to aspirin will have a prolonged time to closure with the epinephrine/collagen and a normal ADP/collagen results. Also, the PFA-100 is very sensitive in the detection of mild type I von Willebrand disease.

The PT and APTT provide an arbitrary dissection of coagulation into the intrinsic (APTT) system and the extrinsic (PT) system. Although in reality this division is artificial, it is very useful from the standpoint of patient evaluation. Isolated prolongation of the prothrombin time is associated with factor VII deficiency, which may be either hereditary or acquired. An abnormal APTT together with a normal PT is indicative of an abnormality in the more proximal part of the intrinsic system (factors XII, XI, VIII, IX, and Fletcher and Fitzgerald). Abnormalities of both test systems are suggestive of a final common pathway defect (factors X, V, II, and I).

The thrombin time is a very useful test and it is often overlooked. The thrombin time may be abnormal in a variety of clinical situations (Table 48.12). Thrombin time is extremely useful in ruling out the presence of heparin in a patient sample. The

finding of a prolonged thrombin time that corrects with the addition of protamine is virtually diagnostic of heparin. Also, the reptilase time may be used to identify heparin as an explanation of prolonged thrombin time. The reptilase time will be normal in the presence of heparin.

TABLE 48.12. ABNORMAL THROMBIN TIME

- Heparin
- Fibrin/fibrinogen degradation products
- Paraproteins
- Dysfibrinogenemia
- Hypofibrinogenemia
- Antibodies to thrombin
- Uremia
- Some patients with lupus anticoagulants

Factor assays are indicated in many instances to establish the exact cause for the abnormal screening studies. These assays, depending on the factor, may utilize either the APTT or PT systems. Appropriate attention is necessary to quality control factor assays. At least two dilutions of the patient's plasma should be used and the patient's values should be checked to ensure parallelism of the patient's values when compared to the reference curve. Nonparallelism of factor assay results is suggestive of an inhibitor.

Synthetic substrate assays have achieved a greater level of importance in the coagulation laboratory. These assays rely on the use of synthetic oligopeptides (three or four amino acids) that mimic amino acid sequences near the site of enzymatic cleavage in the native substrate (1, 10). For instance, the synthetic substrate utilized to evaluate factor Xa activity consists of the four amino acids, isoleucine-glutamine-glycine-arginine. These four amino acids are located adjacent to the peptide bonds in prothrombin where factor Xa activates prothrombin to thrombin. An indicator group is attached at the carboxy-terminal end of the oligopeptide. The indicator group may either be a chromogen or a fluorochrome. In most instances, laboratories utilize chromogenic assays.

Among the advantages of synthetic substrate assays are their ready adaptation to automated instrumentation and their basic enzymatic principles. Thus, utilizing synthetic substrates, it is possible to characterize the reactions with greater precision than the classical coagulation assays relying on a fibrin endpoint. Currently, synthetic substrate assays are most often utilized for evaluation of the fibrinolytic system. Also, there are commercially available kits for many components of the coagulation system (e.g., factor VIII). In addition, synthetic substrates are extremely useful in analysis of the regulatory proteins (AT and protein C).

REFERENCES

1. Rizza C, Lowe G. *Haemophilia and Other Inherited Bleeding Disorders*. London: W.B. Saunders Company, Ltd., 1997.

2. Naylor JA, Green PM, Rizza CR, et al. Factor VIII gene explains all cases of hemophilia. *Lancet* 1992;340:1066–1067.

3. Naylor JA, Green PM, Montandon AJ, et al. Detection of three novel mutations in two hemophilia A patients by rapid screening whole essential regions of the factor VIII gene. *Lancet* 1991;337:635–639.

4. Klein U. Production and characterization of recombinant factor VIII. *Sem Haematol* 1991;28:17–21.

5. Scharrer I, Bray GL, Neutzling O. Incidence of inhibitors in haemophilia A patients — a review of recent studies of recombinant and plasma derived factor VIII concentrates. *Haemophilia* 1999;5:145–154.

6. Schmidt MI, Gamerman S, Smith HE, et al. Recombinant 'activated' factor VII (rFVIIa) therapy for intracranial haemorrhage in haemophilia A patients with inhibitors. *Am J Hematol* 1994;47:36–40.

7. Seligsohn U. High gene frequency of factor XI (PTA) deficiency in Ashkenazi Jews. *Blood* 1978;51:1223–1228.

8. Bolton-Maggs PHB, Wensley RT, Kernoff, PBA, et al. Production and therapeutic use of a factor XI concentrate from plasma. *Thromb Haemost* 1992;67:314–319.

9. Board PG, Losowsky MS, Miloszewski KJA. Factor XIII: inherited and acquired deficiency. *Blood Rev* 1993;7:229–242.

10. Olson JD, Arkin CF, Brandt JT, et al. Laboratory monitoring of unfractionated heparin therapy. College of American Pathologists Conference XXXI on Laboratory Monitoring of Anticoagulant Therapy. *Arch Pathol Lab Med* 1998;122:782–798.

11. Brandt JT, Triplett DA, Alving B, et al. Criteria for the diagnosis of lupus anticoagulants. An Update. *Thromb Haemost* 1995;1185–1190.

12. Triplett DA. Antiphospholipid protein antibodies: laboratory detection and clinical relevance. *Thromb Res* 1995;78:1–31.

13. Tsai HM, Lian ECY. Antibodies to von Willebrand factor — cleaving protease in acute thrombotic thrombocytopenic purpura. *N Engl J Med* 1998;339:1585–1594.

14. Furlan M, Robles R, Solenthaler M, et al. Deficient activity of von Willebrand factor cleaving protease in chronic relapsing thrombotic thrombocytopenic purpura. *Blood* 1997;89:3097–3103.

SUGGESTED READINGS

Gould MK, Dembitzer AD, Doyle RL, et al. Low-molecular-weight heparins compared with unfractionated heparin for treatment of acute deep venous thrombosis. *Ann Intern Med* 1999;130:800–809.

Graham BJM, Moodley I. Economic evaluation of low-molecular-weight heparin in the management of patients with unstable coronary artery disease. *S Afr Med J* 1999;89:C20–C23.

Pettila V, Kaaja R, Leinonen P, et al. Thromboprophylaxis with low molecular weight heparin (dalteparin) in pregnancy. *Thromb Res* 1999;96:275–282.

Ballard JO. Anticoagulant-induced thrombosis. *JAMA* 1999;282:310–312.

Meyer O, Salama A, Pittet N, et al. Rapid detection of heparin-induced platelet antibodies with particle gel immunoassay (ID-HPF4). *Lancet* 1999;354:1525.

Kadidal VV, Mayo DJ, Horne MK. Heparin-induced thrombocytopenia (HIT) due to heparin flushes: a report of three cases. *J Int Med* 1999;246:325–329.

Newman PM, Chong BH. Further characterization of antibody and antigen in heparin-induced thrombocytopenia. *Br J Hematol* 1999;107:303–309.

Izban KF, Lietz HW, Hoppensteadt DA, et al. Comparison of two PF4/heparin ELISA assays for the laboratory diagnosis of heparin-induced thrombocytopenia. *Sem Thromb Hemost* 1999;25:51–56.

Hassell K. Anticoagulation in heparin-induced thrombocytopenia: an ongoing challenge. *Hosp Physician* March 2000:56–61.

Onuigbo M, Cuffy-Hallam M, Mangano A, et al. Successful thrombolytic therapy for massive pulmonary embolism. *So Med J* 2000;93:327–329.

Ozawa T, Takikawa Y, Niiya K, et al. Factor VII Morioka (FVII L-26P): a homozygous missense mutation in the signal sequence identified in a patient with factor VII deficiency. *Br J Haematol* 1998;101:47–49.

Ortel TL. Clinical and laboratory manifestations of anti-factor V antibodies. *J Lab Clin Med* 1999;133:326–334.

Oliver JA, Monroe DM, Roberts HR, et al. Thrombin activates factor XI on activated platelets in the absence of factor XII. *Arterioscler Thromb Vasc Biol* 1999;19:170–177.

Daly HM, Carson PJ, Smith JK. Intracerebral haemorrhage due to acquired factor XIII inhibitor — successful response to factor XIII concentrate. *Blood Coagulation and Fibrinolysis* 1991;2:507–514.

Jones SL. HELLP! A cry for laboratory assistance: a comprehensive review of the HELLP syndrome highlighting the role of the laboratory. *Hematopathol Molec Hematol* 1998;11:147–171.

Mitra D, Jaffe EA, Weksler B, et al. Thrombotic thrombocytopenic purpura and sporadic hemolytic uremic syndrome plasmas induce apoptosis in restricted lineage of human microvascular endothelial cells. *Blood* 1997;89:1224–1234.

Helft G, Bartolomeo P, Zaman AG, et al. The heparin management test: a new device for monitoring anticoagulation during coronary intervention. *Thromb Res* 1999;96:481–485.

49
Thrombophilia

Douglas A. Triplett

This chapter focuses on abnormalities of the regulatory proteins and other clinical states associated with thrombotic disorders. Thromboembolic disease and its complications are the most common cause of death in the western world. Although several hereditary abnormalities of regulatory proteins have now been linked to a familial thrombotic tendency, most patients with thrombosis have a complex multifactorial pathophysiology. The laboratory contribution to the diagnosis of thrombotic disorders, therefore, is most clearly defined in the case of patients who have hereditary thrombophilia. Thrombophilia is an acquired or hereditary condition(s) that predisposes to an increased incidence of thrombosis (1). The hereditary disorders are reviewed and various acquired states in which there is a thrombotic predisposition are discussed.

HEREDITARY DISORDERS OF REGULATORY PROTEINS

Part of "49 - Thrombophilia"

Antithrombin (AT) is a plasma glycoprotein that is a member of the SERPIN family (serine protease inhibitors) (2). AT inhibits all the serine proteases involved in hemostasis. This inhibition is achieved by forming a 1:1 enzyme-inhibitor complex. The reaction between AT and serine protease enzymes is greatly accelerated in the presence of heparin or other related glycosaminoglycans. Physiologically, the presence of heparan sulfate on the surface of endothelial cells is involved in the downregulation of hemostasis in conjunction with AT. Several investigators have utilized activation peptides (fibrinopeptide A, prothrombin fragment 1·2, and factor X activation peptide) to demonstrate there is an ongoing tonic activation of hemostasis. This low level of activation is successfully regulated in the normal individual by a variety of regulatory systems of which the heparan sulfate–AT component is extremely important. Hereditary deficiency of AT was the first plasma-protein deficiency successfully linked to a thrombotic predisposition (3). In contrast to hereditary disorders of coagulation proteins, AT-deficient patients were found to have levels of 40% to 50% of normal. The association of clinical symptoms with a borderline or slight deficiency of this regulatory protein would suggest the physiologic balance of procoagulant/anticoagulant activities is precarious. In contrast, patients with mild hemophilia and levels of factor VIII activity of 5% may be entirely free of bleeding complications unless significantly challenged by trauma or surgery.

AT deficiency is inherited as an autosomal dominant abnormality. The AT gene is located on human chromosome 1q23-25. The gene consists of six exons and five introns distributed over approximately 19 kb of DNA. There is some homology of the AT gene with the α_1-antitrypsin gene. The incidence of AT III deficiency in the general population has been estimated at 1:2,000 to 1:5,000.

Individuals with AT deficiency are usually symptomatic by the age of 40. Typically, the onset of thromboembolic events occurs in the late teens or 20s. On occasion, patients many be encountered who are in their 50s or 60s with no history of thrombosis. Most thromboembolic events are venous. In approximately 50% of cases, the initial event is related to some precipitating factor such as surgery, trauma, use of oral contraceptives, or pregnancy (4,5).

The laboratory evaluation requires specific assays of AT activity and antigen. The usual screening tests for hemostasis are entirely normal. Antigenic measurement of AT may use different techniques: radioimmunoassay, radial immunodiffusion, and Laurell rocket immunoelectrophoresis. Of these, techniques, the Laurell method is most commonly employed. Functional assays may utilize either fibrin end-point or synthetic substrate systems. The most practical functional assay assesses the ability of AT to inhibit an enzyme in the presence of heparin. The most frequently utilized system employs chromogenic substrates (Fig. 49.1). A second functional assay relies on progressive AT activity. In this assay, the patient's plasma containing AT is incubated with a known amount of enzyme (Xa or IIa) in the absence of heparin. Discrepancies between these two functional assays have been observed in a number of patient studies. Patients with normal progressive AT activity and decreased activity in the presence of heparin have gene mutations resulting in abnormalities of heparin binding sites.

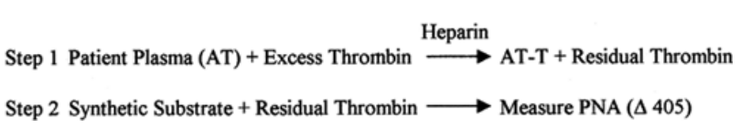


FIGURE 49.1. Antithrombin amidolytic assay. The reaction may be measured using a recording spectrophotometer or by an end-point method when the reaction is stopped by adding acetic acid. PNA, para-nitroaniline.

By utilizing both antigenic and functional assays, it is possible to divide inherited AT deficiency into two distinct types. Type I results in reduced synthesis of a biologically normal protein and type II results in normal antigenic amounts of AT but decreased functional activity in one of the two above functional assay systems (Table 49.1). Patients with type II AT deficiency secondary to an abnormality in the heparin-binding sites often have a milder clinical history and, in fact, may be asymptomatic in the heterozygous state. However, children who are doubly heterozygous for type II AT deficiency may have significant clinical thrombosis.

Acquired AT deficiency may be seen in a variety of clinical conditions. These include liver disease, nephrotic syndrome, disseminated intravascular coagulation, and acute thrombosis. Modest reductions may also be seen in patients using oral contraceptives. L-Asparaginase is associated with a striking decrease of AT.

TABLE 49.1. CLASSIFICATION OF ANTITHROMBIN DEFICIENCY

Type I	(↓ Ag & ↓ Act)
Type II	-RS
Type II	-HBS
Type II	-PE

Type II antithrombin deficiency is characterized by normal levels of antithrombin antigen and decreased inhibitory activity of thrombin and other serine proteases. Several variants of Type II have been described. Type II-RS involves mutations in the reactive site (RS) of the antithrombin molecule. Type II-HBS is characterized by variations in the heparin binding sites due to mutations of the antithrombin gene. In type II-HBS, the ability of heparin to interact with the binding sites is diminished. Type III-PE refers to a variant of antithrombin that has a pleiotropic affect. In this situation, more than one aspect of antithrombin function may be affected.

The clinical management of patients with AT deficiency consists of initial treatment with heparin followed by oral anticoagulants. When the diagnosis has been established, many clinicians maintain the patient on lifelong oral anticoagulant therapy. With the trend toward lower doses of oral anticoagulants and the use of thromboplastins with International Sensitivity Index (ISI) values of 1.0 to 1.7, the complication rate associated with the use of these drugs has decreased. An AT concentrate is also available. The concentrate is valuable in settings of surgery or trauma.

The diagnosis of any of the hereditary disorders of regulatory proteins is often complicated by the initial patient presentation during an acute thrombosis. Frequently, the patient is already receiving heparin before the laboratory is requested to evaluate the patient for possible regulatory protein deficiencies. In this setting, testing other family members may prove to be the best means of establishing a hereditary abnormality.

Protein C Deficiency

Protein C is a vitamin K–dependent protein that is readily converted to active protein C (APC) by a complex of thrombin and thrombomodulin on the surface of endothelial cells (2) (Fig. 49.2). APC in the presence of another vitamin K–dependent protein, protein S, inactivates factors Va and VIIIa. Thus, the protein C system is responsible for the regulation of the two essential cofactors in the intrinsic pathway of coagulation. The remaining cofactor (tissue factor) of the extrinsic pathway is regulated by a relatively newly discovered coagulation protein: extrinsic pathway inhibitor [also known as tissue factor pathway inhibitor (TFPI)].

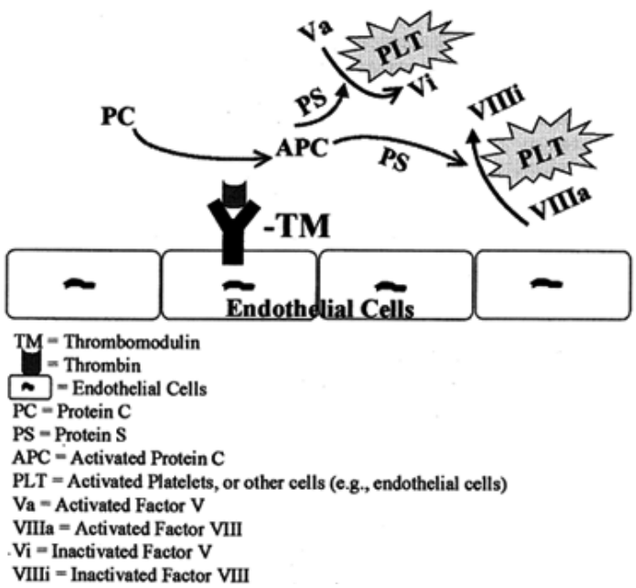


FIGURE 49.2. Protein C/S system. Proteins C and S are vitamin K–dependent proteins that participate in the downregulation of hemostasis. When thrombin is generated *in vivo*, it converts fibrinogen to fibrin. In addition, thrombin may also bind to an endothelial receptor, thrombomodulin. As the name implies, when thrombin binds to thrombomodulin, it loses it procoagulant activity and actively participates in the protein C/S anticoagulant pathway. Thrombin-thrombomodulin converts protein C to activated protein C (APC). APC in the presence of protein S (a cofactor) will inactivate the activated forms of factors V and VIII, thus inhibiting the generation of thrombin. Patients who have hereditary deficiencies of proteins C and S are predisposed to venous thromboembolic events.

Protein S functions in the above reaction as a cofactor for APC. Protein S binds both to endothelial cells and activated platelets and thus serves as a means of localizing APC to an appropriate cell surface. The complex of APC/protein S will then preferentially inactivate cofactors Va and VIIIa. APC is inhibited by at least two plasma proteins: protein C inhibitor and α_1-antitrypsin.

The gene for protein C is found on chromosome 2. The gene contains approximately 11 kb of DNA and has nine exons with eight introns. Protein C deficiency is inherited in an autosomal dominant fashion. Heterozygous patients present clinically with either deep or superficial vein thrombosis. The time of onset and other related clinical findings are very similar to those of AT deficiency. In addition, heterozygous protein C deficiency may be first suspected when a patient who presents with deep vein thrombosis is started on a loading dose of oral anticoagulants.

The appearance of coumadin-induced skin necrosis may alert the clinician to the possibility of protein C or S deficiency states (6). Not all patients who develop this complication prove to have hereditary deficiency of protein C. However, any patient with such a history should be evaluated for this possibility.

Doubly heterozygous or homozygous protein C deficiency is associated with purpura fulminans in the newborn. This is a particularly devastating presentation characterized by extensive areas of skin necrosis together with cerebral and retinal vein thrombosis. Unless appropriate treatment is initiated immediately, this disorder is fatal.

Laboratory evaluation of protein C includes both antigenic and functional assays. The antigenic assays available include Laurell rocket immunoelectrophoresis and enzyme-linked immunosorbent assay (ELISA). Patients with heterozygous protein C deficiency typically have a protein C level of 40% to 50% (reference interval, 65% to 130%). These results are similar to values for heterozygous AT deficiency.

Functional assays for protein C are based on either a fibrin end-point system or synthetic substrate technology. The fibrin end-point assays rely on an activated partial thromboplastin time (aPTT) system (Fig. 49.3). The synthetic substrate assays are more direct, requiring less technical time. Several situations have been described in which the functional coagulation assay is discrepant when compared with the synthetic substrate assay (Table 49.2). Both type I and II deficiency states have been identified. In type I deficiency, the patient has a decreased plasma level of both functional and antigenic protein C. In type II patients, the antigenic level of protein C is within normal limits, although there is decreased functional coagulation and/or substrate activities.

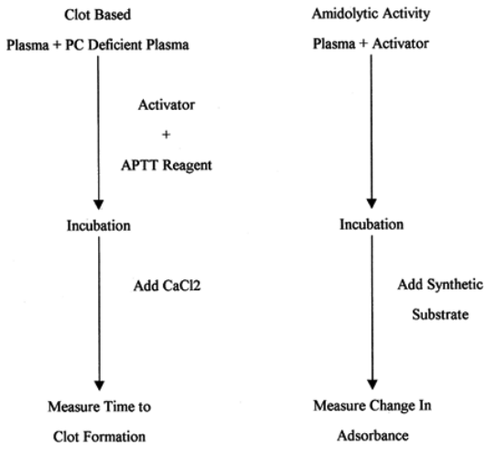


FIGURE 49.3. Protein C functional assays (venom activation).

TABLE 49.2. PROTEIN C/S SYSTEM

	PC Activation	APC Activity
Enzyme	Thrombin	APC
Cofactor	Thrombomodulin	Protein S/factor V
Surface	Endothelium	Plts/endothelium
Substrate	Protein C	Va, VIIIa

PC, protein C; APC, active protein C; Plts, platelets.

When the diagnosis has been established, treatment of protein C deficiency in the heterozygous patient is lifelong oral anticoagulation. In the homozygous patient, replacement of protein C using fresh-frozen plasma is critical in the initial stages of treatment. Oral anticoagulants are indicated for long-term management. The availability of protein C concentrates may prove to be valuable in treating these patients.

Protein S Deficiency

The clinical presentation of protein S deficient patients is similar to AT III and protein C deficiency states (6). In addition to venous thromboembolic events, several recent reports emphasized cerebral vein thrombosis (1,2). Also, warfarin-induced skin necrosis has been reported with protein S deficiency.

There are two forms of protein S present in the circulation: free and bound. The free protein S comprises approximately 40% of the plasma compartment and is the functionally active cofactor. In contrast, the bound protein S circulates complexed to a complement component: C4b-binding protein. C4b-binding protein is an acute phase reactant; consequently, a variety of clinical conditions in which acute phase proteins increase will result in some shift in the equilibrium between free and bound protein S.

The laboratory evaluation of protein S includes antigenic assays and functional assays. For practical purposes, most laboratories are restricted to the use of antigenic determinations. Total protein S may be assayed with a radioimmunoassay or ELISA. Free protein S is based on removing the protein S—C4b-binding complex by precipitation with polyethylene glycol. The supernatant is then assayed for residual protein S by either ELISA or radioimmunoassay.

Protein S deficiency has been categorized into two types. Type I includes patients in whom there are markedly reduced levels of free protein S but a normal amount of protein S complexed to C4b-binding protein. The type II patients have decreased amounts of both free and bound protein S. Most of the patients described have been type I. Acquired deficiencies of protein S have been described in patients with lupus anticoagulants. In addition, protein S has been found to decrease during pregnancy and in the postpartum period. Oral contraceptives have also been found to lower both total and free protein S. Because protein S is a vitamin K–dependent protein, plasma levels are decreased in liver disease and vitamin K deficiency.

Treatment of protein S deficiency requires long-term oral anticoagulant therapy.

Heparin Cofactor II Deficiency

Heparin cofactor II is a member of the SERPIN family. It is a single-chain polypeptide with a molecular weight of approximately 66,000. The gene for heparin cofactor II is located on

chromosome 22. In contrast to AT, heparin cofactor II inhibits thrombin but has no inhibitory activity toward other coagulation serine proteases. The ability to inhibit thrombin is greatly accelerated in the presence of high concentrations of heparin. It appears that the physiologic activator for heparin cofactor II is dermatan sulfate, which accelerates heparin cofactor II activity approximately 1,000-fold. Several patients have been described in whom hereditary deficiency of heparin cofactor II was associated with thrombosis. In these cases, the level of heparin cofactor II was approximately 40% to 60% of normal. Acquired heparin cofactor II deficiency has been seen in patients with disseminated intravascular coagulation, liver disease, and obstetric complications. Laboratory assays for heparin cofactor II rely on the use of dermatan sulfate as a means of preferentially accelerating the thrombin inhibitory activity of heparin cofactor II.

Hereditary Disorders of Fibrinolysis

Several hereditary abnormalities of the fibrinolytic system have been linked to a thrombotic predisposition. Perhaps the most frequently discussed have been dysplasminogenemia and hypoplasminogenemia (7). Unfortunately, the family histories in many of these cases have not proved to be as convincing as those seen in the hereditary disorders of AT, protein C, and protein S. Often, the propositus is the only symptomatic member of the family despite the fact that other family members have similar levels of plasminogen or the same dysfunctional molecule.

Less convincing cases have been reported in which it has been postulated that there is a decreased ability to release tissue plasminogen activator (t-PA) from the endothelial cells or alternatively elevated levels of plasminogen activator inhibitor (PAI). This disequilibrium between t-PA and PAI leads to a hypofibrinolytic state. Again, a hereditary abnormality of this ratio has been convincingly documented in few families. In many acquired situations such as hypertriglyceridemia associated, for example, with diabetes mellitus, systemic lupus erythematosus, abnormalities of the PAI:t-PA ratios have been documented. It appears that, in most instances, the altered PAI:t-PA ratio is acquired. European investigators have emphasized the frequency with which this is encountered when patients with thromboembolic diseases are screened for these two assays in a basal and stimulated state.

The laboratory evaluation of the fibrinolytic system includes measurement of plasminogen (antigenic and functional) and t-PA (antigenic and functional). In addition, it is important to measure PAI levels (antigenic). The stimulation tests are based on the concept of drawing baseline t-PA and PAI levels followed by venous occlusion at 100 mm Hg for 10 to 20 minutes. After stimulation, samples are obtained for repeat PAI and t-PA levels. Intravenous infusion of 1-desamino-8-D-arginine vasopressin (DDAVP) has also been used as a stimulatory test.

Dysfibrinogenemia

Several congenital dysfibrinogenemias have been reported in association with familial thrombosis. Both arterial and venous thromboembolic events have been described. In some cases, the molecular defect has been identified. It may involve abnormalities in either the thrombin-fibrinogen interaction or plasmin-fibrinogen interaction. In the latter case, the binding of plasminogen to the fibrin clot is impaired with resulting increased resistance of fibrin to fibrinolysis.

ACQUIRED DISORDERS

Part of "49 - Thrombophilia"

Various acquired conditions are associated with a predisposition to thrombosis. The presence of a lupus anticoagulant (LA) has been linked to thrombotic events (8). Whether the LA is causative, coincidence, or a consequence is still unresolved. Consequently, patients who present with unexplained thrombosis should be evaluated for the presence of LA. This evaluation includes a screening aPTT. In addition, because LA is merely one member of the family of antiphospholipid antibodies, solid phase assays for anticardiolipin antibodies (ACA) and antibodies to $ß_2$ glycoprotein I and prothrombin should be obtained. In most instances, patients who are LA positive will also be ACA positive. In approximately one third of cases, there will be lack of concordance between these two assay systems. LA-associated thromboembolic events are both venous and arterial. Consequently, the clinical presentation of these patients may be significantly different than the situation encountered with hereditary thrombotic disorders. For instance, young adults with cerebrovascular accidents are frequently found to have LA/ACA (30% of strokes under age 50).

Numerous other clinical conditions are associated with thrombotic predispositions. Among these are Trousseau's syndrome, diabetes mellitus, hyperlipidemia, various vasculitides, paroxysmal nocturnal hemoglobinuria, and Kawasaki's syndrome.

Trousseau's syndrome is the presence of recurrent deep or superficial venous thrombosis in association with malignancies. Often the malignancy may be gastrointestinal or pancreatic in origin. In many cases, the malignancy may be occult and the patient presents initially with thrombosis. Patients who present with deep vein thrombosis or superficial thrombophlebitis that fails to respond to oral anticoagulant therapy or antiinflammatory medications should be suspected of possible Trousseau's syndrome. Typically, heparin is the only satisfactory therapy for management of these patients. The mechanism of increased procoagulant activity in association with malignancy has recently been extensively explored. Initially, it was believed that mucus-secreting tumors produced a substance that directly activated factor X to Xa. In addition, a tumor cell cysteine protease has been identified.

The presence of antiendothelial antibodies has been linked to several thrombotic conditions including Kawasaki's syndrome, hemolytic uremic syndrome, and acute allograft rejection.

LABORATORY EVALUATION OF A THROMBOPHILIA

Part of "49 - Thrombophilia"

Candidates for a thrombotic evaluation include young adults with a positive family history of thromboembolic events who present with a thrombosis. There may or may not be a precipitating

event. The evaluation of these patients requires assays for AT, protein C, protein S as well as molecular assays for factor V Leiden, prothrombin G20210A, among others (Table 49.3 and Table 49.4) (9). Establishing the diagnosis of a hereditary disorder associated with thrombosis is important. Not only from the standpoint of patient management but also family counseling and appropriate precautions for family members who are found to be heterozygous for a deficiency with no history of thrombosis. For instance, young women with a family history of factor V Leiden should be cautioned regarding the use of oral contraceptives. Management of pregnancy in these patients may require the use of low molecular weight heparin.

TABLE 49.3. VENOUS THROMBOSIS RISK FACTORS

Genetic[a]	Acquired/Environmental	Mixed
MTHFR (polymorphism)	History of prior thrombosis	↑ Factor VIII
Factor V Leiden (FVR506Q)	Age	↑ Homocysteinemia
Prothrombin 20210A	Surgery/trauma	↑ Fibrinogen
Protein S deficiency	Extended bed rest	
Protein C deficiency	Malignancy	
Antithrombin deficiency	Oral contraception	
Dysfibrinogenemias	Hormonal replacement	
	Antiphospholipid antibodies	
	Pregnancy/puerperium	
	Myeloproliferative disorders	
	Obesity/inactivity	

[a] Listed in order of frequency. Boldface denotes entries that are most common high risk factors. Note there is a relatively low incidence of venous thrombosis in Asian and African populations.
MTHFR, methylenetetrahydrofolate reductase.
Modified from Rosendaal FR. Risk factors for venous thrombosis: Prevalence, risk and interaction. *Semin Hematol* 1997; 34:171-187.

TABLE 49.4. ARTERIAL THROMBOSIS: RISK FACTORS

Genetic[a]	Acquired/Environmental	Mixed
Hypercholesterolemia	Anticardiolipin antibodies	↑ Factor VIII
↓ HDL	Antioxidized LDL	↑ Fibrinogen
Hypertriglycerides	Diet (↑ fat, ↑ carbohydrates)	↑ Factor VII
Diabetes mellitus	Smoking	↑ Homocysteine
↑ Lipoprotein(a)	Hypertension	↑ C reactive protein
Hemochromatosis (2 genes)	Infections (e.g., chlamydia, CMV)	↑ von Willebrand factor
PLA_2 (glycoprotein IIa/IIIb)	Social class	
Factor VII polymorphisms	Body mass index	
Hyperhomocysteinemia (Variant MTHFRs)		
PAI-1 polymorphisms		
Gender		
Family history of accelerated ASHD		

[a] Boldface denotes genetic based disorders identified by molecular testing in the laboratory.
HDL, high density lipoproteins; CMV, cytomegaloinclusion virus; MTHFRs, methylenetetrahydrofolate reductases; ASHD, arteriosclerotic heart disease.

Because approximately 50% of thrombotic events in patients are unexplained, the future exploration of these patients offers the possibility of a number of new assays. No doubt, a significant number of patients will be found to have abnormalities of endothelial regulatory functions including deficiencies or functional abnormalities of thrombomodulin as well as other regulatory components such as tissue factor expression and upregulation of adhesive proteins.

ACTIVATED PROTEIN C RESISTANCE AND FACTOR V LEIDEN

Part of "49 - Thrombophilia"

The preceding portion of this chapter focused on a chronologic sequence of discovery of abnormalities of plasma proteins that predispose to thrombosis (Table 49.5). The description of Dahlback et al. (10) of a family with a strong autosomal dominant pattern of thrombosis was first reported in 1993. Dahlback et al. evaluated the family with all the available assays that were known to be of value in identifying thrombophilia. All these studies were negative. They subsequently performed an experiment that involved adding APC to the patient's plasma and evaluating the degree of prolongation of the aPTT. In the normal control, the aPTT was significantly prolonged; however, in the patient samples, there was a blunted response to APC. Consequently, Dahlback et al. used the term APC resistance (APC-R), to describe the underlying patient problem that appeared to be linked with thrombosis. Very shortly after their description, Bertina and colleagues (11) at the University of Leiden identified a single amino acid substitution in factor V that led to APC-R.

This mutation affected a cleavage site in factor Va. This site was at 506 and involved the substitution of glutamine for arginine. Consequently, there was decreased ability to downregulate factor Va, thus predisposing to thrombosis. Subsequent studies identified factor V Leiden/APC-R only in the white population. In some parts of southern Sweden, the frequency of factor V Leiden is 15% (12). In southern Europe, the frequency of factor V Leiden was 3% to 4%. There are clear geographic differences in the incidence of factor V Leiden. Based on additional studies, it is thought that the mutation occurred in Europe approximately 30,000 to 40,000 years ago.

TABLE 49.5. HISTORY OF THROMBOPHILIA

1856	Virchow	Thrombosis hypothesis
1965	Egeberg	Antithrombin deficiency
1981	Griffin et al.	Protein C deficiency
1994	Comp et al.	Protein S deficiency
1993	Dahlback et al.	APC-R
1994	Bertina et al.	Factor V Leiden
1996	Poort et al.	Prothrombin G20210A

APC-R, activated protein C-resistance.

Factor V Leiden is found in approximately 20% of patients who present with an initial deep vein thrombosis. In patients who present with a history of two or more venous thromboembolic events, 30% to 50% of these individuals will have factor V Leiden (13). Consequently, it is very important to consider testing for factor V Leiden in patients with venous thromboembolic events. In many cases, there are other contributing factors (e.g., history of recent trauma, use of oral contraception, hormonal replacement therapy, pregnancy, prolonged bed rest).

Heterozygous factor V Leiden patients have approximately a two- to fourfold increased relative risk of thrombosis. It has been estimated that carriers without a history of thrombosis have an absolute risk of thrombosis of approximately 0.45% per year; however, this increases to approximately 1% per year by age 60.

Individuals with homozygosity for factor V Leiden have a significantly increased relative risk of thrombosis. In women who use oral contraception, the presence of homozygosity for factor V Leiden results in a relative increased risk of approximately 80- to 100-fold. In addition, it is important to consider the possibility of underlying factor V Leiden in patients who present with cerebral vein thrombosis.

There are two functional assays available to test for APC-R (14). The original assay utilized patient plasma. APC is added to the patient plasma and the prolongation of the aPTT is determined. A source of normal plasma is used as a reference value. The results are expressed as a ratio:

$$\frac{\text{Patient plasma} + \text{APC}}{\text{Normal plasma}}$$

In most assays, values greater than 2.2 are considered to be within the normal reference interval. Values less than 2.2 are consistent with APC-R. Recently, a second assay has been introduced. In this assay system, the patient plasma is diluted one part to four parts factor V–deficient plasma. This assay is highly specific for identifying factor V Leiden. It eliminates the positive results that one may see in the original assay that are not attributable to factor V Leiden. Conditions that affect the original assay include lupus anticoagulants, pregnancy, and underlying chronic inflammation.

Use of molecular testing to identify factor V Leiden is readily available in most reference laboratories (15). Various techniques have been used including polymerase chain reaction (PCR) and, more recently, the Invader OS(Third Wave Technologies, Madison, WI). The latter assay is quicker and costs less than PCR. It is possible to see discrepancies between gene-based assays and APC-R. There are familial cases in which the baseline level of factor VIII is significantly elevated (greater than 150%). In these cases, an acquired APC-R is readily demonstrated in the original assay for APC-R. The presence of persistent elevated factor VIII levels confirms a risk of thrombosis analogous to or slightly greater than heterozygous factor V Leiden.

PROTHROMBIN GENE MUTATION

Part of "49 - Thrombophilia"

In 1994, Poort and colleagues (16) at the University of Leiden described a mutation in the prothrombin gene. This mutation is found in the white population and is most frequent in southern Europe as opposed to northern Europe. The defect involves the 3′ untranslated region of the prothrombin gene (G20210A). The frequency in the population is approximately 1.5% to 3%. Approximately 4% to 8% of the patients presenting with the first deep vein thrombosis are found to have this mutation. In individuals with recurrent venous thromboembolic events, it occurs at a frequency of 15% to 18%. The risk of thrombosis is these patients appears to correlate with levels of prothrombin activity in plasma.

In contrast to factor V Leiden, most patients with prothrombin G20210A present with superficial venous thrombophlebitis. There is also a relatively high frequency of cerebral vein thrombosis in patients with this mutation (17,18). Often, the cerebral vein thrombotic events are seen in young women who have started using oral contraception.

Unlike factor V Leiden, the prothrombin mutation has been linked to arterial disease. In young women who are smokers, there is approximately a 10-fold increased risk of myocardial infarction (19). Other risk factors such as obesity, hypertension, and diabetes may act synergistically with G20210A.

OTHER HEREDITARY RISK FACTORS PREDISPOSING TO VENOUS/ARTERIAL THROMBOSIS

Part of "49 - Thrombophilia"

With the ready availability of molecular testing, many polymorphisms have been identified as potential causes of both venous and arterial thrombosis. Among these are factor VII polymorphisms, familial elevated levels of factor VIII, abnormalities of PAI-1, and polymorphisms of fibrinogen. Currently, laboratory testing for these is not readily available. There is a need for more data and better clinical studies to evaluate these polymorphisms.

ARTERIAL THROMBOSIS

Part of "49 - Thrombophilia"

Arterial thrombosis and atherosclerosis present an even greater challenge with respect to identifying tests and their relevance to clinical management. More than 50% of patients who present with myocardial infarction are found to have normal lipids (20). Consequently, investigators have been searching for other risk factors. One of the most prominent has been homocysteine (21). Elevated levels of plasma homocysteine may be seen in individuals with underlying molecular defects. Abnormalities of the gene coding for methylenetetrahydrofolate reductase (MTHFR) have been described (22). The C677T polymorphism that leads to a labile variant of MTHFR is the most commonly studied. Approximately 35% to 40% of the white population are heterozygous for labile MTHFR and an additional 10% are homozygous (23). In cases of folate deficiency, patients heterozygous for labile MTHFR may develop increased plasma levels of homocysteine (24).

Homocysteine normal reference intervals are often quoted as 6 to 15 µmol/L. However, it is important to appreciate that the relative risk of thrombosis is based on a gradient approach to reporting homocysteine levels. Optimally, individuals should have a homocysteine level less than 9 µmol/L.

Other arterial risk factors include lipoprotein(a), antiphospholipid antibodies, fibrinolytic capacity, fibrinogen, and apolipoprotein A-I. The clinical utility of these markers remains limited.

REFERENCES

1. Spencer FA, Becker RC. Diagnosis and management of inherited and acquired thrombophilias. *J Thromb Thrombolysis* 1999;7:91–104.

2. Pabinger I, Schneider B. Thrombotic risk in hereditary antithrombin III, protein C, or protein S deficiency. *Arterioscler Thromb Vasc Biol* 1996;16:742–748.

3. Egeberg O. Inherited antithrombin deficiency causing thrombophilia. *Thromb Diath Haemorrh* 1965;13:516.

4. Van Boven HH, Vanderbroucke JP, Briet E, et al. Gene-gene and gene-environment interactions determine risk of thrombosis in families with inherited antithrombin deficiency. *Blood* 1999;94:2590–2594.

5. Martinelli I, Mannucci PM, De Stefano V, et al. Different risks of thrombosis in four coagulation defects associated with inherited thrombophilia: a study of 150 families. *Blood* 1998;92:2353–2358.

6. Gailani D, Reese EP Jr. Anticoagulant-induced skin necrosis in a patient with hereditary deficiency of protein S. *Am J Hematol* 1999;60:231–236.

7. Schuster V, Seidenspinner S, Zeitler P, et al. Compound-heterozygous mutations in the plasminogen gene predispose to the development of ligneous conjunctivitis. *Blood* 1999;93:3457–3466.

8. Triplett DA. Antiphospholipid protein antibodies: laboratory detection and clinical relevance. *Thromb Res* 1995;78:1–31.

9. Rodeghiero F, Tosetto A. Activated protein C resistance and factor V Leiden mutation are independent risk factors for venous thromboembolism. *Ann Intern Med* 1999;130:643–650.

10. Dahlback B, Carlsson M, Svensson PJ. Familial thrombophilia due to a previously unrecognized mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C. *Proc Natl Acad Sci U S A* 1993;90:1004–1008.

11. Bertina RM, Koeleman BPC, Koster T, et al. Mutation in blood coagulation factor V associated with resistance in activated protein C. *Nature* 1994;369:64–67.

12. Lowe GDO, Rumley A, Woodward M, et al. Activated protein C resistance and the FV:R ^{506}Q mutation in a random population sample. *Thromb Haemost* 1999;81:91800-924.

13. Svensson PJ, Zoller B, Mattiasson I, et al. The factor VR506Q mutation causing APC resistance is highly prevalent amongst unselected outpatients with clinically suspected deep venous thrombosis. *J Intern Med* 1997;241:379–385.

14. Rodeghiero F, Tosetto A. Activated protein C resistance and factor V Leiden mutation are independent risk factors for venous thromboembolism. *Ann Intern Med* 1999;130:643–650.

15. Spitzer SG. Molecular testing for the factor V-Leiden mutation in the clinical laboratory. *Lab Med* 1999;30:681–683.

16. Poort SR, Rosendaal FR, Reitsma PH, et al. A common genetic variation in the 3′-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 1996;88:3698–3703.

17. Martinelli I, Sacchi E, Landi G, et al. High risk cerebral-vein thrombosis in carriers of a prothrombin-gene mutation and in users of oral contraceptives. *N Engl J Med* 1998;338:1793–1797.

18. De Stefano V, Chiusolo P, Paciaroni K, et al. Prothrombin G20210A mutant genotype is a risk factor for cerebrovascular ischemic disease in young patients. *Blood* 1998;91:3562–3565.

19. Watzke HH, Schuttrumpf J, Graf S, et al. Increased prevalence of a polymorphism in the gene coding for human prothrombin in patients with coronary heart disease. *Thromb Res* 1997;87:521–526.

20. Ross R. Atherosclerosis—an inflammatory disease. *N Engl J Med* 1999;340:115–126.

21. Lalouschek W, Aull S, Serles W, et al. Genetic and nongenetic factors influencing plasma homocysteine levels in patients with ischemic cerebrovascular disease and in healthy control subjects. *J Clin Lab Med* 1999;133:575–582.

22. Tokgozoglu SL, Alikasifoglu M, Unsal I, et al. Methylene tetrahydrofolate reductase genotype and the risk and extent of coronary artery disease in a population with low plasma folate. *Heart* 1999;81:518–522.

23. McAndrew PE, Brandt JT, Pearl DK, et al. The incidence of the gene for thermolabile methylene tetrahydrofolate reductase in African Americans. *Thromb Res* 1996;83:195–198.

24. Alpert MA. Homocyst(e)ine, atherosclerosis, and thrombosis. *South Med J* 1999;92:858–865.

25. Griffin JH, Evatt B, Zimmerman TS, et al. Deficiency of protein C in congenital thrombotic disease. *J Clin Invest* 1981;68:1370–1373.

26. Egeberg O. Inherited antithrombin deficiency causing thrombophilia. *Thromb Diath Haemorrh* 1965;13:516–530.

27. Virchow R. Phlogose and thrombose in gefassystem. Gessammelte Abhandlungen zur Wissenschaftlichen Medecin. Frankfurt: Staatsdruckerie, 1856.

28. Comp PC, Nixon RR, Cooper MR, et al. Familial protein S deficiency is associated with recurrent thrombosis. *J Clin Invest* 1984;74:2082–2088.

Section 9 Microbiology

Microbiology - Introduction

Gail Woods

Section Chief

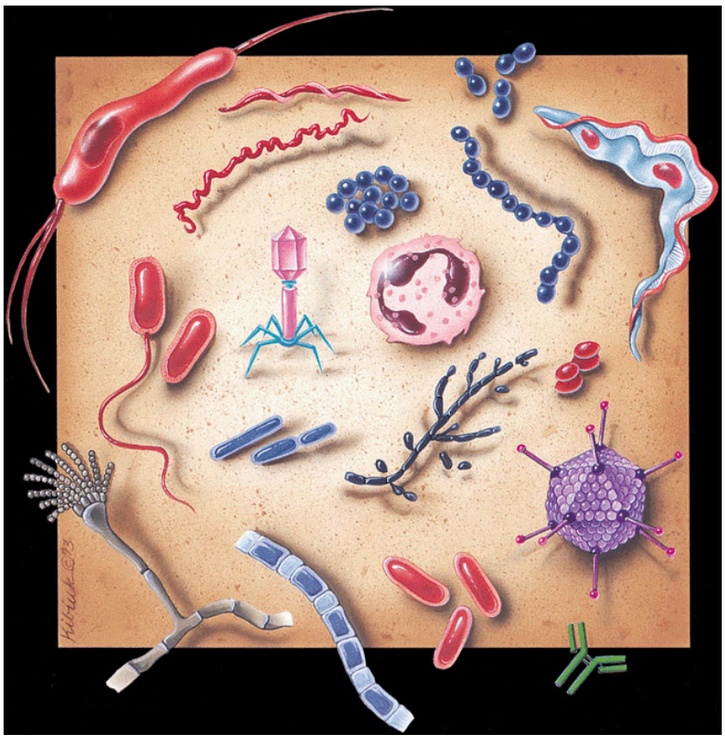


Figure.

During the past two decades, several major changes in the area of clinical microbiology and infectious diseases have occurred. Advances in medical technology and life-support systems have resulted in dramatic increase in the number of patients with profoundly compromised immune status and/or serious underlying diseases. These patients, as well as those who are victims of the AIDS epidemic, are at high risk for infections from both classic “true” pathogens and many opportunistic pathogens. At the same time, we have witnessed the development and application of new technology for the detection and identification of these agents of infection. Developments in immunology, molecular biology, and laboratory automation have both enhanced and complicated the job of the clinical microbiologist. These many changes and advances have placed new demands on the clinical microbiology laboratory. The laboratory must remain flexible enough to develop and offer expanded testing for new and unusual microorganisms as they become recognized as important pathogens in highly immunocompromised patients. The demand for rapid methods and decreased turnaround time has become particularly acute, underscoring the need for prompt therapeutic intervention in the infected immunocompromised patient. Finally, all these tasks must be accomplished in a high-quality, cost-effective manner as the laboratory comes under scrutiny by administrators and third-party payer and external accreditation agencies.

The increased cost of many diagnostic tests, combined with severe financial constraints, has caused many laboratories to send out their high-cost, low-volume tests to a reference laboratory. In other instances, simplification of some tests, such as viral cultures, has allowed laboratories to bring these tests back into the laboratory. Recent developments in immunology and molecular biology have resulted in an increased number of rapid direct-testing methods in which the quality of the specimen submitted is of utmost importance. In each case, these changes make sample collection, handling, and transportation important considerations.

50
Specimen Collection and Processing for Microbiology

Barbara S. Reisner

Gail L. Woods

Successful laboratory diagnosis of a microbial infection depends on many factors, beginning with a well-collected sample. Proper specimen selection, collection, and transport are all essential to ensure that a specimen is representative of the disease process and minimally contaminated with microorganisms present in adjacent tissues. The first part of this chapter focuses on general specimen collection and processing issues; the remainder contains specific recommendations for each specimen type.

One of the most important rules for specimen procurement is to collect the sample from the correct anatomic site; for example, a superficial sample of a lesion is not useful in identifying the cause of a deep wound infection. The timing of sample collection is also important; for example, when submitting a specimen for bacterial culture, samples should be collected before the administration of antibiotics.

Recommended specimen collection techniques may vary by institution; however, all laboratories should have written guidelines for proper specimen collection and these should be followed. In general, specimens other than stool are collected using sterile technique and equipment. Whenever possible, pieces of tissue or aspirated material rather than swab specimens should be obtained. Use of swabs result in the collection of very small volumes of specimen, specimens are more likely to dry out during transport to the laboratory, and specimens are difficult to remove for examination. Sufficient volume of specimen must be collected for the number of tests requested. If the volume is inadequate, the health care provider must prioritize the test requests.

After collection, the specimen must be placed in an appropriately labeled leak-proof container. Use of a specific type of container, containing special media or preservative, may be required by the laboratory. Containers are then placed in a biohazard bag for transport to the laboratory. Each specimen must be accompanied by a requisition slip containing all the information required to evaluate the specimen appropriately and relay the test results back to the health care provider without delay. This includes patient name, age, gender, identification number, location, name of health care provider, time and date of collection, specimen type, diagnosis, and test(s) requested. A mechanism must be in place to match a specimen with a requisition slip. Usually this is accomplished by using a number that is on both the requisition slip and a detachable label that is then placed on the specimen itself.

Transport of specimens to the laboratory should be rapid, optimally in less than 2 hours. For delays in transport, most specimens should be refrigerated; exceptions are blood, cerebrospinal fluid (CSF), and specimens to be examined for anaerobes, fastidious organisms such as *Neisseria gonorrhoeae* and *Bordetella pertussis*, and *Trichomonas vaginalis*, all of which should be maintained at room temperature.

On arrival at the laboratory, the specimen and requisition are examined to ensure that all criteria for acceptance are fulfilled. Minimum criteria for an acceptable specimen are listed in Table 50.1. In addition, some specimens may not be acceptable for particular microbiological procedures. For example, some specimens are not acceptable for anaerobic culture (Table 50.2) because the normal anaerobic flora contained in these specimens make it difficult to distinguish pathogens from nonpathogens. When specimens are rejected, the health care provider is notified so that another specimen may be properly submitted. If information on the requisition is incomplete, laboratory personnel should ask a responsible person to provide the information before processing the specimen further. If a specimen is mislabeled, the sample should be recollected. Relabeling of a specimen is acceptable only for difficult to collect specimens, such as tissue obtained during a surgical procedure or CSF.

TABLE 50.1. SPECIMEN REJECTION CRITERIA

Improper transport temperature
Improper transport container or medium
Prolonged transport time
Unlabeled or mislabeled specimen
Broken or cracked container
Leaking specimen
Dried-out specimen
Inappropriate specimen for test requested
Inadequate volume
Specimen in fixative (for culture)
Duplicate sample in 24-hr period (for urine, sputum, feces culture)

All specimens should be presumed to contain transmissible agents and therefore should be collected and handled using standard precautions. Standard precautions are equivalent to what had previously been termed universal precautions and are designed to reduce the risk of transmission of microorganisms from both recognized and unrecognized sources of infection. This includes the use of gloves, gown, mask, and protective eyewear when there is a risk of coming in contact with the specimen. In most clinical laboratories, a special area is designated for processing clinical samples for culture. Ideally all specimen processing, but at a minimum processing of respiratory secretions and specimens submitted for the detection of mycobacteria or fungi, should be performed in a class II biological safety cabinet.

Many specimen types submitted for detection of bacteria are examined directly using Gram stain; if mycobacteria are suspected, a stain for acid-fast bacilli (AFB) is also performed. Some specimens submitted for detection of fungi will be examined directly with a potassium hydroxide (KOH) and/or calcofluor

white preparation; however, for some specimen types, this is only performed on request. Stool specimens may be examined directly for parasites with a wet mount but more often are placed in a fixative and examined after a concentration procedure. Other techniques for directly examining specimens, such as direct fluorescent antibody stains (DFA), enzyme immunoassays (EIAs), and DNA hybridization or amplification assays are used only in selected situations.

TABLE 50.2. SPECIMENS NOT ROUTINELY ACCEPTED FOR ANAEROBIC CULTURE

Throat, nasopharyngeal, or gingival swabs
Sputa
Bronchial wash, lavage, or brush (except when collected with a protected double lumen catheter)
Gastric and bowel contents
Ileostomy and colostomy effluent
Voided or catheterized urine
Female genital tract specimens collected through the vagina
Surface swabs of ulcers, wounds, and abscesses

A combination of media types is used to isolate bacteria and fungi. These may include enriched, nonselective, selective, or differential media. Viruses are obligate intracellular parasites and can only be cultured within mammalian cells, of which there are three main categories: primary, low-passage finite, and continuous cell lines. With few exceptions, culture of parasites is generally not performed.

Most routine bacterial cultures are incubated for 2 to 3 days. Mycobacterial and fungal cultures are incubated for as long as 6 weeks. Viral cell cultures are incubated for varying lengths of time depending on the specimen source and the growth rate of the viruses that are typically recovered from that site (Table 50.3). The more rapid shell vial technique used for culture of some viruses requires from 24 to 72 hours of incubation. The temperature of incubation is usually 35°C for bacteria and viruses and 30°C for fungi. Various atmospheric conditions may be utilized including ambient, CO_2 enriched, microaerophilic, and anaerobic.

- BLOOD
- CEREBROSPINAL FLUID
- EAR
- EYE
- GASTROINTESTINAL TRACT
- GENITAL TRACT
- LOWER RESPIRATORY TRACT
- UPPER RESPIRATORY TRACT
- TISSUES
- URINE
- SKIN AND SUBCUTANEOUS LESIONS

BLOOD

Part of "50 - Specimen Collection and Processing for Microbiology"

Detection of organisms in blood is one of the most important functions of the clinical microbiology laboratory. Fortunately, there have been many improvements in the media and technology for culturing blood during the past decade, resulting in highly reliable manual and automated systems. To optimize these blood culture systems, important guidelines for collection must be followed. In general, blood for culture should not be obtained using an intravascular device. If a reason exists to culture blood using such a device, for example, to evaluate a possible line-related infection, blood obtained by peripheral venipuncture should be submitted at the same time.

When performing a venipuncture, the skin must be adequately disinfected to minimize contamination with normal skin flora. For maximum antisepsis, the skin is first cleansed with 70% isopropyl alcohol followed by application of an iodophor or iodine tincture that must remain in contact with the skin for 1.5 to 2 minutes. Blood should be collected and inoculated into the blood culture bottles using the same needle. Most routine blood cultures are collected as sets of one anaerobic and one aerobic bottle. In some institutions where the incidence of anaerobic bloodstream infection is extremely low, two aerobic blood culture bottles may be used. Special bottles are available for low-volume blood draws and for mycobacterial or fungal culture.

Blood specimens should be collected before administering antimicrobial agents, although media that contain substances designed to minimize the effect of such agents on bacterial growth are available. Optimally, the specimen should be collected just before a fever spike; however, practically, the specimen should be collected immediately after the spike.

The total volume of blood cultured is one of the most important factors in the recovery of the bacterial pathogen. In one study, increasing the volume of blood cultured over a 24-hour period from 40 to 60 mL increased recovery by 10% (2). For adults, 20 to 30 mL of blood should be collected per venipuncture. Less blood is required for children because there are more microorganisms per milliliter of blood. Most blood culture systems specify the volume of blood that should be placed in the blood culture bottle for optimal recovery. It is particularly important to adhere to these guidelines when using the Isolator system (Wampole Laboratories, Cranbury, NJ), which contains factors that inhibit growth of microorganisms if the minimum blood volume is not present.

For adult patients, two sets of cultures should be collected per

febrile episode to help distinguish probable pathogens from possible contaminants, especially when considering the significance of coagulase-negative staphylococci, corynebacteria, viridans streptococci, and *Bacillus* spp. In general, contaminants are more likely to be recovered from a single set of cultures, whereas pathogens typically are recovered from more than one set. No more than four sets should be submitted in a 24-hour period. This limit allows the evaluation of two febrile episodes per day with two blood cultures sets submitted per episode. Greater numbers of cultures are not necessary based on data that show that more than 95% of bacteremias are detected with the first two to three cultures (3). When culturing blood specifically for mycobacteria, two cultures per week are generally adequate.

TABLE 50.3. USUAL VIRAL GROWTH RATES

Virus	Growth Rate (days)
Adenovirus	2–10
Cytomegalovirus	5–28
Enterovirus	2–8
Herpes simplex	1–7
Influenzavirus	2–10
Mumps	5–10
Parainfluenza	4–10
Respiratory syncytial virus	3–10
Rhinovirus	4–10
Varicella-zoster	5–28

Adapted from Forbes BA, Sahn DF, Weissfeld AS, eds. *Bailey and Scott's* diagnostic microbiology, 10th ed. St. Louis; Mosby, 1998.

Inoculated blood culture vials should be held at room temperature until they reach the laboratory. Direct microscopic examination of blood specimens is not useful because this procedure is not sensitive enough to detect the small numbers of organisms generally present in the blood stream [less than 10 colony-forming units (CFU) per milliliter]. Blood cultures for rapidly growing bacteria and yeast are usually incubated for 5 to 7 days. Cultures for mycobacteria and slowly growing fungi are held for as long as 42 days. As soon as growth is detected from the blood specimen, a stain is performed [Gram, acid-fast, or Giemsa stain] to determine the type of microorganism present. Positive stain results are considered a critical value and called directly to the patient's health care provider.

Many types of blood culture systems are available, including both manual and automated. With conventional manual blood culture systems, blood is inoculated into a culture bottle containing a nutritionally enriched liquid medium. The bottles are examined daily for visual evidence of growth (gas production, lysis of red cells, turbidity) and are also subcultured to solid media. The manual biphasic Septi-Chek system (Becton Dickinson Microbiology Systems, Cockeysville, MD) consists of a bottle containing broth to which an agar paddle is attached. The agar is inoculated by inverting the bottle until the broth comes in contact with the agar medium. The agar and the broth are then visually inspected for evidence of growth. The Isolator lysis centrifugation system is another manual blood culture method that consists of a tube containing the anticoagulant ethylenediaminetetraacetic acid (EDTA) and saponin (a lysing agent). After the tube is inoculated with blood, it is centrifuged and the resulting pellet is plated onto agar media. This system works exceptionally well for the recovery of mycobacteria and fungi, especially *Histoplasma capsulatum*.

The fully automated, continuously monitoring blood culture systems are the newest type of systems developed for the detection of bacteria and fungi in blood. The systems available in the United States are the BacT/Alert (Organon Teknika Corporation, Durham, NC), Bactec 9000 (Becton Dickinson Microbiology Systems), and ESP (Trek Diagnostic Systems Inc., Westlake, OH). The systems are alike in that the culture bottles are incubated in an instrument, where they are continuously monitored (typically at 10-minute intervals) for the production and/or consumption of gas. The data collected are transmitted to a computer and analyzed to allow rapid detection of microbial growth. Each system utilizes a noninvasive method (i.e., colorimetric, fluorescent, or manometric methods for detecting CO_2 or other gases) to monitor growth.

Various bacteria capable of causing bloodstream infections may not be recovered using routine blood culture procedures. Brucellae grow in routine blood culture media in as few as 3 days; however, as much as 21 days of incubation may be required and, in some cases, a terminal subculture is necessary. The HACEK group of bacteria (i.e., *Haemophilus aphrophilus*, *Actinobacillus actinomycetemcomitans*, *Cardiobacterium hominis*, *Eikenella corrodens*, and *Kingella kingae*), associated with bacterial endocarditis, also are recovered from routine media; however, as much as 14 days of incubation and terminal subculture of the broth may be useful. *Leptospira interrogans* requires special media such as Fletcher's, Ellinghausen-McCullough/Johnson-Harris, or Polysorbate 80, and cultures should be incubated for as long as 6 weeks.

Recovery of *Bartonella henselae* from blood is difficult. With the Isolator system, processed blood is plated to freshly prepared enriched blood or chocolate blood agar and incubated at 35°C under conditions of elevated CO_2 and humidity for as long as 4 weeks. Using the Septi-Chek system, growth may be achieved after incubation for 40 days. In broth-only systems, the organisms may be observed by staining a smear of the broth with acridine orange; however, the organism does not produce turbidity or enough CO_2 to be detected using instrumentation.

Culture of catheter tips may be performed to determine the source of a bacteremia. One commonly used technique utilizes the semiquantitative catheter tip culture method (4), in which the segment is rolled across a blood agar plate four times. Cultures yielding organisms present in more than 15 CFU are considered to be significant, potentially indicating a catheter-related infection.

A limited number of viruses are recovered from blood specimens. In most diagnostic virology laboratories, cytomegalovirus (CMV) and the enteroviruses are the most common isolates. Specialized reference laboratories are required to isolate human immunodeficiency virus (HIV), parvovirus B19, and the arthropod-borne viruses. For routine viral culture, 5 to 10 mL of blood should be collected in a sterile tube with an anticoagulant such as citrate; however, enteroviruses and arboviruses may also be recovered from serum.

Processing blood to obtain serum is accomplished by allowing the blood to completely clot at 4°C followed by centrifugation. Processing blood to obtain the leukocyte fraction can be accomplished by harvesting the buffy coat fraction obtained after a low-speed centrifugation; however, more purified preparations are obtained using density-gradient centrifugation with one of several commercially available products.

Direct microscopic examination of peripheral blood neutrophils for CMV can be accomplished using an antigenemia assay, a quantitative DFA that utilizes a monoclonal antibody against the pp65 antigen. This assay is most useful in the management of CMV infection in solid-organ allograft recipients. An alternative direct assay for CMV is the DNA hybrid capture assay (Digene, Silver Spring, MD).

For traditional viral culture, specimens are added to a monolayer of mammalian cells, culture medium is added, and the cells are incubated for as long as 28 days. The cell lines are examined throughout the culture period for viral cytopathic effects. When cytopathic effects are observed, the virus present is determined

with additional tests, e.g., staining the monolayer with a virus-specific antibody. For some viruses, especially CMV, the shell vial technique in which the specimen is centrifuged onto a monolayer of cells growing on a coverslip enhances recovery and allows more rapid reporting of results. With this technique, the coverslips are removed and stained with a monoclonal antibody early in the incubation period (24 to 48 hours).

Several parasites may be found in the bloodstream at some time during their life cycle. These include *Plasmodia* spp, *Trypanosoma* spp, *Babesia* spp, *Leishmania donovani*, and the microfilaria. For some of these organisms (trypanosomes and filaria), examination of fresh whole blood is useful to observe motility; however, in most cases, thick- and thin-stained blood films are required to determine the species of organism present. Blood films of whole, anticoagulated, or various concentrated fractions of blood may be prepared. Buffy coat fractions are useful when looking for *Leishmania* or *Trypanasoma* spp. The time of blood collection is important when evaluating a patient for parasitemia (Table 50.4). Giemsa, Wright's or Wright-Giemsa stain may be used to stain the blood film. Giemsa is the best stain for determining the species of plasmodia. Blood films are first examined under low power to visualize microfilaria and then under progressively higher power to detect other parasites. A thick smear should always be examined along with the thin smear to detect light parasitemia. When malaria is suspected, several smears should be submitted each day for 3 days.

TABLE 50.4. OPTIMAL TIME TO COLLECT BLOOD SAMPLE FOR PARASITE EXAMINATION

Parasite	Collection Time
Babesia microti	Any time
Brugia malayi	Midnight
Leishmania donovani	Any time
Loa loa	Noon
Mansonella ozzardi	Any time
Mansonella perstans	Night
Plasmodium spp	Between chills
Trypanosoma spp	Acute stage
Wuchereria bancrofti	Midnight

Adapted from Miller JM, Holmes HT. Specimen collection, transport, and storage. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*. Washington, DC: ASM Press, 1999:4.

CEREBROSPINAL FLUID

Part of "50 - Specimen Collection and Processing for Microbiology"

CSF is submitted for microbiological analysis when meningitis or encephalitis is suspected. For meningitis, the likely infectious agent differs depending on the duration of symptoms (Table 50.5). The most likely bacterial agent of acute meningitis will also vary with the age of the individual and whether the disease is community or nosocomially acquired. Most infectious cases of encephalitis are a result of viral infection, both arthropod and nonarthropod borne. Parasitic infections of the central nervous system also occur, with varying clinical presentations.

CSF for microbiological examination is usually obtained by lumbar spinal puncture. The specimen should be collected before antibiotic treatment. Each type of procedure (smear, culture, antigen test) generally requires at least 0.5 mL of CSF with the notable exception of the mycobacterial culture, which requires at least 3 mL (greater volumes increase recovery). Because of this, the largest aliquot of specimen should always be sent to the microbiology laboratories, with lesser volumes sent to the chemistry and hematology laboratory. The specimen should be transported to the laboratory promptly and processed as soon as possible. If a delay in processing is unavoidable, the specimen should be held at room temperature.

If greater than 1.0 mL of CSF is received for a given test, the fluid is centrifuged to allow the test to be performed on the concentrated sediment. A cytocentrifuge may be used to prepare a concentrated specimen for direct microscopic evaluation.

A Gram stain should be performed on all CSF specimens immediately on receipt in the laboratory. Rapid tests for bacterial antigens of group B streptococci, *Streptococcus pneumoniae*, some serotypes of *Neisseria meningitidis*, *Escherichia coli* (K1 capsular antigen cross-reacts with that of *N. meningitidis* type B), and *Haemophilus influenzae* type b are available; however, data from several studies indicate that the sensitivity of these tests is less than or equal to that of Gram stain and the results of these tests have little impact on the care of the patient (6,7 and 8). The exception is partially treated meningitis, in which the antigen tests may be more sensitive. For these reasons, many laboratories no longer offer these expensive and labor-intensive tests or strictly limit their use.

For bacterial culture, an aliquot of the specimen is plated onto solid media and incubated for 3 days. Broth cultures of CSF and other sterile body fluids are no longer performed in many laboratories because studies that have shown that most organisms recovered in the broth only are not clinically significant, potentially leading to use of unnecessary antimicrobial agents (9,10). In these studies, broth cultures were found to be useful only for culture of CSF collected for the evaluation of a shunt infection and chronic ambulatory peritoneal dialysis fluid, both examples of clinical conditions in which low numbers of organism may be present. Anaerobic culture is not routinely performed on CSF specimens because anaerobes very rarely cause meningitis.

Leptospires can be seen in CSF during the first 10 days of illness using dark-field microscopy of a concentrated specimen. Leptospires can be recovered in culture by inoculating a special

medium that is incubated in the dark at room temperature for as long as 6 weeks.

TABLE 50.5. PROBABLE INFECTIOUS CAUSE OF MENINGITIS

Duration of Symptoms	Probable Pathogen
<24 hr	Pyogenic bacteria
1–7 days	Enteroviruses
	Pyogenic bacteria
≥4 wk	*Mycobacterium tuberculosis*
	Treponema pallidum
	Brucella spp
	Leptospira interrogans
	Borrelia burgdorferi
	Cryptococcus neoformans
	Coccidioides immitis
	Histoplasma capsulatum
	Candida spp

T. pallidum cannot be grown in routine culture; therefore, the diagnosis must be made based on the clinical presentation and abnormal CSF parameters: pleocytosis, elevated protein, and/or a positive VDRL (Venereal Disease Research Laboratory) test. The VDRL is the only useful test for detecting syphilis antibodies in CSF. Because a positive CSF VDRL is usually associated with a positive serum rapid plasma reagin or VDRL, the serum status should be determined before performing the CSF test.

B. burgdorferi also cannot be grown in routine culture; therefore, the clinical presentation, history, and serologic testing for organism-specific immunoglobulin (Ig) G and IgM are used for diagnosis.

With the exception of specimens from patients with acquired immunodeficiency syndrome, processing CSF for mycobacteria is indicated only for samples with pleocytosis or abnormal glucose or protein values. Because the number of mycobacteria present in CSF is low, a large volume of specimen (3 to 5 mL optimal) is necessary to maximize recovery of the organism in culture. Organisms can be seen directly in the specimen using a special stain for AFB (Kinyoun, Ziehl-Neelson, or auramine-rhodamine). Specimens are plated to both solid and liquid medium for culture.

Most yeast and fungi, with the possible exception of *H. capsulatum*, will be detected in the Gram stain. A special stain such as Wright's or Giemsa may be performed to detect *H. capsulatum*. A direct antigen test for the capsular antigen of *Cryptococcus neoformans* is available and is the most sensitive test for diagnosing cryptococcal meningitis. The India ink test is no longer recommended for this purpose because it lacks sensitivity and specificity (white cells present in CSF may be misidentified as encapsulated yeast). For culture, the specimen is inoculated to solid medium only.

Viral culture of the CSF is generally useful only for recovery of enteroviruses. Culture for other viruses, such as HIV and arboviruses, are only performed by specialized laboratories. For the arboviruses, diagnosis requires serologic testing for viral antibodies in most cases. Use of a nucleic acid amplification test is the most reliable way to diagnose herpes simplex virus (HSV) encephalitis. Amplification tests are also available for the enteroviruses and in the future may become part of the routine testing in cases of possible enteroviral meningitis.

A wet mount examination and the Giemsa stain are used to examine CSF for free-living amebae (*Naeglaria* spp, *Acanthamoeba* spp, *Balamuthia mandrillaris*). Culture of *Naeglaria* and *Acanthamoeba*, using nonnutrient agar with a bacterial overlay, may also be available in some laboratories. For neurocysticercosis, detection of antibodies against *Taenia solium* in the CSF, along with clinical history and characteristic findings on head computed tomography scan, is diagnostic.

Body Fluids Other Than CSF

Specimens in this category include pericardial, pleural, peritoneal, and synovial fluids. For optimal microbiological analysis, it is essential that fluid, rather than a fluid-saturated swab, be submitted. A volume of 1 to 5 mL is adequate for bacterial culture; however, larger volumes (10 to 15 mL) should be submitted for fungal or mycobacterial culture. For diagnosing peritonitis associated with chronic ambulatory peritoneal dialysis, collection of at least 50 mL of fluid may improve recovery of the responsible pathogen.

Body fluids should be submitted in a sterile container. If the syringe used to collect the specimen is submitted to the laboratory, the needle must first be removed. Transferring the specimen to a tube containing an anticoagulating agent is not recommended for bacterial culture because heparin, sodium citrate, and EDTA have all been shown to inhibit particular bacterial species (11,12). For culture of bacteria other than mycobacteria, a sample may be directly inoculated into a blood culture bottle at the bedside. For viral culture, the specimen should be placed in viral transport medium and placed on ice. Body fluids should be transported promptly to the laboratory.

Direct examination of the specimen by Gram stain should always be performed. If more than 1 mL of specimen is received, it should be concentrated by centrifugation first. An acid-fast stain or calcofluor white/KOH preparation of the sediment from a concentrated specimen should be performed if mycobacteria or fungi are suspected. Solid media used for culture of body fluid specimens should be selected to allow recovery of fastidious organisms. As previously mentioned with regard to CSF, inoculating a liquid medium for routine bacterial culture of body fluid specimens may be necessary only when culturing dialysate from patients undergoing chronic ambulatory peritoneal dialysis (9,10). Both liquid and solid culture media should be inoculated for mycobacterial culture.

Pericarditis owing to Coxsackie A and B virus may be diagnosed by viral culture of the pericardial fluid; however, virus may also be recovered from throat and rectal specimens. Parasites may rarely be found in body fluids and may be seen in either a wet mount or Giemsa-stained smear. These include microfilariae, *Strongyloides stercoralis* in cases of hyperinfection, *Entamoeba histolytica* as a result of a ruptured liver abscess, or scoleces and hooklets (hydadid sand) of *Echinococcus granulosis* after rupture of a hydatid cyst.

EAR

Part of "50 - Specimen Collection and Processing for Microbiology"

Types of ear specimens include swab specimens for the diagnosis of otitis externa and middle ear fluid obtained by tympanocentesis for the diagnosis of otitis media. The potential pathogens at these two sites differ.

The causes of an infection of the external ear are similar to those of other infections of skin and soft tissue. Localized infections may be a result of *Staphylococcus aureus* and group A streptococci. Diffuse infections are often owing to gram-negative bacilli, especially *Pseudomonas aeruginosa*. Chronic infections may occasionally be a result of tuberculosis, syphilis, or leprosy. Fungal infections may be a result of *Aspergillus* spp and *Candida albicans*.

Otitis media is most often caused by organisms derived from the respiratory flora, including *S. pneumoniae*, *H. influenzae*, and *Moraxella catarrhalis*, as well as *S. aureus* and gram-negative bacilli. Anaerobic bacteria can be involved in middle ear infections;

therefore, anaerobic culture may be useful. Respiratory syncytial virus (RSV), influenza viruses, enteroviruses, and rhinoviruses have also been associated with otitis media.

Direct examination of a smear of the middle ear fluid using Gram stain may be performed if an adequate volume is submitted. Direct examination of other types of ear specimens is not usually done. Specimens from the inner ear should be plated to media capable of supporting the growth of fastidious bacteria. *Candida* spp will grow on routine media used for bacterial culture, but special fungal media should be used to recover molds. Viral culture generally is not performed because antiviral therapy typically is not administered.

EYE

Part of "50 - Specimen Collection and Processing for Microbiology"

Several types of specimens may be submitted for the microbiological analysis of eye infections, including conjunctival specimens collected with a swab or a sterile spatula for the diagnosis of conjunctivitis, corneal scrapings collected with a sterile spatula for the diagnosis of keratitis, vitreous fluid for the diagnosis of endophthalmitis, and purulent material for the diagnosis of cellulitis. If a swab is used, it should be dipped in sterile broth media before collecting the specimen. Multiple samples should always be obtained, one for each stain and media type. Because it is difficult to transport the small volumes of specimen collected, samples are applied directly to glass slides and growth media at the collection site. Smears should be fixed immediately after preparation.

Special procedures are required when collecting specimens for *Chlamydia trachomatis* and viral culture. Using swabs with a wooden shaft should be avoided because the wood may be toxic to the organisms, and specimens should be placed directly in transport medium containing antimicrobial agents (usually gentamicin, vancomycin, and nystatin or amphotericin B; penicillin should not be used for *C. trachomatis*) to inhibit overgrowth of bacteria and fungi. Specimens should be placed on ice for transport to the laboratory.

Direct examination of smears is important for rapid diagnosis of eye infections and should always include a Gram stain for detection of bacteria. To detect *C. trachomatis* directly in a conjunctival specimen, a Giemsa stain of conjunctival scrapings may be examined for the presence of epithelial cells containing basophilic intracytoplasmic inclusions. However, monoclonal antibodies for use in direct immunofluorescent assay are preferred because they are more sensitive and specific than Giemsa stain. For optimal results with the direct fluorescent antibody tests, the collection kit provided by the manufacturer should be used, and the specimen should contain at least 10 columnar or metaplastic squamous cells. A commercial EIA (Clamydiazyme, Abbott Laboratories, Abbott Park, IL) and a nucleic acid probe assay (Pace, Gen-Probe Inc., San Diego, CA) are also available for detection of *C. trachomatis* in conjunctival specimens.

An AFB stain should be performed if mycobacteria are suspected. Calcofluor white and/or a KOH preparation should be examined if fungi are suspected. In fungal keratitis, fungi may not be seen unless tissue from deep in the corneal parenchyma is examined. Rapid detection of adenovirus, HSV, and varicella zoster virus (VZV) can be accomplished using fluorescent immunostains. Stains used to detect parasites include Giemsa and hematoxylin and eosin (H&E), which are useful for detecting *Acanthamoeba* spp and microfilaria, and modified trichrome stain to detect microsporidia.

Whether inoculated at the bedside or in the laboratory, the media for routine bacterial culture of eye specimens should include an enriched medium for the recovery of fastidious organisms (e.g., *S. pneumoniae*, *N. gonorrhoeae*, and *H. influenzae*). Cell culture is required to recover *C. trachomatis*. The rapidly growing mycobacteria may be recovered on media used for routine bacterial culture; however, the laboratory should be notified so that cultures can be incubated for additional time to ensure recovery. If mycobacterial culture is specifically requested, special solid and liquid media will be inoculated. *Candida* spp will be recovered using procedures for routine bacterial culture; however, recovery of molds requires fungal media and incubation conditions. Cell culture is required to recover viruses. The more rapid technique of shell vial culture is often used to detect CMV. *Acanthamoeba* are one of the only parasites for which culture is attempted (using nonnutrient agar with a bacterial overlay), but the procedure is not available in all laboratories.

GASTROINTESTINAL TRACT

Part of "50 - Specimen Collection and Processing for Microbiology"

Feces, and in some cases rectal swabs, are submitted to the laboratory primarily to determine the etiologic agent of infectious diarrhea or food poisoning. With the availability of a new EIA for *Helicobacter pylori* antigen (Meridian Diagnostics Inc., Cincinnati, OH), submission of stool to determine the etiologic agent of gastritis and peptic ulcers may become common practice in the future. Feces should be collected in a clean container with a tight lid and should not be contaminated with urine, barium, or toilet paper. Feces should optimally be examined within 2 hours of collection. If this is not possible, a portion should be placed in Stuart's or Amies transport medium for bacterial culture or in preservative for parasitology examination. Rectal swabs should be placed in a tube transport system containing modified Stuart's medium. Unpreserved stool specimens should be maintained at refrigerator temperature during storage and transport to the laboratory.

It is becoming standard practice to reject stool specimens for bacterial culture and parasite examination from patients who have been hospitalized longer than 3 days because diarrhea that develops in the hospital is not likely to be caused by food- or water-borne pathogens. For such patients, examination for the toxins produced by *Clostridium difficile* is recommended.

For bacterial culture, the media inoculated should allow the detection of *Salmonella*, *Shigella*, and *Campylobacter* spp at a minimum. This usually consists of a selective and differential medium such as MacConkey agar, a more selective medium such as Hektoen enteric or xylose-lysine-desoxycholate agar, and a medium for growth of *Campylobacter* spp, such as *Campylobacter* agar with 10% sheep blood. Media for the recovery of *Campylobacter* spp are incubated in a microaerophilic environment at 42°C for as long as 3 days.

The prevalence of gastroenteritis caused by *Aeromonas* spp,

Plesiomonas spp, *Vibrio* spp, and *Yersinia enterocolitica* is sufficiently low in most countries that cultures for these organisms generally are performed on request only. Media routinely used to culture for enteric pathogens are sufficient to recover *Aeromonas* spp and *Plesiomonas* spp. *Vibrio* spp and *Y. enterocolitica* will also grow on routine stool culture media, but the selective medium thiosulfate citrate bile salts sucrose agar should be inoculated for optimal recovery of *Vibrio* spp, and cefsulodin-irgasan-novobiocin agar should be inoculated and incubated at room temperature for the optimal recovery of *Y. enterocolitica*.

The prevalence of enterohemorrhagic *E. coli*-associated enterocolitis varies in different parts of the United States, and in many areas of the country, it is so low that laboratories look for it on request only. However, in 1993, the Council of State and Territorial Epidemiologists recommended that bloody stools be routinely cultured for this organism. To detect enterohemorrhagic *E. coli*, the stool specimen is inoculated onto sorbitol-MacConkey agar, a medium that differentiates isolates of enterohemorrhagic *E. coli*, which do not ferment sorbitol, from almost all other *E. coli*, which are sorbitol fermentation positive. In addition, stool filtrates can be tested for toxin production in Vero cells or by a commercial EIA.

Diseases associated with *C. difficile*, such as pseudomomembranous colitis and antibiotic-associated diarrhea, are caused by the toxins produced by the organism and are diagnosed by detecting toxin in feces. The reference method for detection of the cytotoxin is cell culture assay. To extract toxin, the stool specimen is clarified by centrifugation and filtered through a 0.45-µm membrane filter. Serial dilutions of the filtrate are prepared and inoculated to monolayers of human diploid lung fibroblasts. Alternatively, toxin may be detected in stool by one of many commercial EIAs, which are less sensitive than cell culture but provide results within a few hours. A latex agglutination test, which detects glutamate dehydrogenase produced by the organism, is available; however, this enzyme is not found exclusively in toxin-producing isolates of *C. difficile*, and cross-reactivity with other bacterial species occurs.

For epidemiologic purposes, stool or rectal swabs placed in an anaerobic transport system may be cultured anaerobically to isolate *C. difficile*. The sample is inoculated to a selective medium such as cycloserine-cefoxitin-fructose agar and incubated anaerobically for 48 hours. Isolates identified as *C. difficile* must be tested for toxin production to determine the association of the organism with disease

Stool specimens or gastric contents collected from persons with short-incubation food poisoning should be evaluated for *S. aureus* and *Bacillus cereus*. Specimens should be examined by Gram stain, and, because both of these organisms may be present normally in food, quantitative cultures are performed. A series of dilutions of the specimen is prepared, and 0.1 mL of the undiluted specimen and each of the dilutions is planted onto colistin nalidixic–acid or phenylethyl alcohol blood agar. The presence of 10^5 CFU or more of *S. aureus* or *B. cereus* per gram of specimen is of potential significance.

The clinical diagnoses of food-borne botulism and infant botulism may be confirmed by detecting botulinal toxin, *Clostridium botulinum*, or both in feces. Optimally, 25 to 50 mL of stool, 15 to 20 mL of serum, and a sample of the suspect food should be collected. Most clinical laboratories are not properly equipped to process specimens from persons with suspected botulism. In the United States, when a case of botulism is suspected, investigators at the Centers for Disease Control and Prevention (CDC) should be notified to assure appropriate diagnosis, treatment, and investigation of the potential outbreak.

With regard to mycobacterial culture, stool specimens usually are submitted for isolation of *Mycobacterium avium* complex (primarily in patients with acquired immunodeficiency syndrome). Processing the specimen involves decontamination, concentration, preparation of smears, and inoculation of appropriate media. Gastric aspirates, representing swallowed sputum, are occasionally submitted for mycobacterial culture from infants, young children, and persons who are obtunded. These should be obtained in the early morning before the patient eats and collected in a container containing sodium carbonate to neutralize the pH.

Fungal culture of stool is not recommended. As many as 40% of healthy individuals, and 75% of immunocompromised individuals may be colonized with yeast. Biopsy of tissue should be performed to diagnose definitively fungal infections of the gastrointestinal tract.

Viral culture of stool or rectal swabs may be used to determine the cause of many types of viral illness (e.g., fever, rash, gastroenteritis) in which the infectious agent is shed for weeks after an illness. Direct immunoassays are available to detect rotavirus in stool or in material collected on a rectal swab. Electron microscopy is also used to detect rotavirus, adenovirus, calicivirus, astrovirus, or Norwalk-like viruses in stool.

To detect parasites, stool is examined microscopically for the presence of protozoa, helminth eggs, and larvae. Stool from patients who have received medications such as barium sulfate, mineral oil, bismuth, some antidiarrheal preparations, antimalarials, and tetracycline should be rejected because they may obscure the parasites. A single stool specimen should be submitted initially, and, if negative, collection of at least two additional stool specimens is recommended. Fresh stool must be examined immediately after collection to visualize protozoa. If this is not possible, stool should be placed in a preservative for transport. Generally, a portion of the specimen is placed into each of two vials, one containing formalin and one containing polyvinyl alcohol; however, one-vial systems and systems utilizing other types preservatives may also be used.

On receipt in the laboratory, the specimen is treated to remove fecal debris and concentrated by centrifugation. If more than one specimen is received from a patient, the specimens may be pooled into one centrifuge tube. A wet mount with saline or iodine is used to examine the sediment, and a permanent stain (i.e., trichrome or iron hematoxylin) of either concentrated or unconcentrated specimen is prepared. Although many organisms are observed first in the wet mount preparation, a definitive identification often requires examination of the stained material. Many parasites observed in feces are not pathogenic; however, their presence may indicate that the patient has been exposed to contaminated food or water or that a pathogenic parasite is present at an undetectable concentration.

Special stains must be performed to detect some parasites. Microsporidia are detected with the modified trichrome stain or

calcofluor white. *Cryptosporidium parvum* and *Cyclospora cayatenensis* are best observed with the modified acid-fast stain. A DFA for *Giardia lamblia* and *C. parvum* is available, as are EIAs for *G. lamblia*, *C. parvum*, and *Entamoeba histolytica*.

The Scotch tape preparation, in which a piece of Scotch tape is applied to the skin across the anal opening and applied to a glass slide for microscopic evaluation is the optimal method for detecting *Enterobius vermicularis* (pinworm).

GENITAL TRACT

Part of "50 - Specimen Collection and Processing for Microbiology"

Genital tract specimens are sent to the laboratory for determining the cause of various clinical syndromes, including vulvovaginitis, bacterial vaginosis, genital ulcers, urethritis, cervicitis, endometritis, salpingitis, and ovarian abscess in females, and urethritis, epididymitis, prostatitis, and genital ulcers in males. Many specimens will be contaminated with the normal microbiota of the genital tract or skin; therefore, the microbiologist must differentiate the normal flora from potential pathogens. Organisms such as *N. gonorrhoeae*, *C. trachomatis*, and *Haemophilus ducreyi* are always pathogenic, whereas organisms such as the Enterobacteriaceae, *S. aureus*, and group B streptococci are pathogenic only in some clinical situations.

A direct Gram stain should be performed on any aspirated material. Gram stains of other specimens are useful in only a few situations. For example, in urethral discharge from males, the presence of gram-negative diplococci within polymorphonuclear leukocytes is sufficient for the presumptive diagnosis of gonorrhea. A decreased ratio of lactobacilli to *Gardnerella* spp, *Bacteroides* spp, and curved, gram-variable bacilli in a Gram-stained smear of a vaginal specimen can be useful for the diagnosis of vaginosis.

A wet mount preparation of vaginal secretions is a rapid and simple method for determining the etiologic agent of bacterial vaginosis as well as vulvovaginitis. For optimal results, the wet mount should be performed and examined within 1 to 2 hours of specimen collection. To prepare the wet mount, a sample of the discharge is mixed with saline, a portion of the mixture is placed on a glass slide, covered with a coverslip, and examined under low and high power for the presence of clue cells (epithelial cells covered with small coccobacillary bacteria), consistent with the diagnosis of vaginosis, pseudohyphae, suggestive of vaginal candidiasis, and motile trichomonads. Two other diagnostic tests performed in conjunction with the wet mount are vaginal pH and the “whiff test.” The vaginal pH is usually 4.5 or less in the absence of infection and in women with vulvovaginal candidiasis but more than 4.5 in those with bacterial vaginosis or trichomoniasis. A positive whiff test (i.e., generation of a pungent, fishy odor on addition of 10% KOH to the specimen) is associated predominantly with bacterial vaginosis but occasionally occurs with trichomoniasis.

Direct examination by dark-field or DFA (reagents are available from the CDC) is required for the detection of *Treponema pallidum* in tissues and tissue exudates. For dark-field microscopy, the specimen should be examined within 20 minutes of collection to ensure retention of motility. The test requires a microscope with a dark-field condenser and well-trained and experienced personnel who are able to recognize *T. pallidum* spirochetes based on the tightness and regularity of the spirals and its characteristic corkscrew movement. DFA, conversely, is performed on air-dried smears of tissues, body fluids, secretions, and lesion exudates, and is specific for *T. pallidum*, making it easier to interpret.

Both EIA and DFA tests are commercially available for the direct detection of *C. trachomatis* in genital specimens. If these kits are used, the specimen must be collected using the procedures and collection kits recommended by the manufacturer.

Nucleic acid probe and amplification methods for direct detection of *N. gonorrhoeae* and *C. trachomatis* in endocervical and urethral specimens, and a combination probe assay (Affirm VPIII microbial identification test, Becton Dickinson Microbiology Systems), for the detection of *Candida* spp, *Gardnerella vaginalis*, and *T. vaginalis* in vaginal secretions, are commercially available. As with the immunoassays, the specimen must be collected using the procedures and collection kits recommended by the manufacturer.

HSV may be detected directly in a smear of a specimen collected from the base of a herpetic lesion. For the traditional Tzanck smear, the material is stained with Wright or Giemsa stain and examined for the presence of characteristic multinucleated giant cells with molded nuclei and homogeneously staining “ground glass” nuclear chromatin (Fig. 50.1). The sensitivity of this approach ranges from 23% to 70%. Alternatively, a DFA may be used. The sensitivity of the DFA varies with the stage of the lesion, ranging from 10% to 87%. Several EIAs are commercially available. The sensitivity of these assays varies, but they generally perform very well in symptomatic populations. *Molluscum contagiosum* is detected by direct examination of specimens using histopathologic techniques. Human papillomavirus is detected by histology or with the DNA hybrid capture assay (Digene).

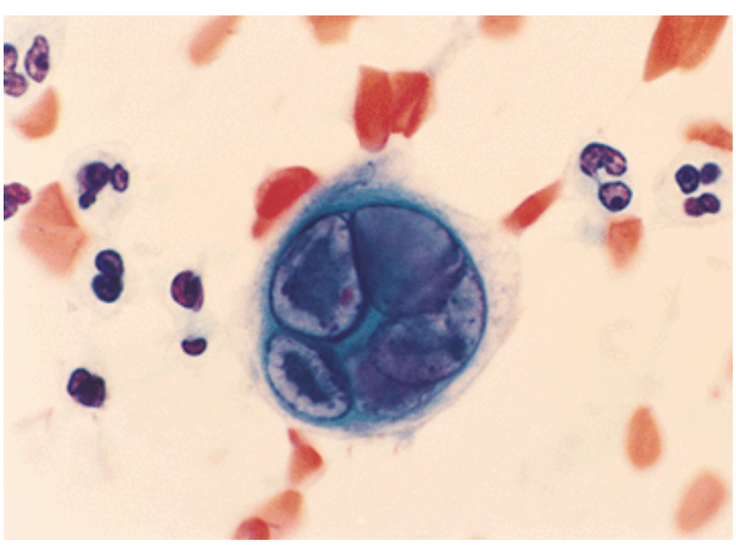


FIGURE 50.1. A multinucleate squamous cell demonstrates typical herpes simplex virus cytopathic effect (Papanicolaou stain, original magnification ×250). (Courtesy of Vicki J. Schnadig, M.D., Department of Pathology, University of Texas Medical Branch, Galveston, TX.)

The optimal media and incubation conditions for bacterial culture of genital specimens depend on the source and the organisms likely to cause disease at that site. Tissue and aspirates

should be plated to media capable of recovering fastidious organisms. Specimens from the cervix, vagina, and urethra should at a minimum be evaluated for *N. gonorrhoeae* and *C. trachomatis* by culture or a direct detection method. Culture is currently the only acceptable diagnostic procedure in a medicolegal case. The optimal approach for isolating *N. gonorrhoeae* is direct inoculation of a selective agar medium, such as modified Thayer-Martin, at the patient's bedside, with transport to the laboratory in a CO_2-containing environment. Alternatively, the swab specimen (cotton swabs should be avoided because they may be toxic) can be placed in a transport system containing Stuart's or Amies medium and delivered to the laboratory within 24 hours. If a delay in transport cannot be avoided, the swab should be left at room temperature, never refrigerated. For recovery of *C. trachomatis*, the specimen should be placed in liquid transport medium and transported to the laboratory without delay, where it is inoculated to cells using the shell vial technique.

It is important to detect group B streptococci in vaginal and perineal specimens from obstetric patients because this organism may cause sepsis and/or meningitis in neonates. For optimal recovery of group B streptococci, the procedure outlined by the CDC (13) should be followed: a swab specimen of the vaginal introitus and anorectum is inoculated into Lim broth, an enrichment broth for streptococci, which is incubated at 35°C and subcultured to a blood agar plate after overnight incubation.

If infection with *H. ducreyi* is suspected, material from the base of the ulcer is collected and held at room temperature until processed. One swab is used to prepare a smear for Gram staining. The presence of many small pleomorphic gram-negative bacilli and coccobacilli arranged in chains and groups suggests *H. ducreyi*. Recovery of the organism on an enriched medium such as chocolate agar supplemented with IsoVitaleX (Becton Dickinson Microbiology Systems) is necessary to confirm the diagnosis.

Actinomyces spp may cause pelvic inflammatory disease in women who use an intrauterine contraceptive device (IUD). An IUD submitted for culture should be placed in a sterile liquid medium, vortexed, and the liquid used to inoculate the culture medium. Cultures for *Actinomyces* spp should be incubated anaerobically for 14 days.

Fungal culture of female genital tract specimens is not productive. *Candida* spp are part of the normal flora in the genital tract of as many as 20% of healthy women; however, culture may be a useful tool after treatment of candidiasis.

Viral culture remains the gold standard for detection of HSV. The shell vial culture technique, in which the monolayer is stained after 16 to 48 hours with a monoclonal antibody, yields a more rapid result; however, traditional cell culture remains the most sensitive technique. Adenovirus may also be isolated by cell culture.

T. vaginalis is most often detected by examination of a wet mount examination of vaginal or urethral discharge, prostatic secretions, or urine sediment. However, culture may also be performed by inoculating the specimen at the bedside into commercially available pouches containing a medium designed to support the growth of trichomonads. The inoculated pouch must remain at room temperature during transport to the laboratory because cold temperatures will cause the organisms to lose viability. On receipt in the laboratory, the medium is examined immediately for motile trichomonads and held for an additional 7 days at 35°C, before issuing a final report.

LOWER RESPIRATORY TRACT

Part of "50 - Specimen Collection and Processing for Microbiology"

Specimens from the lower respiratory tract are submitted primarily to determine the etiologic agent of pneumonia. The most common respiratory specimen received in the laboratory is sputum, expectorated or induced. Other types of specimens are tracheal aspirates, transtracheal aspirates, bronchial washes, bronchial brushings, and bronchoalveolar lavage fluids

Lower respiratory tract specimens should be delivered promptly to the laboratory, but if delays in transport or processing are unavoidable, the specimen should be refrigerated. In general, expectorated sputa and tracheal aspirates should be screened microscopically before processing for bacterial culture to determine whether they are representative of lower respiratory secretions or saliva. This is accomplished by examining a Gram-stained smear prepared from the purulent portion of the specimen under low-power magnification to determine the number of squamous epithelial cells and/or neutrophils present. There are many ways that this information may be used to assess the quality of the specimen. A simple screening method involves assessment of the squamous epithelial cells only: more than 10 squamous epithelial cells per low-power field indicate that the specimen is contaminated with saliva and hence is not acceptable for culture. This criterion is particularly useful when evaluating specimens from neutropenic patients because the presence of neutrophils is not required for a specimen to be acceptable. Expectorated sputa for detection of *Mycoplasma pneumoniae*, *Legionella* spp, and mycobacteria do not need to be screened for adequacy.

For all acceptable lower respiratory tract specimens, the Gram-stained smear is examined under oil immersion to determine the relative amounts of organisms present. Intracellular organisms should be specifically noted. With regard to culture, use of both selective and nonselective media is recommended. In addition, a medium capable of recovering fastidious organisms, such as *Haemophilus* spp, may be inoculated if gram-negative coccobacilli are observed in the Gram-stained smear.

It is recommended that bronchoalveolar lavage and bronchial brush specimens from patients with suspected ventilator-associated pneumonia be cultured quantitatively to optimally evaluate the significance of the organisms recovered (14,15 and 16). Bronchial brush specimens, which contain approximately 0.01 to 0.001 mL of secretions, are placed in 1 mL of sterile saline or broth immediately after collection, and the sample is promptly transported to the laboratory, where a smear for Gram staining is prepared by cytocentrifugation, and 0.01 mL of the specimen is plated to appropriate media using a pipette or calibrated loop. Colony counts greater than 1,000 organisms per milliliter of broth (corresponding to 10^6 CFU/mL of original specimen) appear to correlate with infection. During bronchoalveolar lavage, 10 to 100 mL of fluid is collected. A portion of this sample is transported to the laboratory, where a smear is prepared by cytocentrifugation and Gram stained. The Gram-stain report should

include a statement regarding the presence or absence of intracellular organisms. A 0.001-mL aliquot of the specimen is inoculated onto agar media, and the recovery of 10,000 colonies or more of a specific organism per milliliter of fluid correlates with pneumonia.

Legionella spp are not an uncommon cause of community-acquired or nosocomial pneumonia. Legionellosis can be diagnosed by culture, detection of antigens in the urine, or serologic testing. Culture is preferred because, unlike the other methods, it is not limited to detection of particular species or serotypes. Because legionellae grow slowly, their optimal isolation from highly contaminated specimens, such as sputa or tracheal aspirates, is achieved by decontaminating the specimens with acid before plating. Specimens then are inoculated to buffered charcoal yeast extract agar with and without antimicrobial agents and are incubated for a minimum of 5 days.

M. pneumoniae is a major cause of primary atypical pneumonia. Because mycoplasmas are fastidious and grow very slowly, a definitive diagnosis often is based on results of serologic tests. When culture is required, the specimen of choice is a throat swab; however, sputa or other respiratory specimens are also acceptable. The specimen should be placed immediately into a transport medium containing protein, such as albumin, and penicillin to reduce growth of contaminating organisms. Specimens may be stored in the transport medium for as long as 48 hours at 4°C or frozen for longer periods at -70°C. Mycoplasma are cultured using a biphasic culture system that is incubated for 3 weeks.

Burkholderia (formerly *Pseudomonas*) *cepacia* is an important respiratory pathogen in persons with cystic fibrosis. This organism grows well on routine media; however, selective media such as PC (for *Pseudomonas cepacia*) and oxidative-fermentative-polymyxin B-bacitracin-lactose agars are useful for its optimal recovery from respiratory secretions.

C. trachomatis can cause serious respiratory disease in infants, and *Chlamydia pneumoniae* causes illness in all age groups, but the most severe disease generally occurs in young children and the elderly. *Chlamydia psittaci* is primarily an animal pathogen but occasionally causes disease in humans. *C. pneumoniae* is difficult to culture and identify and *C. psittaci* presents a significant laboratory hazard; therefore, diagnosis of infections with these two agents usually relies on serology. Specimens for *C. trachomatis* culture are collected in a medium that contains antimicrobial agents. To maintain viability of chlamydiae, the specimen should be transported to the laboratory immediately, where it is inoculated to McCoy or buffalo green monkey cells using the shell vial technique. For delays in processing of as long as 48 hours, the specimen may be stored in the refrigerator; for longer delays, the specimen should be stored at -70°C or colder.

Nocardia spp may cause respiratory disease. Specimens for culture should be transported promptly to the laboratory; however, for short delays, storage at 4°C is acceptable. Direct examination of a Gram-stained smear of the specimen may reveal thin, beaded, gram-positive organisms that are branching, filamentous, or coccobacillary in appearance (Fig. 50.2). These organisms also are partially acid fast. There are no specific media for recovery of *Nocardia*. These organisms grow on Sabouraud dextrose agar with heart infusion supplementation, brain heart infusion agar, sheep blood agar, or tryptic soy agar and usually survive mycobacterial decontamination procedures and are recovered on media designed for mycobacterial culture. For heavily contaminated specimens, such as sputa, selective buffered charcoal yeast extract agar is recommended. Cultures are incubated for as long as 3 weeks at temperatures ranging from 25° to 37°C, but 37°C is optimal.

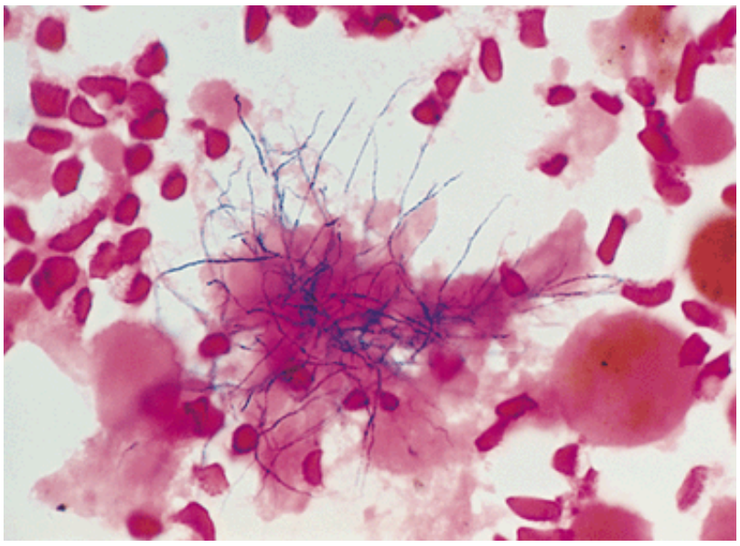


FIGURE 50.2. *Nocardia* spp in bronchoalveolar lavage fluid (Gram-Weigert stain, original magnification ×250). (Courtesy of Vicki J. Schnadig, M.D., Department of Pathology, University of Texas Medical Branch, Galveston, TX.)

To recover mycobacteria from contaminated respiratory tract specimens, the specimen must first be decontaminated and concentrated. A smear prepared from the concentrated specimen is stained for AFB. Fluorescent stains containing rhodamine and/or auramine are recommended for optimal sensitivity. Commercial nucleic acid amplification methods for direct detection of *Mycobacterium tuberculosis* complex in respiratory specimens are available. For culture, specimens are inoculated to at least one liquid and one solid medium and incubated for at least 6 weeks.

Viruses that cause respiratory tract disease in the nonimmunocompromised patients include RSV, influenza and para viruses, and adenoviruses. Diagnosis of infection with these viruses is generally accomplished utilizing a nasopharyngeal specimen; however, the virus may also be present in specimens obtained from the lower respiratory tract. Immunoassay reagents that can be used directly on specimens are available for all the viruses listed above. In most cases, culture is more sensitive than a direct assay. Traditional or shell vial culture may be performed. CMV is an important pathogen in immunocompromised individuals. For these patients, a more invasive specimen, such as the bronchoalveolar lavage, should be obtained for culture. Serology is the method of choice for diagnosing hantavirus infection.

Important fungal pathogens of the lower respiratory tract include *Coccidioides imitis*, *H. capsulatum*, *C. neoformans*, *Blastomyces dermatididis*, and *Aspergillus* spp. Fungi may be observed in a direct microscopic examination of a calcofluor/KOH preparation; however, fungal culture should also be performed. Specimens are plated to media containing antibacterial agents to prevent bacterial overgrowth and to media with and without

cycloheximide to inhibit saprophytic fungi. *Pneumocystis carinii* is an important fungal pathogen in the immunocompromised host. No culture methods are available for this organism; therefore, diagnosis relies on microscopic examination of a smear stained with Gomori methenamine silver stain (for detection of cysts) and Giemsa stain (for trophozoites), or DFA (detects both cysts and trophozoites).

Various parasites may be identified by microscopic examination of stained preparations of lower respiratory tract specimens. The modified AFB stain or an immunostain is recommended for detection of *C. parvum*, and the modified trichrome stain is recommended for microsporidia. Several stains including H&E and periodic acid–Schiff stain may be useful for detecting strongyloides, other helminths or helminth eggs, and amebae.

UPPER RESPIRATORY TRACT

Part of "50 - Specimen Collection and Processing for Microbiology"

Nasopharynx

Nasopharyngeal aspirates, washings, and swab specimens are primarily used for the diagnosis of viral respiratory infections but may also be submitted to diagnose pertussis, diphtheria, chlamydia infections, and candidiasis, as well as to identify carriers of *N. meningitidis* or *S. aureus*. Nasopharyngeal cultures have little utility in the diagnosis of otitis media; however, they may have epidemiologic value for monitoring the antibiotic susceptibility of *S. pneumoniae* in children.

The preferred specimen for the recovery of *B. pertussis* is nasopharyngeal cells collected with a small-tipped calcium alginate or Dacron swab. Rayon or cotton swabs should not be used because they may contain fatty acids that are toxic to the organism. Optimally, media should be inoculated at the bedside; however, if this is not feasible, the swab specimen may be transported to the laboratory in special transport media. For delays in transport of less than 2 hours, a 1% casamino acid solution is suitable. For delays as long as 24 hours, Amies medium with charcoal can be used. If the transport time is more than 24 hours, Regan-Lowe transport medium should be used.

Direct examination of the nasopharyngeal specimen for *B. pertussis* is performed using a DFA procedure. Depending on the antibodies used, these assays may or may not distinguish between *B. pertussis* and *Bordetella parapertussis*. In general, DFA has relatively low sensitivity and specificity for the diagnosis of pertussis and should only be performed as a supplement to and not a replacement for culture.

The preferred medium for culture of *B. pertussis* is Regan-Lowe charcoal agar containing 10% horse blood with and without cephalexin. At the bedside, the specimen is inoculated onto one third of the agar surface, and in the laboratory, the inoculum is streaked further for isolation. Cultures are incubated at 35°C for 5 to7 days in a humid atmosphere without CO_2.

For the optimal recovery of *Corynebacterium diphtheriae*, both nasopharyngeal and throat specimens should be submitted for culture. When specimens are processed for culture with no delay, no special transport medium or conditions are required. For transport to a reference laboratory, it is recommended that the specimen be sent dry in a container with dessicant. Alternatively, the specimen may be placed in Stuart's or Amies transport medium or plated on a Loeffler's serum or tellurite medium and preincubated overnight at 35°C in 5% CO_2.

A smear of specimens for diagnosis of diphtheria may be stained with the Gram stain and examined for pleomorphic gram-positive rods or with Loeffler's methylene blue and examined for pleomorphic beaded rods with swollen ends and reddish-purple metachromatic granules. The presence of organisms having the described morphology by either staining method is consistent with but not specific for *C. diphtheriae*. The specimen should be planted onto Loeffler's serum medium or a medium containing potassium tellurite for the recovery of *C. diphtheriae*.

Infections with *C. pneumoniae* most often are diagnosed serologically; however, when culture is required, a nasopharyngeal swab specimen is optimal.

To identify carriers of *S. aureus*, nasal secretions are collected from the anterior nares with a polyester tipped swab, which is placed in a tube transport system and promptly delivered to the laboratory. The specimen is planted onto 5% sheep blood agar, an agar medium selective for gram-positive organisms (i.e., colistin-nalidixic acid agar or phenylethyl alcohol agar), or mannitol salt agar, a medium selective for staphylococci and helpful in differentiating coagulase-positive from coagulase-negative species.

To identify carriers of *N. meningitidis*, a nasopharyngeal swab specimen is collected and transported to the laboratory in Amies or Stuart's medium or plated directly onto medium and transported in a CO_2-containing system. A Gram-stained smear is not useful because *N. meningitidis* cannot be distinguished morphologically from other commensal neisseriae in the upper respiratory tract. Specimens should be inoculated onto an enriched medium such as blood or chocolate agar; modified Thayer-Martin, Martin-Lewis, or New York City medium may be used as well. The cultures are incubated in a humidified atmosphere at 35°C in the presence of 5% CO_2.

Nasopharyngeal aspirates and washes are superior to nasal swab specimens for the diagnosis of viral respiratory tract infection. Specimens collected for viral culture should be placed directly into a viral transport medium and maintained on ice during storage and transport to the laboratory. Specimens are centrifuged and the sediment is used to prepare a smear for immunostaining or for other immunoassays as well as to inoculate shell vial or traditional cell cultures.

Throat

Throat swab specimens are generally collected to diagnose group A streptococcal pharyngitis or to detect shedding of viruses such as enteroviruses, HSV, or CMV. The specimen is collected by depressing the tongue, inserting the swab between the tonsillar pillars and behind the uvula, and obtaining material from the pharynx without contamination with saliva. For bacterial culture, swabs should be placed in a tube transport system containing modified Stuart's medium, and for viral culture into viral transport medium and stored on ice during transport to the laboratory.

Throat swabs for bacterial culture are routinely only evaluated for group A streptococci. Many rapid direct tests for group A streptococci are commercially available, including EIA, optical

immunoassay, and nucleic acid–based assays. The reported sensitivities of the EIA tests vary between 31% and 95% (17,18). The sensitivity of the optical immunoassay also varies, but some reports suggest that it is higher than that of EIA (19,20). In general, the nucleic acid–based test has a sensitivity greater than 90% (21,22). The specificity of all the direct tests generally is 95% to 100%. When a rapid, direct test is requested, two throat swabs should be collected. If the direct test is positive, the second swab may be discarded, but if the direct test is negative, the second swab must be cultured because the sensitivity of the direct tests is less than 100%.

To culture group A streptococci, either blood agar or selective blood agar may be used. Cultures should be incubated in an environment with reduced oxygen tension, achieved by incubating anaerobically, in 5% to 10% CO_2, or in air with a coverslip placed over the primary inoculum. These culture conditions may also allow the recovery of groups C and G streptococci, both of which may cause pharyngitis but do not cause the serious sequelae associated with group A streptococci.

Throat swabs may be helpful in determining the etiologic agent of epiglottitis, a rapidly progressing cellulitis with the potential to cause obstruction of the airway (almost always due to *H. influenzae* type b but occasionally *S. aureus* or *S. pneumoniae*). The specimen should be collected only by a physician in a setting where intubation of the patient may be performed immediately, if necessary. Specimens should be plated to an enriched medium. Alternatively, the etiologic agent of epiglottitis may be determined using blood cultures.

Culture of throat specimens is important in the diagnosis of diphtheria. Direct evaluation of the specimen and culture should be performed as previously described for nasopharyngeal specimens.

Arcanobacterium haemolyticum may cause pharyngitis and peritonsillar abscess. The organism will be recovered on the selective media used to recover group A streptococci.

To detect *N. gonorrhoeae* in the throat, the swab specimen should be inoculated at the bedside or transported promptly and inoculated as soon as possible onto a selective medium, such as modified Thayer-Martin agar.

Vincent's angina is an acute necrotizing ulcerative tonsillitis, caused by *Fusobacterium* spp, *Borrelia vincentii*, and other anaerobes, that usually occurs in adults who have poor mouth hygiene or serious systemic disease. The laboratory diagnosis is made by direct examination of a smear of a swab specimen collected from the ulcerated lesion and stained with dilute carbol fuchsin (Ziehl-Neelson stain diluted 1:10 with water) or Gram stain. The presence of many spirochetes, fusiform bacilli, and polymorphonuclear leukocytes is presumptive evidence of this disease. Cultures of the involved area usually are not helpful because many species of anaerobes are present in the oral cavity; however, blood cultures should be collected because the illness commonly is accompanied by sepsis.

The enteroviruses and CMV do not cause pharyngitis but are shed from the throat for prolonged periods after illness. Both viruses as well as HSV, which may cause symptomatic disease or may be shed asymptomatically during periods of reactivation, may be recovered in cell culture.

TISSUES

Part of "50 - Specimen Collection and Processing for Microbiology"

Tissue specimens obtained surgically are procured at great expense and considerable risk to the patient; therefore, for optimal evaluation enough material should be collected to allow both histopathologic and microbiologic examination of the specimen. The histopathology of the lesion serves to differentiate between infection and malignancy and to distinguish between acute and chronic infectious processes. For chronic infections, potential pathogens include actinomycetes, *Brucella*, mycobacteria, fungi, and parasites, any one of which may be present in small numbers, emphasizing the need for adequate samples for examination and culture. Swabs are rarely adequate, and their use should be discouraged.

After collection, tissues should be placed in a sterile container and transported rapidly to the microbiology laboratory to prevent drying. For most procedures, the specimen is placed in a small amount of sterile medium or sterile saline and homogenized by mincing with a sterile scalpel, grinding with a mortar and pestle or tissue grinder, or using a stomacher. The resulting homogenate is used to prepare smears for Gram stain or other stains as indicated by the clinical picture and to inoculate culture media. The Gram stain should be examined for the presence of microorganisms, leukocytes, and squamous epithelial cells (suggestive of surface contamination). For routine culture, tissue should be inoculated to a liquid medium (9,10) and an enriched agar medium to recover fastidious organisms.

Optimally, bone marrow aspirates are submitted to the microbiology laboratory in collection tubes for the lysis centrifugation blood culture system. Pediatric tubes should be used when less than 7 mL is collected. Bone marrow may also be submitted in a sterile container or containers containing anticoagulants. However, because anticoagulants inhibit growth of some organisms (11,12), their use is not recommended. If the laboratory chooses to process a specimen containing an anticoagulant, the ordering physician must be informed about the limitations of the results.

Gastric biopsies or brushings may be submitted for detection of *H. pylori*, an important cause of gastritis and peptic ulcer disease. The organism can be observed in tissue sections using Giemsa, H&E, or Warthin-Starry silver stain. In addition, organisms can be visualized in stained touch preparations of minced tissue. The urease test is an indirect method for diagnosing *H. pylori* gastritis or peptic ulcer disease. The organism produces large amounts of this enzyme, which can be detected by inoculating urea broth or media with a portion of the homogenized tissue. Alternatively, a piece of tissue can be placed directly into a commercially available test kit. The urease test may be positive in as little as 15 minutes but should be held for 24 hours before issuing a negative report. To recover *H. pylori* in culture, the biopsy is homogenized in 0.9% saline and inoculated onto chocolate, brain heart infusion, or brucella agar supplemented with horse or rabbit blood. The optimal agar medium for recovery of *H. pylori* has not yet been determined. The cultures are incubated at 35°C in a microaerophilic, humid environment for 7 days.

B. henselae is one agent of cat scratch disease and bacillary angiomatosis. Because culture of this organism is very difficult, diagnosis

is most often made by clinical manifestations and exclusion of other diseases or by serologic analysis. The organism may be observed in sections of fixed tissue specimens stained with Warthin-Starry silver stain or a tissue Gram stain. This organism may rarely be recovered by inoculating the specimen onto freshly prepared heart infusion agar containing 5% to 10% defibrinated rabbit or horse blood, chocolate agar, or sheep blood agar. The plates are incubated in high humidity at 35°C for as long as 1 month.

Legionella spp, *Nocardia* spp, *C. diphtheriae*, and mycobacteria are recovered using methods previously described. If an infection with a zygomycete is suspected, the tissue should be carefully minced rather than ground before microscopic examination and culture to preserve the hyphal elements. To recover pathogenic fungi, media with antibacterial agents and cycloheximide should be inoculated if there is a possibility of bacterial and fungal contamination of the specimen.

Tissue biopsies are recommended for the diagnosis of tissue parasites. In the liver and spleen, *Echinococcus*, *E. histolytica*, *L. donovani*, and microsporidia may be observed. In muscle, *Taenia solium*, *Trichinella spiralis*, *Onchocerca volvulus*, and *T. cruzi* and microsporidia are possible pathogens. Impression smear and tease or squash preparations may be examined in addition to the standard histologic preparation.

URINE

Part of "50 - Specimen Collection and Processing for Microbiology"

Acceptable methods of urine collection include midstream clean catch (preferably a first voided morning specimen), catheterization, and suprapubic aspiration. Foley catheter tips should not be accepted for culture because they usually are contaminated with urethral organisms. Most commonly, the midstream flow of a clean-catch urine is collected. All urine specimens should be transported promptly to the laboratory and processed within 2 hours of collection. If a delay in transport or processing cannot be avoided, specimens may be refrigerated for as long as 24 hours. Collection kits containing preservatives to maintain the bacterial population stable for 24 hours at room temperature are commercially available, but they offer no advantage over refrigeration.

The urinary tract above the urethra is sterile in healthy humans, but the urethra normally is colonized with many different bacteria, so urine specimens collected by a noninvasive method become contaminated during passage. Commensal bacteria are differentiated from potential pathogens in urine by quantitative culture. Originally, growth of 10^5 CFU/mL or greater was considered highly indicative of infection, but this criterion has been modified for different situations. For example, in young, sexually active women with acute urethral syndrome, as few as 10^2 CFU/mL is significant in the presence of concomitant pyuria. True urinary infections associated with fewer than 10^5 CFU/mL may occur in infants and children; in males; and in persons who are catheterized, recently treated with antimicrobial agents, drink large amounts of fluids, have symptoms and concomitant pyuria, have urinary obstruction, or have pyelonephritis acquired from hematogenous spread.

Several methods for examination of urine to identify quickly those samples that most probably will not yield clinically significant organisms are available. These screening methods correlate well with culture when growth of 10^5 CFU/mL or greater is the reference, but they compare less favorably in the presence of lower colony counts (23). Screening urine specimens by staining with Gram stain is rapid and economical with regard to reagents but is very time-consuming. Finding one or more organisms per oil immersion field in a smear prepared from an uncentrifuged specimen correlates with growth of 10^5 CFU/mL or greater. Commercially available dipstick tests that combine nitrate reductase (an enzyme produced by most gram-negative bacilli that cause urinary tract infections) and leukocyte esterase (an enzyme produced by neutrophils) are rapid, inexpensive, and simple to perform, but their sensitivity is low. Several automated urine screening systems are also commercially available including the FiltraCheck-UTI colorimetric filtration system (Meridian Diagnostics Inc., Cincinnati, OH), the UTI screen Bacterial ATP Assay (Coral Biotechnology, San Diego, CA), and the Urine ID card (bioMerieux Vitek, Inc., Hazelwood, MO).

Quantitative bacterial culture of a urine specimen is accomplished by inoculating appropriate media (e.g., blood and MacConkey agars) with a measured amount of urine, most commonly with a plastic or wire calibrated loop designed to deliver a known volume. A 0.001-mL loop is used to inoculate all urine specimens, except those collected from women with suspected acute urethral syndrome and suprapubic aspirates. Both of these latter specimens are inoculated with a 0.01-mL loop. The loop is inserted vertically into the well-mixed specimen, and the loopful of urine removed is spread over the surface of the agar plate by streaking from top to bottom in a vertical line and again from top to bottom perpendicular to the line in a back-and-forth fashion to facilitate the determination of an accurate colony count.

L. interrogans may be recovered from urine after the first week of illness and for several months thereafter. Urine should be processed as soon as possible after collection because acidity may harm the organisms. If a delay in processing is expected, the urine should be diluted 1:10 in 1% bovine serum albumin and stored at 5° to 20°C. Both undiluted urine and urine diluted 1:10, 1:100, and 1:1,000 in sterile buffered saline should be inoculated to Fletcher's, Ellinghausen-McCullough/Johnson-Harris, or polysorbate 80 medium with and without neomycin. Cultures are incubated at room temperature for at least 6 weeks.

Mycobacteria can be recovered from urine using procedures previously discussed.

Many commercial bacterial antigen test kits developed to diagnose bacterial meningitis include procedures for use on urine specimens. In general, these kits should not be used on urine specimens for the diagnosis of bacterial meningitis, and their use on CSF should be limited, as discussed previously. Of note, in 1996, the U.S. Food and Drug Administration issued a product alert specifically cautioning against the use the group B streptococci antigen kits on urine specimens because of the risk of both false-positive and -negative results.

A radioimmunoassay and an EIA are commercially available for the detection of the *Legionella pneumophila* serogroup 1 antigen in urine. Antigen may be detectable for months in the urine after an infection. These assays have sensitivities of 80% or greater when performed on concentrated urine and specificities

of greater than 98%. The drawback of both assays is that only one *L. pneumophila* serogroup is detected.

Recently, nucleic acid amplification assays have become commercially available for the detection of *N. gonorrhoeae* and *C. trachomatis* in both male and female urine samples. The assays have different specimen collection and handling instructions (refrigeration of specimen or addition of a special reagent immediately after collection) that must be followed to obtain accurate results.

Candida spp may be isolated from routine bacterial urine cultures. Recovery of other fungi from a urine specimen is best achieved by centrifuging the specimen and plating the sediment to fungal media. The *H. capsulatum* urine antigen test, which is performed in one reference laboratory, may be useful for diagnosis of disseminated histoplasmosis.

The most commonly isolated viruses from urine are CMV, HSV, and adenovirus. The shell vial technique is often used for CMV, and traditional cell culture is used to recover HSV and adenovirus.

T. vaginalis is probably the most frequently observed parasite in urine. The organism is identified in a wet mount of the specimen, examined under low power magnification for organisms with the characteristic shape and motility. *T. vaginalis* may also be recovered in culture using one of several commercially available media that is inoculated with the specimen at the time of sample collection. For both the wet mount and culture, the specimen should never be refrigerated because this causes the organisms to rapidly lose motility and die. Other parasites that may be seen in a wet mount or stained preparation of urine specimens are *Schistosoma haematobium*, microsporidia, or the microfilariae.

SKIN AND SUBCUTANEOUS LESIONS

Part of "50 - Specimen Collection and Processing for Microbiology"

Ideally, the infected material is aspirated with a needle and syringe. For transport, the material is expelled into a sterile container that is tightly capped and promptly delivered to the laboratory. If an aspirate cannot be obtained, swab specimens of exudate collected from the deep portion of the lesion are acceptable. A minimum of two swab specimens should be collected, one for culture, the other for preparation of smears for staining. For bacterial and fungal cultures, swabs may be placed in tube transport systems containing modified Stuart's medium. To recover anaerobes, an additional swab specimen must be collected and placed in an anaerobic transport system. For viral culture, the specimen (aspirate or swab) should be placed in viral transport media and kept on ice. Specimens should be transported promptly to the laboratory. If a delay in processing is unavoidable, specimens may be stored in the refrigerator, except those for recovery of anaerobes, which should be maintained at room temperature. Processing specimens involves preparation of a smear for Gram-stain examination and inoculation of appropriate media for culture. When swab specimens are processed, it is not necessary to inoculate a broth medium because organisms recovered from broth cultures of swab specimens are rarely significant (10,24).

If detection of mycobacteria is requested, specimens should be decontaminated and concentrated. The sediment is used to prepare a smear for staining for AFB and to inoculate mycobacterial media. If the specimen is from a skin lesion, a separate set of media should be inoculated and incubated at 30°C to enhance recovery of mycobacteria that preferentially grow at lower temperatures such as *Mycobacterium marinum*, *M. ulcerans*, *M. chelonae*, and *M. haemophilum*. To recover *M. haemophilum*, inoculation to a medium containing blood is also required. *M. leprae* may be observed in the acid-fast smear but will not grow *in vitro*. A nucleic acid amplification test for detection of *M. leprae* in tissue biopsies is available by special request through the Gillis W. Long Hansen's Disease Center in Carville, LA.

Exudates submitted for fungal culture should be examined carefully for granules, and, if present, they must be crushed and examined for fungal elements. Nail specimens submitted for dermatophyte examination should be cut into small pieces before microscopic examination and culture. Skin (and hair) may be examined without processing. Other specimens are minced and then examined microscopically for fungal elements in a calcofluor/KOH preparation. Specimens should be inoculated to media containing antimicrobial agents and cycloheximide to prevent overgrowth with bacteria and saprophytic fungi.

Viruses that cause a maculopapular exanthem will not be present in the lesion, which is caused by the host's hypersensitivity response to the infection. In contrast, virus is present in vesicular lesions. For vesicles of VZV, a Tzanck preparation should be prepared by carefully removing the top of the vesicle with a needle, soaking up the excess fluid with gauze, and pressing a slide against the base of the lesion. The smear is optimally examined using a direct immunofluorescent stain. The Tzanck smear may also be used to detect HSV, but the most sensitive method of detection is cell culture of the vesicular fluid and material collected from the base of the lesion. In most cases, culture is positive within 48 hours. Enteroviruses may also be isolated from samples of vesicular lesions.

Parasites present in ulcers and other skin lesions include *Leishmania* spp, *Acanthamoeba* spp, microfilaria, and *O. vulvulus*. As for other specimens, detection of parasites relies on microscopic evaluation of stained material.

REFERENCES

1. Forbes BA, Sahn DF, Weissfeld AS, eds. *Bailey and Scott's diagnostic microbiology*, 10th ed. St. Louis: Mosby, 1998:999.

2. Li J, Plorde JJ, Carlson LG. Effects of volume and periodicity on blood cultures. *J Clin Microbiol* 1994;32:2829–2831.

3. Weinstein MP, Reller LB, Murphy JR, et al. The clinical significance of positive blood cultures: a comprehensive analysis of 500 episodes of bacteremia and fungemia in adults. I. Laboratory and epidemiologic observations. *Rev Infect Dis* 1983;5:35–53.

4. Maki DG, Weise CE, Sarafin HW. A semiquantitative culture method for identifying intravenous-catheter-related infection. *N Engl J Med* 1977;296:1305–1309.

5. Miller JM, Holmes HT. Specimen collection, transport, and storage. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*. Washington, DC: ASM Press, 1999:4:57.

6. Granoff DM, Murphy TV, Ingram DL, et al. Use of rapidly generated results in patient management. *Diagn Microbiol Infect Dis* 1986;4:157S–166S.

7. Maxson S, Lewno MJ, Schutze GE. Clinical usefulness of cerebrospinal fluid bacterial antigen studies. *J Pediatr* 1994;125:235–238.

8. Perkins MD, Mirrett S, Reller LB. Rapid bacterial antigen detection is not clinically useful. *J Clin Microbiol* 1995;33:1486–1491.

9. Morris AJ, Wilson SJ, Marx CE, et al. Clinical impact of bacteria and fungi recovered only from broth cultures. *J Clin Microbiol* 1995;33:161–165.

10. Morris AJ, Smith LK, Mirrett S, et al. Cost and time savings following introduction of rejection criteria for clinical specimens. *J Clin Microbiol* 1996;34:355–357.

11. Evans GL, Cekoric T, Searcy RL. Comparative effects of anticoagulants on bacterial growth in experimental blood cultures. *Am J Med Technol* 1968;34:103–112.

12. Rosett W, Hodges GR. Antimicrobial activity of heparin. *J Clin Microbiol* 1980;11:30–34.

13. Schuchat A, Whitney C, Zangwill K. Prevention of perinatal group B streptococcal disease: a public health perspective. *MMWR Morb Mortal Wkly Rep* 1996;45:1–24.

14. Baselski VS, El-Torky M, Coalson JJ, et al. The standardization of criteria for processing and interpreting laboratory specimens in patients with suspected ventilator-associated pneumonia. *Chest* 1992;102:571S–579S.

15. Chastre J, Fagon J, Bornet-Lecso M, et al. Evaluation of bronchoscopic techniques for the diagnosis of nosocomial pneumonia. *Am J Respir Crit Care Med* 1995;152:231–240.

16. Sanchez-Nieto JM, Alcaraz AC. The role of bronchoalveolar lavage in the diagnosis of bacterial pneumonia. *Eur J Clin Microbiol Infect Dis* 1995;14:839–850.

17. Carroll K, Reimer L. Microbiology and laboratory diagnosis of upper respiratory tract infections. *Clin Infect Dis* 1996;23:442–448.

18. Wegner DL, Witte DL, Schrantz RD. Insensitivity of rapid antigen detection methods and single blood agar plate culture for diagnosing streptococcal pharyngitis. *JAMA* 1992;267:695–697.

19. Harbeck RJ, Teague J, Crossen GR, et al. Novel, rapid optical immunoassay technique for detection of group A streptococci from pharyngeal specimens: comparison with standard culture methods. *J Clin Microbiol* 1993;31:839–844.

20. Heiter BJ, Bourbeau PP. Comparison of two rapid streptococcal antigen detection assays with culture for diagnosis of streptococcal pharyngitis. *J Clin Microbiol* 1995;33:1408–1410.

21. Heelan JS, Wilbur S, Depetris G, et al. Rapid antigen testing for group A streptococcus by DNA probe. *Diagn Microbiol Infect Dis* 1996;24:65–69.

22. Heiter BJ, Bourbeau PP. Comparison of the Gen-Probe group A streptococcus direct test with culture and a rapid streptococcal antigen detection assay for diagnosis of streptococcal pharyngitis. *J Clin Microbiol* 1993;31:2070–2073.

23. Pezzlo M. Detection of urinary tract infections by rapid methods. *Clin Microbiol Rev* 1988;1:268–280.

24. Silletti RP, Ailey E, Sun S, et al. Microbiologic and clinical value of primary broth cultures of wound specimens collected with swabs. *J Clin Microbiol* 1997;35:2003–2006.

51
Bacteriology

David L. Sewell

The purpose of a clinical microbiology laboratory is to provide accurate, timely, clinically relevant information about the identity, antimicrobial susceptibility, and significance of microorganisms isolated from a given specimen and to help prevent the spread of infection. The issue is complex because of the ubiquitous presence of microorganisms that colonize skin and mucous membranes; the transient presence of organisms in normally sterile sites; the enormous diversity of specimen type and source; the importance of the appropriate selection, collection, and transport of the specimen to the laboratory; and the need to interpret the significance of the recovery of a microorganism on the basis of the patient's history and symptoms. The routine workup of a specimen must be flexible enough to accommodate the special situations related to the patient's history and underlying condition. Compounding the problem is the rapidly increasing diversity of methods to perform these tasks, ranging from basic culture techniques to miniaturization and automation of conventional identification schemes to immunoassays and molecular methods. These techniques must be evaluated in the target population while remaining cost-effective. The operation of an efficient laboratory requires close communication between the laboratory and the clinician to expedite the early diagnosis and treatment of the patient.

This chapter attempts to address some of these points. The emphasis is on the type of disease produced, the diagnostic procedures used to identify the causative agent, interpretation of clinical significance, and some consideration of antimicrobial therapy. Detailed descriptions of methods are not presented, but the reader is directed to the excellent texts listed as suggested readings.

STAPHYLOCOCCI AND RELATED GENERA

Part of "51 - Bacteriology"

Microbiology

Until recently, the family Micrococcaceae consisted of the four genera *Staphylococcus, Micrococcus, Planococccus*, and *Stomatococcus* (1). However, the staphylococci and micrococci are not closely related based on DNA base composition, nucleic acid hybridization, or 16S ribosomal RNA analysis. The genus *Staphylococcus* is closely related to the new genus *Macrococcus* (2) and the broadly defined *Bacillus-Lactobacillus-Streptococcus* cluster (1,3). Members of the genus *Micrococcus* have been divided into the six genera *Micrococcus, Kocuria, Kytococcus, Nesterenkonia, Arthrobacter*, and *Dermacoccus* and are referred to as micrococci in this chapter. Members of the genera *Planococcus* and *Macrococcus* are infrequently, if ever, isolated from clinical specimens and are not discussed further. Micrococci and *Stomatococcus mucilaginosus* are common inhabitants of the skin and mucous membranes of humans and are infrequently implicated in significant infections.

Staphylococci are catalase-positive, gram-positive cocci, 0.5 to 1.5 μm in diameter that occur singly, in pairs, and in tetrads, and usually divide in more than one plane to form irregular grapelike clusters. The cell wall contains peptidoglycan and teichoic acid. Most species are facultative anaerobes, capable of growing in the presence of 10% NaCl and between 18° and 40°C. *Staphylococcus aureus* subsp *anaerobius* and *S. saccharolyticus* are catalase negative and grow best under anaerobic conditions. Most staphylococci grow well on noninhibitory media, but infrequently some strains may require hemin, menadione, or CO_2 for growth. Staphylococci are susceptible to lysis by lysostaphin but resistant to lysozyme.

Currently, there are 32 species and 15 subspecies of staphylococci, seven of which are coagulase positive (3). *S. aureus, S. epidermidis*, and *S. saprophyticus* are the most frequently isolated species of staphylococci, but more than half the other coagulase-negative species have been implicated as opportunistic pathogens (Table 51.1).

Spectrum of Disease

Because of the prevalence of staphylococci on skin and mucous membranes, they are implicated in various infections ranging from skin infections to bacteremia and meningitis. *S. aureus* often produces an acute, pyogenic infection that may lead to bacteremia and metastatic foci in other organs or sites in the body. Most common localized infections involve the skin and soft tissue, including pustules, furuncles, impetigo, cellulitis, and postsurgical wounds. Dissemination may result in a more serious infection such as bacteremia, endocarditis, meningitis, osteomyelitis, septic arthritis, pneumonia, pericarditis, or renal abscess. Factors predisposing to *S. aureus* infections include impaired integrity of the skin barrier (e.g., burns, surgery, ulcers, minor trauma, poor personal hygiene) and underlying conditions (e.g., eczema, acne, impaired humoral and cellular immunity, presence of foreign bodies) and diseases (e.g., malignancy, viral infections).

In addition to pyogenic infections, *S. aureus* is associated with toxin-induced diseases. The most common form is gastroenteritis. This self-limited disease occurs 2 to 6 hours after the consumption of food containing thermostable enterotoxins produced by *S. aureus* after contamination and improper storage of food. Symptoms consist of acute onset of nausea and vomiting, followed by diarrhea.

TABLE 51.1. PATHOGENIC SIGNIFICANCE OF *STAPHYLOCOCCUS* SPECIES

Species	Pathogenic Significance
S. aureus	Common
S. epidermidis	Common
S. saprophyticus	Common
S. cohnii	Uncommon
S. haemolyticus	Uncommon
S. hominis	Uncommon
S. saccharolyticus	Uncommon
S. simulans	Uncommon
S. warneri	Uncommon
S. intermedius[a]	Rare/undetermined
S. capitis	Uncommon
S. caprae[a]	Uncommon
S. auricularis	Rare/undetermined
S. xylosus	Uncommon
S. arlettae[a]	Rare/undetermined
S. equorum[a]	Rare/undetermined
S. gallinarum[a]	Rare/undetermined
S. kloosit[a]	Rare/undetermined
S. carnosus[a]	Rare/undetermined
S. sciuri[a]	Rare/undetermined
S. lentus[a]	Rare/undetermined
S. hyicus[a]	Rare/undetermined
S. lugdunensis	Uncommon
S. schleiferi	Uncommon
S. chromogenes[a]	Rare/undetermined
S. delphini[a]	Rare/undetermined
S. pasteuri[a]	Rare/undetermined
S. muscae[a]	Rare/undetermined
S. vitulus[a]	Rare/undetermined
S. piscifermentans[a]	Rare/undetermined
S. felis[a]	Rare/undetermined
S. lutrae[a]	Rare/undetermined

[a] Isolated from animals or animal products. Modified from Pfaller MA, Koontz F. Coagulase-negative staphylococci: rationale for species identification. *Am Soc Clin Pathol Check Sample* 1989; 32:1-7.

The other toxin-mediated diseases, toxic shock syndrome (TSS) and scalded skin syndrome, are often associated with infection or colonization with *S. aureus*. Scalded skin syndrome occurs primarily in children and is caused by an exfoliative toxin. TSS results from the *in situ* production of TSS toxin-1 (TSST-1) and is manifested by high fever, rash, profound hypotension, diarrhea, and shock. TSST-1 is a superantigen that stimulates T cells and induces the production of tissue necrosis factor and interleukin-1 (4). Initially, most patients were young menstruating women who used highly absorbent tampons, but now TSS is seen in men and women.

Coagulase-negative staphylococcal infections are usually indolent, associated with implanted prosthetic devices, usually caused by *S. epidermidis*, and hospital acquired. The exception is *S. saprophyticus*, which causes urinary tract infections in young sexually active women (5). Infections caused by other coagulase-negative staphylococci are increasing as the numbers of catheters and indwelling foreign bodies placed in patients increase. Usually these infections are associated with intravenous catheters, peritoneal dialysis catheters, vascular grafts, cerebrospinal fluid shunts, prosthetic joints, and heart valves. Other infections include endocarditis, bacteremia, surgical wound infections, osteomyelitis, peritonitis, infections after ocular surgery, or genitourinary infections associated with catheterization. *S. intermedius* (coagulase positive) is associated with dog-bite infections and may be misidentified as *S. aureus* (6).

Epidemiology

Staphylococci are ubiquitous in the human ecosystem, inhabiting the skin, mucous membranes, and gastrointestinal (GI) tract, and are a leading cause of both community- and hospital-acquired infections.

S. aureus colonizes the skin, perineal area, umbilical stumps, and sometimes the GI tract of neonates shortly after birth. The anterior nasal vestibule is the primary colonization site in adults. Approximately 20% to 40% of adults in the community are carriers, with much higher rates observed in hospital personnel, patients receiving periodic hemodialysis, and injection drug abusers. Although *S. aureus* can survive on inanimate objects in the environment, person-to-person transmission is the most important route of spread, especially on the hands of the hospital staff.

Staphylococcal food poisoning follows the contamination of foodstuffs by food handlers who are either infected or colonized, and the subsequent inadequate refrigeration of the contaminated food. Foods most commonly associated with this disease include meats, dairy products, and salads.

Infections by coagulase-negative staphylococci are usually hospital acquired, caused by *S. epidermidis*, and occur in patients with an indwelling foreign body. The patient's endogenous flora is the usual reservoir for coagulase-negative staphylococci infections, but transmission from hospital personnel to the patient has been documented. Hospital strains are usually multiple antibiotic resistant and colonize the skin and GI tract of hospitalized patients and hospital staff. Frequent hand washing and the use of gloves by hospital personnel are the most effective methods for minimizing the spread of staphylococci within the hospital environment. Most likely, coagulase-negative staphylococci gain access to deeper tissues during the insertion of the foreign body. Strains of coagulase-negative staphylococci often exhibit a propensity to adhere to plastics and other biomaterials by producing a capsular polysaccharide-adhesion followed by formation of a biofilm (7). This may allow the organisms to escape host defenses.

Of the coagulase-negative staphylococci, *S. epidermidis* is most frequently isolated from the skin (nares, head, axillae, leg, arms, toe webs) and represents 65% to 90% of all staphylococci recovered (8). *S. hominis* is the next most frequently isolated species. Other coagulase-negative staphylococcal species are only occasionally isolated from the skin (*S. simulans, S. xylosus, S. cohnii, S. warneri, S. saccharolyticus,S. haemolyticus*) or occupy special niches on the skin [*S. capitis* (head), *S. auricularis* (ear), *S. saprophyticus* (genitourinary skin)].

It is often difficult to differentiate infection from colonization and to document transmission and spread of specific strains within the hospital environment. Historically, phage typing, phenotypic markers, or antibiograms were used to identify an outbreak strain with a unique pattern. These techniques lack sensitivity and specificity for following transmission of coagulase-negative staphylococci. More recently, molecular and genetic techniques have provided excellent discrimination between species and strains. These methods include endonuclease restriction analysis of plasmid and chromosomal DNA, ribotyping, and nucleic acid probe hybridization.

Diagnostic Procedures

Microscopy

Gram-stained smears of clinical specimens provide rapid and presumptive evidence of staphylococcal infection. Staphylococci appear as gram-positive cocci (0.5 to 1.5 µm in diameter) and occur singly or in pairs, short chains, or small clusters (Fig. 51.1). These features are not always observed in clinical specimens in which staphylococci have been ingested by phagocytes or exposed to cell wall–active antimicrobial agents or are in the stationary phase of growth. In these situations, the bacteria may appear gram variable or negative. The Gram-stain morphology cannot definitively differentiate staphylococci from other gram-positive cocci seen in clinical specimens.

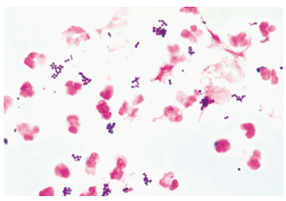


FIGURE 51.1. Gram stain of material containing polymorphonuclear neutrophils and *Staphylococcus aureus*.

Culture Techniques

Most staphylococci grow well on nonselective agar and broth media containing peptone (e.g., sheep blood agar, brain-heart infusion, thioglycollate). On sheep blood agar, growth occurs within 24 hours at 35° to 37°C. Colonies are 1 to 3 mm in diameter, usually opaque, smooth, and circular, and have a butyrous consistency. Colonies of *S. aureus* may be pigmented (yellow or yellow-orange) and exhibit β hemolysis on sheep blood agar, but these observations may also occur with some other staphylococcal species. Colonies of *S. epidermidis* are usually gray-white and nonhemolytic. Nutritionally variant staphylococci require supplementation of the medium with hemin, menadione, thiamine, or pantothenate for growth.

Specimens likely to be contaminated with other bacteria are inoculated onto a selective medium such as colistin-nalidixic acid (CNA) agar, phenylethyl alcohol (PEA) agar, mannitol-salt agar (MSA), Schleifer-Kramer agar, or lipase-salt-mannitol agar or tellurite-glycine agar. These media inhibit the growth of gram-negative bacteria. On MSA, *S. aureus*, *S. simulans*, *S. capitis*, and *S. saprophyticus* are differentiated from other clinical species of coagulase-negative staphylococci by their ability to produce acid from mannitol. Selective media should be incubated at least 48 hours for optimal growth.

Identification Methods

The initial identification of a suspected colony begins with colonial morphology, Gram-stain morphology, and a spot-catalase test to eliminate organisms such as streptococci and corynebacteria. When it is clinically relevant to differentiate between staphylococci and micrococci, three relatively simple tests can be used in the clinical laboratory (9). They include susceptibility to bacitracin (micrococci), lysostaphin (staphylococci), and furazolidone (staphylococci). Some clinically significant *Staphylococcus* species are identified by the colony morphology, growth characteristics, enzyme activities, and other biochemical characteristics listed in Table 51.2. In addition to these conventional methods, several commercial identification systems and automated instruments provide clinical laboratories with the capability of identifying the staphylococci with an accuracy of 70% to 90% (3, 10).

TABLE 51.2. IDENTIFICATION OF CLINICALLY SIGNIFICANT *STAPHYLOCOCCUS SPECIES*

Character	S. aureus	S. epidermidis	S. hominis	S. haemolytucs	S. warneri	S. saprophyticus	S. simulans	S. lugdunensis
Colony pigment	+[a]	-	d	d	d	d	-	d
PYR	-	-	-	+	-	-	+	+
Coagulase	+	-	-	-	-	-	-	-
Deoxyribonuclease	+	-	-	-	-	-	-	-
Hemolysis	+	(d)	-	(+)	(d)	-	(d)	(d)
Phosphatase (alkaline)	+	+	-	-	-	-	(d)	-
Polymyxin B resistance	+	+	-	-	-	-	-	d
Novobiocin resistance	-	-	d	-	-	+	-	+
Acid (aerobically) from								
Maltose	+	+	+	+	+	+	±	+
D-Trehalose	+	-	d	+	+	+	d	+
D-Mannitol	+	-	-	d	d	d	+	+
Sucrose	+	+	(+)	+	+	+	+	+
D-Mannose	+	(+)	-	-	-	-	d	+

[a] PYR, L-pyrrolidonyl-β-naphthylamide; +, 90% or more strains positive; –, 90% or more strains negative; d, 11–89% of strains positive; ±, 90% or more strains weak positive; (), delayed reaction requiring 48 to 72 hr.
Modified from Kloos WE, Bannerman TL. *Staphylococcus* and *Macrococcus*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:272.

S. aureus, *S. epidermidis*, and *S. saprophyticus* can also be identified with a high degree of accuracy on the basis of colony morphology; coagulase and phosphatase activities; acid production from D-mannitol, D-trehalose, maltose, sucrose, and D-xylose; and novobiocin resistance. *S. aureus* is identified in most cases by the coagulase test. Although *S. intermedius*, *S. schleiferi* subsp *coagulans*, *S. lutrae*, *S. delphini*, and *S. hyicus* are also coagulase positive, they are rarely isolated from humans. The standard tube test detects free staphylocoagulase and is the more definitive test, whereas the slide coagulase test detects clumping factor. These two tests have been replaced in large part by the commercially available particle agglutination tests that measure clumping factor and protein A. These systems are quite accurate but may miss some methicillin-resistant *S. aureus* (11). Other alternative methods to the coagulase test are available (10).

Direct Antigen or Nucleic Acid Detection

Rapid identification of *S. aureus* in culture can be accomplished by the Accuprobe method (GenProbe, Inc., San Diego, CA). This test has also been used for the direct identification of *S. aureus* in blood cultures (12). Currently, there is no commercially available nucleic amplification test for the direct detection of *S. aureus* in clinical specimens.

Serologic Procedures

The serologic response to various extracellular products and cell-wall components of *S. aureus* has been measured in an attempt

to diagnose some forms of staphylococcal disease. With the exception of antibodies to TSST-1, most of these tests have little clinical utility.

Antibodies to TSST-1, measured by immunodiffusion (ID), radioimmunoassay (RIA), and enzyme immunoassay (EIA), may aid in the diagnosis of TSS, but TSS is usually diagnosed by clinical signs and symptoms. The patient also has TSS if no antibodies are detected, but the strain produces TSST-1 because convalescent antibodies are detected in only approximately 40% of TSS patients.

Interpretation of Results

Staphylococci are frequently isolated in the clinical microbiology laboratory because of their ubiquitous presence on the skin and mucous membranes of humans. The isolation of *S. aureus* from an infected site should be considered clinically relevant until proven otherwise.

The complete identification of coagulase-negative staphylococci should be considered on clinically significant isolates recovered from normally sterile body sites, peritoneal dialysates, tissue, or multiple sets of blood cultures. The clinical relevance of a specific isolate is based on clinical, microbiological, and epidemiologic criteria such as the immune status and presence of a foreign body in the patient, the species of *Staphylococcus* isolated, and the number of specimens from which the identical strain is isolated.

Therapeutic Considerations

Accurate antimicrobial susceptibility testing must be performed on all clinically significant isolates of staphylococci because of their variable resistance to commonly used antimicrobial agents. Resistance to methicillin and other semisynthetic penicillins is widespread throughout the world. In addition, methicillin-resistant *S. aureus* and other staphylococci usually exhibit cross-resistance to other β-lactam antibiotics as well as increased resistance to macrolides, clindamycin, aminoglycosides, trimethoprim-sulfamethoxazole, tetracyclines, and quinolones. Most strains are susceptible to rifampin, but this drug cannot be used alone because of the high mutation rate to resistance. Currently, the drug of choice for the empiric treatment of infections caused by methicillin-resistant staphylococci is vancomycin until susceptibility to other β-lactam antibiotics can be determined.

The recent appearance of *S. aureus* and coagulase-negative staphylococci with reduced susceptibility to vancomycin [minimum inhibitory concentration (MIC), 4 to 16 mg/mL] is a serious concern (13,14 and 15). Infection control guidelines for the prevention and control of staphylococci with increased resistance to vancomycin have been developed (16,17).

Summary and Conclusions

S. aureus, S. epidermidis, and *S. saprophyticus* are associated with the majority of staphylococcal infections. *S. aureus* causes infections of skin, soft tissue, bloodstream, and internal organs; staphylococcal food poisoning; and toxic shock syndrome. *S. epidermidis* causes nosocomial infections, especially in patients with indwelling foreign bodies. *S. saprophyticus* causes urinary tract infections in young women.

Staphylococci are usually transmitted from person to person with the primary reservoir being the skin and mucous membranes of humans and animals. Most staphylococci are easily isolated on routine media and are identified by morphology, conventional biochemical and metabolic characteristics, and commercial identification kits. Because of the variability of resistance to antimicrobial agents (especially the penicillinase-resistant penicillins), susceptibility testing must be performed.

STREPTOCOCCI, ENTEROCOCCI, AND RELATED GENERA

Part of "51 - Bacteriology"

Microbiology

Molecular taxonomic studies have clarified, to some extent, and expanded the taxonomy of the streptococci and related genera.

The taxonomic picture is improving but will continue to change in the coming years. Although the new taxonomic information is critical for a better understanding of streptococci and related genera, the older identification schemes based on phenotypic characteristics are still useful in the clinical laboratory. The preliminary classification of streptococci is based on their hemolytic action on blood agar and differentiates β-hemolytic, α-hemolytic, and nonhemolytic (γ-hemolytic) groups. Further differentiation (especially of the β-hemolytic streptococci) is achieved by serologic classification based on the presence of group-specific antigens and currently performed by commercial particle agglutination kits. With the exception of the group D streptococci, the majority of the α-hemolytic (viridans streptococci) and nonhemolytic species are classified by physiologic tests. The interrelationship of these three schemes is depicted in Table 51.3.

TABLE 51.3. CLASSIFICATION OF *STREPTOCOCCUS* SPECIES MOST FREQUENTLY CAUSING HUMAN INFECTIONS

Species	Group Antigen	Hemolysis	Sherman Classification	Clinical Features
S. pyogenes (large colony)	A	β, γ	Pyogenic	Pharyngitis, tonsillitis, otitis media, mastoiditis, sinusitis, scarlet fever, erysipilus, cellulitus, impetigo, pneumonia, toxic shock-like syndrome, endometritus, bacteremia. Nonsuppurative sequelae: acute rheumatic fever, acute glomerulonephritis
S. agalactiae	B	β, γ	Pyogenic	Neonatal sepsis and meningitis, chorioamnionitis, puerperal sepsis, bacteremia, skin and soft tissue infection, urinary tract infection
S. dysgalactiae subsp. *equisimilis* (large colony)	C, G	β	Pyogenic	Upper respiratory, skin and soft tissue infections, bacteremia, deep tissue infection
E. faecalis (other enterococci)	D	α, β, γ	Enterococci	Endocarditis, urinary tract infection, bacteremia, intraabdominal abscess, soft-tissue infection
S. bovis	D	γ	Nonenterococci	Endocarditis, bacteremia
S. pneumoniae	—	α	—	Pneumonia, bacteremia, meningitis, otitis media, sinusitis
S. anginosus/milleri (*S. intermedius, S. constellatus*)	F,A,C,G, None	α, β, γ	Pyogenic viridans	Endocarditis, bacteremia, deep-tissue infection
S. mitis, S. mutans *S. salivarius* *S. morbillorum*	None or A-O	α, γ	Viridans	Endocarditis (*S. mutans* major cause of caries)
Abiotrophia	None	α	Viridans	Endocarditis
Other viridans streptococci	None	α	Viridans	Rare

The β-hemolytic streptococci with Lancefield group A, C, or G antigen are further divided into large (more than 0.5 mm in diameter) and small (less than 0.5 mm in diameter) colonies (18). The large colony groups are pyogenic, and the small colony streptococci belong to the *S. milleri* or *S. anginosus* group (*S. anginosus, S. intermedius, S. constellatus*), which also includes non-β-hemolytic, viridans streptococci. Other β-hemolytic streptococci are principally animal pathogens and include *S. dysgalactiae* subsp *dysgalactiae* (groups C and L), *S. equi* subsp *equi* (group C), *S. equi* subsp *zooepidemicus* (group C), and *S. canis* (group G).

Historically, the group D streptococci were divided into enterococcal and nonenterococcal strains based on tolerance to salt. Currently, the enterococci are placed in the genus *Enterococcus*, which is composed of 17 species (19). *E. casseliflavus* and *E. flavescens* are closely related at the species level (20). *E. faecalis*, and *E. faecium* are isolated most often from human sources. The nonenterococcal group D isolates *S. bovis, S. alactolyticus*, and *S. equinus* remain as streptococci.

The non-β-hemolytic strains (viridans) include *S. pneumoniae* and five groups: the Bovis group (*S. bovis, S. alactolyticus, S. equinus*), the Mutans group (*S. mutans, S. sobinas, S. cricetus, S. rattus, S. downei, S. macacae*), the Salivarius group (*S. salivarius, S. vestibularis, S. thermophilus*), the Mitis group (*S. mitis, S. sanguis, S. parasanguis, S. gordonii, S. crista, S. oralis*), and the Anginosus group (*S. anginosus, S. intermedius, S. constellatus*) (18). The nutritionally variant streptococci were considered forms of *S. mitis* or *S. sanguis* but are not genetically related to the streptococci and are placed in a new genus, *Abiotrophia*.

The medically important streptococci are catalase-negative, gram-positive cocci and spherical to ovoid (less than 2 μm in diameter) and tend to form pairs or chains. Some species may appear as short rods on initial isolation. Most species are facultative anaerobes, exhibit a fermentative metabolism, and have complex but variable nutritional requirements. Growth on solid agar is enhanced by blood, serum, or glucose.

The examination of gram-positive cocci by molecular-based taxonomic methods has characterized several genera that are opportunistic pathogens in compromised patients but are infrequently isolated from clinical specimens (21). These organisms can be preliminary differentiated based on morphology of cells grown in broth. The genera *Leuconostoc, Lactococus*, and *Globicatella* appear as coccobacilli in pairs and chains, whereas *Aerococcus, Alloiococcus, Gemella, Helcococcus*, and *Pediococcus* are spherical cocci in pairs, tetrads, or clusters. These bacteria are catalase negative or weakly positive. *Leuconostoc* and *Pediococcus* (and *Lactobacillus*) are resistant to vancomycin and may be confused with viridans streptococci (22). Additional characteristics useful for differentiation of these genera include gas produced from glucose (positive only for *Leuconostoc*),

pyrrolidonyl arylamidase activity (negative for *Leuconostoc, Pediococcus*, and *A. urinae*), and leucine aminopeptidase activity (negative for *Leuconostoc, A. viridans, Helcococcus*, and *Globicatella*).

Spectrum of Disease

Streptococci are part of the normal flora that cause a wide variety of infections, including upper respiratory infections, skin and soft-tissue infections, bacteremia, meningitis, urinary tract infections, deep-tissue infections, endocarditis, arthritis, and osteomyelitis, as well as nonsuppurative sequelae observed after infection by *S. pyogenes* (Table 51.3).

Group A streptococci (*S. pyogenes*) are important human pathogens and a major cause of upper respiratory and skin infections. Suppurative complications of streptococcal pharyngitis include parapharyngeal space abscesses, otitis media, sinusitis, mastoiditis, and meningitis. Bacteremia may lead to infections in the joints, bones, or other foci. In addition, group A streptococci can cause a severe invasive infection associated with a TSS-like syndrome. It is believed that the pyrogenic exotoxins function as superantigens, capable of inducing the release of lymphokines and cytokines, causing the TSS-like syndrome (23). Nonsuppurative sequelae after pharyngitis include acute rheumatic fever (ARF) and acute glomerulonephritis (AGN). Skin infections precede the development of AGN but not ARF.

Group B streptococci (*S. agalactiae*) are a major cause of infections in neonates and postpartum women, including neonatal sepsis and meningitis, puerperal sepsis, and chorioamnionitis. The most common adult infections occur in postpartum women and are associated with the genital tract, but various other infections also occur. Age, diabetes mellitus, malignancy, and liver disease are important risk factors.

Group C and G streptococci are not frequent human pathogens but are associated with infections of the respiratory tract and skin, otitis media, bacteremia, pneumonia, meningitis, neonatal sepsis, puerperal sepsis, endocarditis, and arthritis.

Of the enterococci, *E. faecalis* causes disease more frequently than the other species and is implicated in the following infections: endocarditis, intraabdominal infections, soft-tissue infections, urinary tract infections, and bacteremia. It is rarely associated with meningitis or pneumonia.

The group D streptococci (nonenterococcal) *S. bovis* causes primarily endocarditis and bacteremia, which is often associated with malignancies or instrumentation of the GI tract.

S. pneumoniae is a major pathogen in infants and adults, especially the elderly or persons with asplenia, sickle cell disease, or immune disorders. It is the most frequent cause of community-acquired pneumonia and may cause mastoiditis, sinusitis, otitis media, meningitis, arthritis, and peritonitis.

S. anginosus/milleri group is associated with abscesses in various areas of the body (e.g., skin, brain, liver, abdomen, cervicofacial area, and lung).

The viridans streptococci are usually associated with cases of endocarditis but are increasingly recovered from neutropenic patients (24,25).

Epidemiology

The epidemiology of the streptococci is as diverse as the number of species in this group, which reside on the mucous membranes of the respiratory, genital, and GI tracts. Group A streptococcal pharyngitis is a common infection of young school-aged children, with a peak incidence occurring during the winter and spring. Transmission is ordinarily by person-to-person contact and via nasal or salivary droplets. Outbreaks owing to food or water contamination are also documented. Crowding, such as occurs in schools and military barracks, favors transmission. Asymptomatic carriage of group A streptococci occurs in the throats of as many as 20% of school-aged children. The epidemiology of ARF parallels that of streptococcal pharyngitis.

Streptococcal pyoderma occurs primarily as an endemic disease in children but can occur in epidemics. Crowding, a tropical or subtropical climate, and poor living conditions favor the spread of skin streptococci. AGN may occur after streptococcal pharyngitis or pyoderma.

Group B streptococci colonize the genital and/or lower GI tract of pregnant women at a rate ranging from 5% to 40%. Colonization of the newborn occurs either *in utero* or more commonly during delivery. Nosocomial transmission also occurs, especially in the setting of heavy maternal colonization and crowded nursery conditions.

Enterococci and group D streptococci are found as normal flora in the GI and genital tract of humans. Nearly all adults harbor *E. faecalis* as well as *E. faecium* (26). *S. bovis* occurs in 5% to 10% of normal adults but is found more frequently in patients with colon cancer.

S. pneumoniae is part of the normal oropharyngeal flora in 5% to 70% of adults. Transmission is person to person and is increased by crowding such as occurs in military barracks. Of the 84 capsular serotypes, the majority of infections are caused by the lower number serotypes.

The viridans streptococci constitute 30% to 60% of the oral flora, with various species residing in different ecological niches of the oropharynx. They also colonize the mucous membranes of the urogenital and GI tracts.

Diagnostic Procedures

Microscopy

Streptococci and enterococci are typically ovoid or lancet shaped but may elongate to resemble short rods. Most often they occur in pairs or chains when grown in broth (Fig. 51.2). *S. pneumoniae* appears with the adjacent cell wall flattened, lancet shaped, and encapsulated and occurs predominantly in pairs but may form short chains (Fig. 51.3). It may occasionally appear to be gram-negative if observed in phagocytes, if the patient is receiving antimicrobial agents, or if the culture is old. The other related genera usually appear coccoid except that *Gemella haemolysans* is easily decolorized and may appear gram negative.

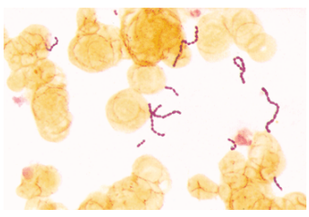


FIGURE 51.2. Gram stain of blood culture broth positive for streptococci.

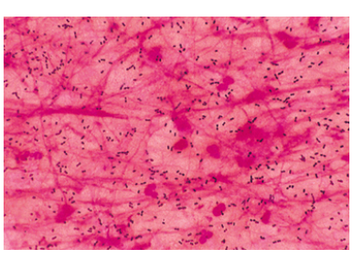


FIGURE 51.3. Gram stain of *Streptococcus pneumoniae* in sputum.

Culture Techniques

Streptococci and enterococci grow well on rich, nonselective broths and agars such as trypticase soy, heart infusion, or Todd-Hewitt,

which may be supplemented with blood. Selective agars such as PEA or CNA can be used with specimens containing gram-negative organisms. Trimethoprim-sulfamethoxazole is added to sheep blood agar to improve the recovery of group A streptococci from throat cultures. *Abiotrophia* species are recovered by satelliting around a *Staphylococcus* streak, by adding pyridoxal to the growth medium, or by placing a disk containing pyridoxal on the agar plate.

Streptococci and enterococci are facultative anaerobes, often requiring 5% to 10% CO_2 (especially *S. pneumoniae*) for optimal growth. Approximately 90% to 95% of *S. pyogenes* isolates from patients with pharyngitis can be recovered from sheep blood agar incubated in air or sheep blood agar (with or without trimethoprim-sulfamethoxazole) incubated anaerobically for 48 hours (27).

The Centers for Disease Control and Prevention (CDC) recommend that all pregnant women should be screened for the perineal vaginal carriage of group B streptococci at 35 to 37 weeks of gestation with a selective broth medium (28). To achieve a more rapid detection of group B streptococcal colonization, growth in the selective broth is detected by a molecular nucleic acid probe assay (29,30).

Identification Methods

The initial characterization of streptococci and enterococci is based on catalase reaction, hemolysis, colony morphology, and source of the specimen. Streptococci are catalase negative, but weak reactions may occur, especially with enterococci and some related genera. The type of hemolytic reaction is usually observed with surface colonies, but the most accurate method uses subsurface colonies. Group A streptococci produce large zones of ß hemolysis; group B streptococci produce large colonies with a small zone of ß hemolysis; enterococci usually have large gray colonies with no hemolysis (may be α- or rarely ß-hemolytic), and the viridans streptococci are small, raised colonies with α or no hemolysis. *S. pneumoniae* is α-hemolytic with dome-shaped colonies that may have a depressed center or appear mucoid.

Cultures of the *S. anginosus*/*S. milleri* group may produce a buttery or caramel odor (31). Clinically significant streptococci and enterococci are presumptively identified by the tests listed in Table 51.4. Group A streptococci are inhibited by bacitracin and hydrolyze L-pyrrolidonyl-ß-naphthylamide (PYR). Group B streptococci are hippurate and CAMP positive and PYR negative. Other ß-hemolytic streptococci are negative for these tests but are susceptible to trimethoprim-sulfamethoxazole. The ß-hemolytic streptococci are definitively identified by detection of the group-specific cell-wall carbohydrate antigen by any of a number of commercially available particle agglutination kits.

TABLE 51.4. PRESUMPTIVE IDENTIFICATION OF STREPTOCOCCI AND ENTEROCOCCI

		Susceptibility		Hydrolysis of						
Category	Hemolysis	Bacitracin	SXT[a]	Hippurate	PYR[a]	CAMP	Bile Esculin	Growth in 6.5% NaCl	Optochin and Bile[b]	VP[a]
Group A	ß	+	-	-	+	-	-	-	-	-
Group B	ß	-[c]	-	+	-	+	-	+[c]	-	-
ß-Hemotytic streptococci (not group A, B, or D)	ß	-[c]	+	-	-	-	-	-	-	-
Enterococcus	α, ß, none	-	-	-[c]	+	-	+	+	-	-
Group D, non-*Enterococcus*	α, none	-	+[c]	-	-	-	+	-	-	+
Viridans group	α, none	-[c]	+	-[c]	-	-	-[c]	-	-	-[c]
S. pneumoniae	α	±	?	-	-	-	-	-	+	?
S. milleri group	α,ß	?	?	-	-	-	-	-	-	+

[a] SXT, trimethoprim-sulfamethoxazole; PYR, L-pyrroidonyl-ß-naphthylamide; VP, voges-proskauer, +, positive reaction or susceptible: -, negative reaction or resistant.

[b] Optochin susceptibility and bile solubility.

[c] Exceptions occasionally occur.

α-Hemolytic colonies are tested for bile solubility or susceptibility to optochin. *S. pneumoniae* is positive for both tests; other streptococci are negative. The slide bile solubility test is useful for the rapid presumptive identification of *S. pneumoniae* from blood cultures (32). Enterococci and group D streptococci are presumptively identified by hydrolysis of esculin in the presence of bile, growth in 6.5% NaCl, and hydrolysis of PYR. Further identification of enterococci, group D streptococci, and the viridans streptococci require additional identification schemes (18,19,21), or identification by commercial kits. Strains belonging to the *S. anginosus*/*S. milleri* group can be presumptively identified by positive reactions for Voges-Proskauer and arginine hydrolysis, esculin hydrolysis, and lack of growth in 6.5% NaCl (33).

Direct Antigen or Nucleic Acid Detection

Several commercial particle agglutination kits are available for the direct detection of streptococcal cellular or capsular antigens in body fluids [e.g., cerebrospinal fluid (CSF), urine, serum] or throat swabs. Many factors affect the sensitivity and specificity of the test, including the type of specimen and the particular kit.

The identification of group A streptococcal antigen from throat swabs is a rapid method to aid in the diagnosis of streptococcal pharyngitis. A positive test is an excellent predictor of infection, but the sensitivity is often lower than 90%, especially when few organisms are present. A negative test should be followed with a culture.

The direct detection of other streptococcal antigens in clinical specimens is possible but generally not clinically useful. The Neufeld capsular swelling reaction (or Quellung reaction) is another method to detect directly pneumococci in sputum or CSF.

Antiserum is not widely available but can be obtained from the Statens Seruminstitut, Copenhagen, Denmark.

Molecular probe assays are commercially available for the direct detection of group A streptococci in throat swabs, group B streptococci in vaginal specimens, and culture identification of enterococci and *S. pneumoniae*.

Serologic Procedures

The detection of antibodies to streptococcal cell-wall antigen or extracellular products aids in the diagnosis of recent group A streptococcal infection. These tests are used most frequently for the diagnosis of ARF or AGN in the absence of a recently documented infection. The most helpful tests are antistreptolysin O, antihyaluronidase, and antideoxyribonuclease B. The lack of significant titers to these three antigens makes the diagnosis of ARF very doubtful. In cases of AGN, antistreptolysin O titers are usually very low.

Interpretation of Results

The clinical relevance of the isolation of any of the streptococcal species depends on the type of specimen from which the organism was recovered and knowledge of the spectrum of infections caused by the organism. In addition, one must consider the clinical signs and symptoms of the patient because all the streptococci are either normal human flora or can be carried asymptomatically. In general, the isolation of ß-hemolytic streptococci and *S. pneumoniae* should be interpreted as clinically relevant unless shown otherwise.

The clinical relevance of the recovery of enterococci and viridans streptococci from specimens is related to the severity of the infection. However, it is important to identify the *S. anginosus*/*S. milleri* group and *S. bovis* from body fluids and blood cultures because of their association with deep-seated abscesses and GI cancer, respectively.

Therapeutic Considerations

The ß-hemolytic streptococci remain uniformly susceptible to penicillin and vancomycin and thus do not require routine susceptibility testing. If erythromycin is considered, the isolate should be tested for susceptibility to the drug. Nearly all strains of *S. bovis* are susceptible to penicillin or ampicillin. The enterococcal isolates exhibit intrinsic resistance to the cephalosporins, clindamycin, aminoglycosides, and trimethoprim-sulfamethoxazole at standard concentrations. In addition, the most active agents such as the penicillins and vancomycin are not bactericidal, requiring the use of combination therapy (penicillin plus aminoglycoside) for life-threatening infections. High-level resistance to aminoglycosides is increasing. More disturbing is that several enterococcal species now exhibit resistance to vancomycin, the penicillins, and high levels of aminoglycosides.

A significant change has occurred in the antimicrobial susceptibility patterns of the viridans streptococci and *S. pneumoniae*. Penicillin resistance, resistance to the extended spectrum cephalosporins, and other antimicrobial agents are increasing worldwide. Pneumococcal strains demonstrating multiple resistance to erythromycin, tetracycline, trimethoprim-sulfamethoxazole, and chloramphenicol have been isolated. The extent of resistance to antimicrobics varies with the geographic area.

Although most strains of viridans streptococci, including the *S. anginosus*/*S. milleri* group, remain susceptible to penicillin, the frequency of penicillin resistance is increasing. Because of this unpredictable resistance to antimicrobics, susceptibility testing should be performed on clinically relevant isolates.

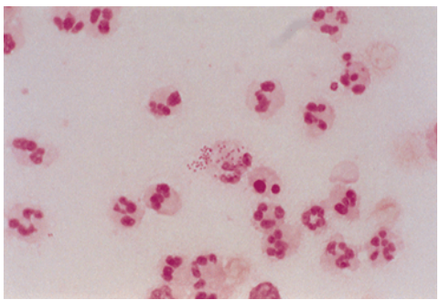


FIGURE 51.4. Gram stain of *Neisseria gonorrhoeae* in urethral exudate.

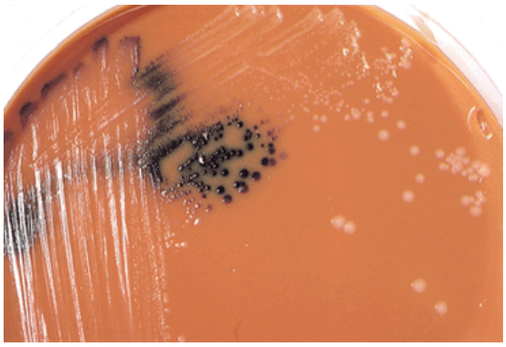


FIGURE 51.5. Colonies of *Neisseria gonorrhoeae* on Martin-Lewis medium after addition of oxidase reagent. Oxidase-positive colonies turn purple.

Summary and Conclusions

The streptococci and related genera represent a heterogeneous collection of medically important bacteria that cause a broad spectrum of infections (e.g., pharyngitis, skin and soft-tissue infections, deep-tissue infections, meningitis, and urinary tract infections). The application of nucleic acid techniques has improved the classification of the streptococci and related genera. Most strains can be isolated readily from clinical specimens with the use of enriched laboratory media and can be identified with a combination of serologic, physiologic, and biochemical tests. Commercial identification systems are also available as are kits for the direct detection of streptococcal antigen in clinical specimens. Although many of the streptococci remain susceptible to penicillin, increasing resistance to penicillins, cephalosporins,

aminoglycosides, and, in the case of enterococci, to vancomycin is being observed.

NEISSERIA AND BRANHAMELLA (*MORAXELLA*)

Part of "51 - Bacteriology"

Microbiology

The family Neisseriaceae contains four genera, *Neisseria, Moraxella, Acinetobacter*, and *Kingella*. The latter three genera are discussed later. In addition to the two commonly recognized pathogens *Neisseria gonorrhoeae* and *N. meningitidis*, the genus *Neisseria* contains several usually nonpathogenic species of human origin. These include *N. lactamica, N. sicca, N. subflava, N. mucosa, N. polysaccharea, N. flavescens, N. cinerea*, and *N. elongata*. *N. canis* and *N. caviae* are found in the upper respiratory tract of cats and *N. denitrificans* in guinea pigs. *N. kochii* is most likely a subspecies of *N. gonorrhoeae* (34).

The genus *Branhamella* contains four species, of which only *B. catarrhalis* is medically important. Because *Branhamella* has been designated as a subgenus within *Moraxella*, this species is referred to as *Moraxella (Branhamella) catarrhalis* (35). Because the taxonomic position remains uncertain and similar morphology to *Neisseria* and the common identification systems, the terminology *B. catarrhalis* is used here.

Neisseria and *Branhamella* species are gram-negative, oxidase- and catalase-positive cocci that appear most often in pairs with the adjacent sides flattened (kidney-bean shape). *N. elongata* exists as a small coccobacillus and is catalase negative. These bacteria are aerobic and grow best on enriched media containing blood or serum incubated in a humid atmosphere (5% to 7% CO_2) at 35°to 37°C.

Spectrum of Disease

The clinical spectrum of gonococcal infections involves primarily the mucous membranes of the lower genitourinary tract and less frequently those of the rectum, oropharynx, and conjunctiva. A small proportion of men may be asymptomatic, whereas as many as 80% of women may be asymptomatic or minimally symptomatic, contributing to the transmission of gonorrhea.

Gonorrhea in men usually presents as acute urethritis. If untreated, it may progress to prostatitis, epididymitis, or periurethral abscess, but these complications are rare. Anorectal gonorrhea occurs in approximately 40% of homosexual men and pharyngeal infection in approximately 25%.

Mucopurulent cervicitis is the primary manifestation in women, although involvement of the urethra, rectum, periurethral glands, and the Bartholin gland ducts may occur. The genital infection in 10% to 20% of women progresses to pelvic inflammatory disease (e.g., endometritis, salpingitis, tuboovarian abscess, peritonitis). Infected mothers may transmit the disease to their babies *in utero*, during delivery, or postpartum. The conjunctiva is the major site of infection in newborns. Gonococcal vulvovaginitis may occur in prepubertal girls.

Spread of gonococci from the genitourinary tract, the rectum, or the pharynx to the bloodstream occurs in 0.3% to 5% of infected persons. Disseminated gonococcal infection results in arthritis and dermatitis, with rare complications including meningitis, osteomyelitis, and endocarditis.

N. meningitidis causes bacteremia (meningococcemia), meningitis, chronic meningococcemia (with arthritis and dermatitis), arthritis, and pneumonia. Rarely, it has been implicated as the cause of sinusitis, endocarditis, pharyngitis, and urethritis. The presentation may be subacute to fulminant.

The other *Neisseria* species rarely cause diseases such as meningitis, bacteremia, endocarditis, pericarditis, and ocular infections. *B. catarrhalis* is an important cause of otitis media and sinusitis in children and bronchopulmonary infections in the elderly or persons with underlying pulmonary deficiency, malignancy, or immune deficit. Infrequently, *B. catarrhalis* is the cause of meningitis, septic arthritis, endocarditis, bacteremia, and ocular infections.

Epidemiology

Gonorrhea is a common, worldwide sexually transmitted disease, second only to *Chlamydia trachomatis* infections in the United States (36). Because *N. gonorrhoeae* colonizes only human mucous membranes, close contact between individuals is required for transmission. Risk factors include young age (younger than 24 years old), urban residence, low educational and socioeconomic background, nonwhite race, prostitution, and multiple sexual partners.

Humans are the only natural host for *N. meningitidis*. The organism resides on the nasopharyngeal mucosal membranes and is transmitted via respiratory droplets. The disease predominates in infants and children during endemic periods and shifts to older populations during epidemic periods. Although 13 serogroups are recognized, strains belonging to the serogroups A, B, C, Y, and W-135 are most frequently associated with disease.

The other *Neisseria* species are found as part of the oropharyngeal flora in humans. *B. catarrhalis* is not frequently isolated from the oropharynx of adults, but carriage is much higher in children, elderly adults, and patients with chronic pulmonary or respiratory tract disease (34,37).

Diagnostic Procedures

Microscopy

A Gram-stained smear of the appropriate clinical specimen revealing numerous polymorphonuclear leukocytes and many intracellular and extracellular Gram-negative diplococci is nearly diagnostic for *B. catarrhalis* pulmonary infection, gonococcal urethritis and cervicitis, and meningococcal meningitis (Fig. 51.4). The sensitivities of the Gram stain relative to the culture for the detection of gonococci in urethral exudates from symptomatic men, endocervical exudates, and rectal specimens are 95% to 100%, 40% to 60%, and 40% to 60%, respectively (38). The specificity of the smear is 95% to 100% for specimens other than rectal.

Culture Techniques

Neisseria species and *B. catarrhalis* can be isolated on 5% sheep blood and enriched chocolate agars. *N. gonorrhoeae* requires chocolate agar (grows poorly on blood agar). *N. meningitidis*

grows on blood agar, and the other *Neisseria* species and *B. catarrhalis* grow on media devoid of blood. Optimal growth occurs in 5% to 7% CO_2-enriched humid atmosphere at 35° to 37°C. Plates should be incubated for 72 hours. A selective medium (e.g., modified Thayer-Martin, Martin-Lewis, New York City medium) is used for the isolation of *N. gonorrhoeae* and *N. meningitidis* from sites that are heavily colonized with other bacterial flora. Some strains of *N. lactamica, N. cinerea, N. polysaccharea, N. subflava*, and *B. catarrhalis* grow on these selective media, and some strains of *N. gonorrhoeae* are inhibited.

N. gonorrhoeae appears as gray to white opaque glistening colonies of 0.5 to 1.0 mm in diameter (Fig. 51.5). *N. meningitidis* produces a round, translucent blue-gray colony that is approximately 1 to 2 mm in diameter. *B. catarrhalis* colonies are grayish-white, opaque, and smooth. The other *Neisseria* species may or may not produce pigmented colonies.

Identification Methods

The definitive identification of *Neisseria* species and *B. catarrhalis* is based on carbohydrate utilization tests and other physiologic characteristics (Table 51.5). *Kingella denitrificans* is included in the table because it grows on the *Neisseria* selective agar and produces acid from glucose. Consequently, it may be mistaken for pathogenic *Neisseria*. *N. gonorrhoeae* grows on selective media and produces acid from glucose, whereas *N. meningitidis* produces acid from both glucose and maltose. *B. catarrhalis, N. flavescens, N. elongata*, and *N. cinerea* are all asaccharolytic. *B. catarrhalis* is differentiated by the DNase reaction. *N. flavescens* produces polysaccharide in 1% sucrose, whereas *N. cinerea* does not. *N. elongata* is catalase negative.

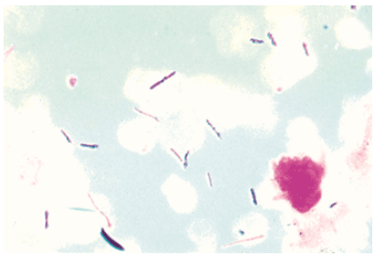


FIGURE 51.6. Gram stain of blood culture broth containing *Bacillus* species.

TABLE 51.5. CHARACTERISTICS OF HUMAN *NEISSERIA* SPECIES, *BRANHAMELLA CATARRHALIS*, AND *KINGELLA DENITRIFICANS*

Species	Pigment[a]	Superoxol[b]	Acid Produced from					Polysaccharide from ≥1% Sucrose	Reduction of		DNase	Extra CO_2 Needed[e]	Growth on		
			Glucose	Maltose	Fructose	Sucrose[c]	Lactose (ONPG)		NO_3	NO_2[d]			MTM, ML, or NYC Medium	Chocolate, Blood Agar at 22°C	Nutrient Agar at 35°C
N. gonorrhoeae	-	+	+	-	-	-	-	-	-	-	-	VI	+[f]	-	-
N. meningitidis	-	-	+	+	-	-	-	-	-	d	-	I	+	-	+
N. lactamica	-	-	+	+	-	-	+	-	-	d	-	d	+	-	+
N. cinerea	-	-	-[g]	-	-	-	-	-	-	+	-	d	-[h]	-	+
N. polysaccharea	-	-	+	+	-	-	-	+	-	d	-	d	+	-	+
N. kochii	-	+	+	-	-	-	-	-	-	-	-	No	+	-	+
N. flavescens	+	-	-	-	-	-	-	+	-	-	-	I	+	-	+
N. sicca	d	-	+	+	+	+	-	+	-	+	-	No	-	+	+
N. subflava[i,j] Biovar *subflava*	+	-	+	+	-	-	-	-	-	+	-	No	-	+	+
Biovar *flava*	+	-	+	+	+	-	-	-	-	+	-	No	-	+	+
Biovar *perflava*	+	-	+	+	+	+	-	+	-	+	-	No	-[j]	+	+
N. mucosa	+	-	+	+	+	+	-	+	+	+	-	No	-	+	+
B. catarrhalis	-	-	-	-	-	-	-	-	+	-	+	No	d	+	+
K. denitrificans	-	-	+	-	-	-	-	-	+	-	-	I	+	-	-

[a] Pigment observed in colonies on nutrient agar. Strains of *N. cinerea* and *N. lactamica* are yellow-brown and yellow pigmented when growth is harvested on a cotton applicator or smeared on filter paper.

[b] All *Neisseria* species and *B. catarrhalis* give a positive catalase best with 3% H_2O_2; *N. gonorrhoeae* strains give strong reactions with 30% H_2O_2 [superoxol] whereas other species are negative.

[c] Some strains may be inhibited by 5% sucrose; reactions may be obtained on a starch-free medium containing 1% sucrose. Strains of *N. gonorrhoeae, N. meningitidis*, and *N. kochii* do not grow on this medium.

[d] Results for tests in 0.1% (wt/vol) nitrite; *N. gonorrhoeae* strains and strains of some other species that are negative in 0.1% nitrite can reduce 0.01% (wt/vol) nitrite.

[e] Extra CO_2: VI, very important; I, important for growth; No, not needed for growth.

[f] ≥90% of vancomycin-susceptible strains of *N. gonorrhoeae* may not grow on TM or MTM medium.

[g] Some strains of *N. cinerea* may give a weak reaction in glucose in some rapid tests for the detection of acid from carbohydrates.

[h] Some strains of *N. cinerea* have been isolated on gonococcal selective medium, but are colistin susceptible and will not grow when subcultured on selective media. Colistin-resistant mutants of *N. cinerea* have not been described.

[i] Strains of *N. subflava* biovars give consistent patterns of acid production when tested in appropriate media.

[j] Some strains of *N. subflava* biovar *perflava* grow on gonococcal selective media in primary culture, are colistin resistant, and grow on selective media on subculture.

From Knapp JS. Historical perspectives and identification of *Neisseria* and related species. *Clin Microbiol Rev* 1988;1:417.

ONPG, *o*-Nitrophenyl-β-D-galactopyranoside; DNase, deoxyribonuclease; MTM, modified Thayer-Martin medium; ML, Martin-Lewis medium; NYC, New York City medium; +, most strains (≥90%) positive; -, most strains (≥90%) negative: d, some strains positive, some strains negative.

Commercial kits are available for the rapid identification of *Neisseria* and *B. catarrhalis* based on reactions with carbohydrate or chromogenic substrates. The accuracy and the number of species identified vary with the system. Confirmation of an isolate as *N. gonorrhoeae* can be done by fluorescent antibody techniques, particle agglutination tests, enzyme immunoassay (EIA), or by use of nucleic acid probe.

Direct Antigen or Nucleic Acid Detection

The direct detection of gonococci in clinical specimens can be performed by EIA, molecular probe, or nucleic acid amplification methods. The EIA (Abbott Laboratories, North Chicago, IL) is as sensitive and specific as the Gram stain for detecting gonococci in urethral and urine specimens (collected from symptomatic men) but less sensitive and specific than culture in women (39,40). The nucleic acid probe (GenProbe) is sensitive (more than 95%) and specific (more than 99%) for the detection of gonococci in endocervical and urethral specimens (41,42) but less sensitive for pharyngeal specimens (43). Highly sensitive and specific nucleic amplification tests are available for the detection of gonococci in urine (single and pooled), endocervical, and urethral specimens (44,45 and 46). Because of the use of urine specimens, the ease of transport, and the ability to detect concurrently *Chlamydia trachomatis*, the molecular methods are replacing the traditional culture.

Commercial particle agglutination kits are used for the direct detection of *N. meningitidis* in body fluids. The sensitivity of the procedure (70% to 90%) varies with the concentration of antigen and the kit. Most kits exhibit a high degree of specificity.

Serologic Procedures

There are no reliable commercially available tests to diagnose gonococcal infections serologically. The serologic determination of an individual's exposure to *N. meningitidis* is primarily of epidemiologic interest.

Interpretation of Results

The isolation of *N. gonorrhoeae* is always clinically relevant. Because *N. meningitidis* and *B. catarrhalis* are normal flora in the upper respiratory tract, their significance depends on the type and quality of the specimen from which they were recovered and the nature of the infectious process. Assessing the clinical significance of the other *Neisseria* species is more difficult because they rarely cause disease.

TABLE 51.6. MICROSCOPIC MORPHOLOGY OF GRAM-POSITIVE BACILLI

Genus	Microscopic Appearance	Comments
Kurthia	Rods in chains, cocci in old cultures	Strict aerobe
Arthrobacter	Marked rod-coccus cycle	Strict aerobe
Brevibacterium	Short rod, cocci in old cultures	Strict aerobe
Bacillus	Medium to large rods	Spores, decolorizes easily
Listeria	Small rods, coccobacilli, short chains	
Oerskovia	Coccoid, extensive branching filaments	
Rothia	Diphtheroid, coccoid, branching filaments	
Mycobacterium	Small to medium rods	Acid-fast, aerobe
Rhodococcus	Rod, coccoid, branching	Pink-coral; sometimes acid fast
Corynebacterium	Diphtheroid, coccobacilli	
C. matruchotii	Rod attached to filament (whip), branching	
Arcanobacterium	Diphtheroid, curved, pleomorphic	
Erysipelothrix	Rods, chains, coccobacilli, pleomorphic	
Nocardia	Branching filaments, short rods	Weakly acid-fast
Lactobacillus	Medium-thick rods, chains	Anaerobe but some grow in CO_2
Propionibacterium	Diphtheroid, some branching	Anaerobe but some grow in CO_2
Actinomyces	Rods, filaments with branching	Anaerobe but some grow in CO_2
Clostridium	Medium to large rods, some filaments	Anaerobe but some grow in CO_2

FIGURE 51.7. Gram stain of blood culture broth containing *Listeria*.

The isolation of a gram-negative, oxidase-positive diplococcus

on *Neisseria*-selective agar from a genital specimen is presumptively *N. gonorrhoeae*. The identity must be confirmed by additional testing, especially in low-risk patients because of the medicolegal implications of a diagnosis of gonorrhea (47).

Therapeutic Considerations

All clinically relevant isolates of *Neisseria* species and *B. catarrhalis* should be tested for ß-lactamase production because of the increasing incidence of plasmid or chromosomally mediated resistance to the penicillins. Ceftriaxone is recommended for the treatment of uncomplicated gonorrhea in adults. Additional consideration should be given for other coexisting sexually transmitted diseases. A follow-up culture should also be obtained after treatment. Penicillin, the extended spectrum cephalosporins, or chloramphenicol is recommended for the treatment of severe meningococcal infections. Rifampin is recommended for the chemoprophylaxis of the household contacts. Generally, the other *Neisseria* species are susceptible to penicillin.

The emergence of resistance to penicillins, tetracyclines, and spectinomycin in *N. gonorrhoeae* has prompted the CDC to recommend treatment with the extended spectrum cephalosporins and the newer fluoroquinolones (48,49). Resistance to ciprofloxacin and ofloxacin has been reported, especially in the Far East. If patients fail to respond to treatment, the isolate should be recovered from culture and susceptibilities determined. As always, the patient should be evaluated for other sexually transmitted diseases.

Approximately 85% of strains of *B. catarrhalis* produce ß-lactamase and should be considered resistant to penicillin, ampicillin, and amoxicillin (50). The majority of isolates are susceptible to cephamycins, ß-lactamase-resistant penicillins, trimethoprim-sulfamethoxazole, fluoroquinolones, and the newer macrolides.

Summary and Conclusions

The two clinically important species of *Neisseria* are *N. gonorrhoeae* (causing gonorrhea with all the clinical manifestations) and *N. meningitidis* (causing meningococcemia and meningitis). *B. catarrhalis* is an important cause of otitis media and sinusitis in children and bronchopulmonary infections in adults.

The definitive laboratory diagnosis is made by isolation and identification of the etiologic agent. Numerous commercial systems are available for identification of *Neisseria* and *B. catarrhalis*. Direct detection of gonococci by EIA and molecular methods are available.

The emergence of *N. gonorrhoeae* strains resistant to the penicillin, tetracyclines, and fluoroquinolones has resulted in the use of more third-generation cephalosporins for the treatment of infections.

GRAM-POSITIVE AEROBIC TO FACULTATIVELY ANAEROBIC BACILLI

Part of "51 - Bacteriology"

Microbiology

The gram-positive aerobic to facultatively anaerobic bacilli are a taxonomically diverse group of bacteria that have been grouped together based primarily on cellular and colony morphology, endospore formation, and aerobic growth. When a gram-positive rod is isolated from a clinical specimen, the genera that should be initially considered include *Bacillus, Erysipelothrix, Corynebacterium, Mycobacterium, Nocardia, Lactobacillus*, and occasional strains of *Clostridium, Actinomyces*, and *Propionibacterium* because these genera are more commonly isolated or associated with infections. *Lactobacillus, Clostridium, Actinomyces*, and *Propionibacterium* are discussed in the anaerobe section. *Mycobacterium* and *Nocardia* are addressed in another chapter. Other genera that need to be considered but are rarely isolated or infrequently associated with human infections include *Arcanobacterium, Brevibacterium, Kurthia, Rothia, Arthrobacter, Oerskovia, Rhodococcus, Turicella, Dermabacter, Curtobacterium, Microbacterium, Aureobacterium, Cellulomonas, Exiguobacterium, Gordona, Tsukamurella, Dermatophilus, Dietza, Micromonspora*, and the actinomycetes. Most of these organisms are either part of the normal skin or mucous membranes of humans or are found in soil, water, or foodstuffs. For a discussion of these genera, see references 51 and 52.

It has been proposed that the genus *Bacillus* be divided to include five new genera: *Alicyclobacillus, Paenibacillus, Brevibacillus, Aneurinibacillus*, and *Virgibacillus* (53). It is likely that further taxonomic changes will occur.

The genus *Bacillus* contains more than 50 recognized species, arranged into rRNA sequence groups. *B. anthracis* and *B. cereus* are the two most important pathogens. They are gram-positive straight rods of variable size (2 to 10 µm in length and 0.5 to 2.5 µm in width) that occur singly or in chains (Fig. 51.6). The major characteristic for inclusion in this genus is the formation of endospores (which does not always occur on routine laboratory media). The organisms grow on simple media and are strict aerobes or facultative anaerobes, catalase positive, oxidase variable, and motile (except *B. anthracis* and *B. mycoides*).

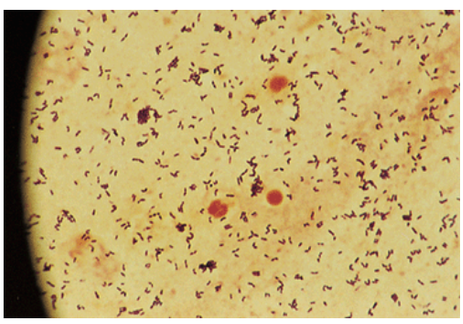


FIGURE 51.8. Gram stain of *Corynebacterium* in sputum.

Based on DNA homology and 16 S rRNA sequencing, the genus *Listeria* contains six species and two subspecies (54). Only *Listeria monocytogenes* and rarely *L. ivanovii* are pathogenic for humans. Because species other than *L. monocytogenes* are seldom isolated from clinical specimens, they are not discussed. *Listeria* is a short gram-positive rod that occurs singly, in short chains, or

in parallel or V forms (Fig. 51.7). It grows on most laboratory media and is facultatively anaerobic, catalase positive in most strains, and motile at room temperature.
The genus *Erysipelothrix* contains two species, *E. rhusiopathiae* and *E. tonsillarum*. They are catalase-negative, nonmotile, gram-positive, straight or slightly curved rods that occur singly, in pairs, in short chains, or in filaments. Only *E. rhusiopathiae* causes infections in humans.
The genus *Corynebacterium* is most closely related to *Mycobacterium* and the partially acid-fast genera *Nocardia* and *Rhodococcus*, based on cell-wall and fatty-acid composition (1). There are currently 46 species of which 31 cause infections (55). Corynebacteria are gram-positive, straight or slightly curved rods with tapered or club-shaped ends that occur singly or in angular (V or L) forms and palisade arrangements (Fig. 51.8). They are nonmotile, catalase positive, and facultatively anaerobic and form metachromatic granules. *Arcanobacterium haemolyticum* may morphologically resemble *Corynebacterium* and is catalase negative and nonmotile.
Rothia dentocariosa is a gram-positive rod that may appear coccoid, diphtheroid, or filamentous. It is catalase variable, facultatively anaerobic, and nonmotile and is part of the normal oropharyngeal flora.

Spectrum of Disease

B. anthracis is the etiologic agent of anthrax that, depending on the portal of entry, presents in three forms: (a) cutaneous anthrax (most common), (b) pulmonary anthrax (rare), or (c) GI anthrax (no reported cases in the United States). If untreated, anthrax is often fatal. Anthrax is rarely seen in the United States. *B. cereus*, the most frequently isolated *Bacillus* species, causes food poisoning and along with the numerous other *Bacillus* species, causes a variety of opportunistic infections such as localized infections related to trauma (especially fulminant eye infections) and systemic infections (e.g., meningitis, endocarditis, bacteremia, and pneumonia).
L. monocytogenes causes a mild febrile illness, especially in pregnant women; it also causes severe disease, including meningitis, bacteremia, *in utero* infections, or localized infections (including skin, ocular, arthritis, abscess, osteomyelitis, and endocarditis), which occur most often in immunocompromised patients.
E. rhusiopathiae infections usually involve localized cutaneous lesions (erysipeloid) or less frequently a disseminated form that is often associated with endocarditis.
C. diphtheriae causes infections of the human upper respiratory tract (diptheria) and skin. The other corynebacteria cause a wide array of infections in patients who are immunocompromised or have a prosthesis (56). *C. jeikeium* is associated with endocarditis, bacteremia, foreign body infections, and wound infections; *C. pseudodiphtheriticum* with respiratory tract infections and endocarditis, and *C. urealyticum* with urinary tract infections.
A. haemolyticum is associated with pharyngitis in young adults. *Rothia* strains are associated with endocarditis, bacteremia, abscesses, dental caries, and periodontal disease but are infrequently encountered.

Epidemiology

Anthrax occurs worldwide and is primarily a disease of herbivores. Transmission to humans occurs by contact with infected animals or their products (e.g., wool, hides, goat hair). The reservoir is contaminated soil where spores can survive for long periods. The other *Bacillus* species are inhabitants of soil and therefore common contaminants of nearly all articles. Contamination of foodstuffs and subsequent improper storage may result in food poisoning.
L. monocytogenes is widespread in nature and has been isolated from the environment (e.g., soil, water, sewage), animals (e.g., mammals, birds, fish, and humans), and foodstuffs (e.g., vegetables, dairy products, meats). The source of the infection in many cases is unknown, but food-borne transmission is recognized as an important factor. Although large numbers of people are exposed, most infections occur in immunocompromised individuals. It is also an occupational disease of veterinarians and abattoir workers.
E. rhusiopathiae is found in animals, fish, and birds. Transmission occurs by contact with contaminated meat or other objects. Infection is seen most often in fish handlers, meat processing workers, veterinarians, farmers, and housewives.

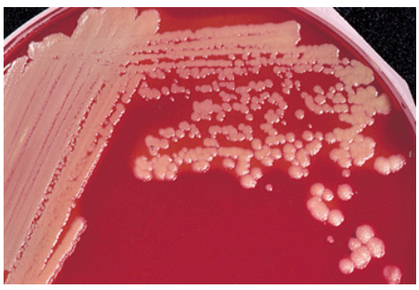


FIGURE 51.9. Colony of *Bacillus cereus* on sheep blood agar.

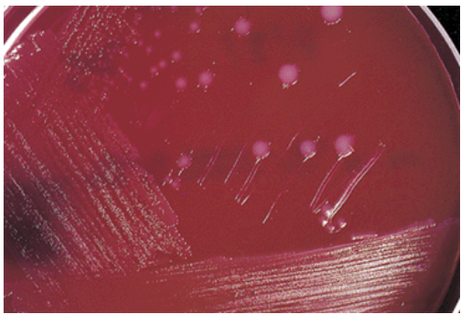


FIGURE 51.10. Colony of *Listeria* on sheep blood agar after 72 hours of incubation. Note the diffuse zone of ß hemolysis surrounding the colonies.

The upper respiratory tract and skin of humans serve as the

reservoir for *C. diphtheriae*. Transmission occurs via respiratory secretions or direct contact with exudate or contaminated fomites. The incidence of diphtheria is nearly absent in developed countries owing to immunization programs but is endemic in other underdeveloped countries. The other *Corynebacterium* species are common inhabitants of the skin and mucous membranes and are often isolated in the laboratory and dismissed as contaminants. Most infections are nosocomial or occur as the result of contact with animals or their products. *A. haemolyticum* is found in the pharynx and skin of humans.

Diagnostic Procedures

Microscopy

The identification of this heterogeneous group of organisms begins with a Gram stain of the clinical specimen. A predominant single morphology is sought, as well as correlation with the source of the specimen and the signs and symptoms of the patient (Table 51.6).

In suspected cases of diphtheria, growth from the Loeffler's slant is stained with alkaline methylene blue. The typical coryneform morphology and the presence of metachromatic granules suggest *C. diphtheriae*.

Some strains of *Bacillus* stain gram negative as they age. In addition, because they may be oxidase positive, they are sometimes confused with gram-negative rods. Spores produced by *Bacillus* appear clear with a Gram stain but must be confirmed by phase-contrast microscopy or a spore stain.

Culture Techniques

Bacillus species grow well on all noninhibitory blood-containing media and most nutrient media. They produce a variety of colony morphologies and pigments (Fig. 51.9). Heat-shock or ethanol spore selection methods followed by plating on PEA agar may be used for contaminated sites. Cultures are incubated at 35°C in air or CO_2. Strict safety precautions (e.g., biological safety cabinet, gloves) must be used for specimens and cultures suspected of containing *B. anthracis*.

Listeria grows on most laboratory media. After 24 hours of incubation on sheep blood agar, it produces a small (less than 1 mm), translucent, whitish-gray colony with a narrow zone of ß hemolysis, which is best seen after removal of the colony from the medium (Fig. 51.10). Columbia CNA agar is recommended for isolation from contaminated specimens. New, shortened enrichment procedures have been developed for the isolation of *Listeria* from food and are now recommended for tissues, feces, or environmental cultures (57). Listeria grows at 35°C in air or CO_2 but will also grow at 4°C.

E. rhusiopathiae grows on blood-containing media (e.g., blood, chocolate, CNA), but growth in infusion broth containing 1% glucose or trypticase soy broth subcultured at 2 and 7 days may be superior to direct plating. Media should be incubated at 35°C in CO_2 for 3 to 7 days.

TABLE 51.7. CHARACTERISTICS OF SELECTED GRAM-POSITIVE BACILLI

Genus	Catalase	Growth	Motility	Esculin	H_2S (TSI)	TSI s/b[a]	Nitrate Reduction	Urease	Hemolysis	Comments
Kurthia	+	A	+	-	-	-/-	-	+	N	Grows best 20–30°C
Brevibacterium	+	A	-	+	-	+/+	v	-	v	
Bacillus	+	FA	v	+	-	v/v	v	v	v	Easily decolorized spore formation
Listeria monocytogenes	+	FA	+	+	-	+/+	-	-	B	Motile at room temperature
Oerskovia	+	FA	+	+	-	+/+	v	v	N,A	Yellow pigment
Rothia	+	FA	-	+	-	+/+	+	-	N	Coccoid in broth
Mycobacterium	+	A	-	-	-	-/-	+	v	N	Acid-fast
Rhodococcus	+	FA	-	v	-	-/-	v	v	N	Pink coral pigment
Corynebacterium	+	FA	-[b]	-	-	v/v	v	v	v	
Arcanobacterium	-	FA	-	v	-	+/+	-	-	B	
Nocardia	+	FA	-	+	-	NA	v	v	N	Weakly acid-fast
Erysipelothrix	-	FA	-	-	+	+/+	-	-	N,A	
Lactobacillus	-	AN[c]	-	v	-	+/+	-	-	N,A	
Propionibacterium	+	AN[c]	-	v	-	+/+	-	-	v	
Actinomyces	-[b]	AN[c]	-	+	-	+/+	v	-	N	
Arachnia	-	AN[c]	-	-	-	+/+	-	-	N	
Clostridium	-	AN[c]	v	v	+	v/v	+	v	v	Easily decolorized spore formation

[a] s/b, slant/butt; +, acid; -, no change.

[b] Some species positive.

[c] Some species grow in CO_2.

TSI, triple-sugar iron; +, >90% of strains positive; -, <10% of strains positive; v, variable; A, strict aerobe; FA, facultative anaerobe; AN, anaerobe; B, beta: A, alpha; N, nenhemolytic.

Corynebacterium grows well on most routine enriched media (e.g., blood, CNA) and produces colonies that are translucent, opaque, white, or gray. For lipophilic corynebacterium, Tween

80 should be added to the plate (55). In cases of suspected diphtheria, selective agar (e.g., cystine-tellurite agar, modified Tinsdale agar) should be used. Loeffler's or Pai agar should be inoculated for demonstration of metachromatic granules. Cultures are incubated at 35°C in air or CO_2 for at least 72 hours.

Identification Techniques

Differential characteristics to presumptively group the different genera of gram-positive rods are depicted in Table 51.7. Additional testing is usually required for definitive identification.

Overdecolorized *Bacillus* species may be differentiated from gram-negative rods by spore formation (stimulated on esculin or triple sugar iron (TSI) agar), susceptibility to vancomycin, and resistance to KOH (potassium hydroxide) (57). Other tests useful for the identification of clinically relevant isolates are shown in Table 51.8. *Bacillus* species that produce medusa head colonies and are nonmotile, penicillin susceptible, and nitrate positive are presumptively *B. anthracis* and need to be confirmed by the CDC. The laboratory diagnosis of *B. cereus* food poisoning requires quantitative examination of the food, vomitus, or feces.

TABLE 51.8. IDENTIFICATION OF SELECTED *BACILLUS*, *BREVIBACILLUS*, AND *PAENIBACILLUS* SPECIES

Species	Width ≥ µm	Penicillin	Motility	Lecithinase	V-P	Nitrate Reduction	Starch Hydrolysis	Anaerobic Growth	Growth in 7% NaCl	Growth at 50°C
B. anthracis	+	s	-	+	+	+	+	+	+	-
B. cereus	+	R	v	+	+	+	+	+	+	-
B. mycoides	+	R	-	+	+	+	+	+	+	-
B. thuringiensis	+	R	v	+	+	+	+	+	+	-
B. megaterium	+	v	v	-	-	v	+	-	+	-
B. subtilis	-	s	+	-	+	+	+	-	+	+
B. pumilus	-	s	+	-	+	-	-	-	+	+
B. licheniformis	-	s	+	-	+	+	+	+	+	+
B. firmus	-	s	v	-	-	+	+	-	+	-
B. coagulans	v	s	+	-	v	v	+	+	-	+
B. circulans	-	NA	v	-	-	v	+	v	v	-
B. sphaericus	v	NA	+	-	-	+	-	-	v	-
B. laterosporus[a]	-	NA	+	+	-	+	-	+	-	-
B. brevis[a]	-	NA	+	-	-	v	-	-	-	-
P. polymyxa	-	NA	+	-	+	+	+	+	-	-
P. alvei	v	NA	+	-	+	+	+	+	-	-
P. macerans	-	NA	+	-	-	+	+	+	-	+

[a] Brevibacillus.

+, >90% of strains positive; -, <10% of strains positive; v, variable reactions; V-P, Voges-Proskauer reaction; NA, no data available.

The misidentification of *Listeria* as *Enterococcus*, group B streptococci, or *Corynebacterium* should not occur. All gram-positive rods or coccobacilli, especially from sterile body fluids, are screened to rule out *Listeria* by positive tests for catalase, tumbling motility at room temperature, esculin hydrolysis, and ß hemolysis on sheep blood (Table 51.7). *L. monocytogenes* should be identified to species level when isolated from food sources.

E. rhusiopathiae is nonmotile and catalase negative. It is the only gram-positive rod that produces H_2S in triple sugar iron or Klingler's iron agar slant. Additional characteristics are presented in Table 51.7.

C. diphtheriae (including other stains of *Corynebacterium, S. aureus*, and *Listeria*) produces grayish-black colonies on tellurite-containing media and a brown halo of Tinsdale medium. *C. ulcerans* and*C. pseudotuberculosis*, closely related if not actually strains of *C. diphtheriae*, also form a brown halo. All suspected isolates of *C. diphtheriae* must be identified (Table 51.9) and shown to produce diphtheria toxin by *in vivo* or *in vitro* tests. *C. diphtheriae, C. ulcerans*, and *C. pseudotuberculosis* lack pyrazinamidase activity.

TABLE 51.9. IDENTIFICATION OF MEDICALLY RELEVANT *CORYNEBACTERIUM* SPECIES

Species	Fermentation Oxidation	Lipophilism		Nitrate Reduction	Urease	Esculin Hydrolysis	Pyrazinamidase	Alkaline Phosphatase	Acid Production from Glucose	Maltose	Sucrose	Mannitol	Xylose	CAMP Reaction	Other Traits
C. accolens	F[b]	+	+	-	-	V	-	+	-	V	V -	-			
C. afermentans subsp. *afermentans*	O	-	-	-	-	+	+	-	-	-	- -	V			
C. afermentans subsp. *lipophilum*	O	+	-	-	-	+	+	-	-	-	- -	V			
C. amycolatum	F	-	V	V	-	+	+	+	V	V	- -	-	Most O/129 resistant		
C. argentoratense	F	-	-	-	-	+	V	+	-	-	- -	-	Chymotrypsin positive		
C. auris	O	-	-	-	-	+	+	-	-	-	- -	+	Slight adherence in agar, cleaved mycolics		
C. confusum	F	-	+	-	-	+	+	(+)	-	-	- -	-	Tyrosine negative		
C. coyleae	F	-	-	-	-	+	+	(+)	-	-	- -	+			
CDC group F.1	F	+	V	+	-	+	-	+	+	+	- -	-			
CDC group G	F	+	V	-	-	+	+	+	V	V	- -	-	Fructose positive, anaerobic growth positive		
C. diphtheriae	F	V	V	-	-	-	-	+	+	-	- -	-	Glycogen positive		
C. durum	F	-	+	(V)	(V)	+	-	+	+	+	V -	-	Adherence to agar		
C. falsenii	F	-	-	(+)	-	(+)	+	(+)	V	-	- -	-	Yellowish		
C. glucuronolyticum	F	-	V	V	V	+	V	+	V	+	- V	+	ß-Glucuronidase positive		
C. imitans	F	-		-	-	(+)	+	+	+	(+)	- -	+	Tyrosine negative, O/129 resistant		
C. jeikeium	O	+	-	-	-	+	+	+	V	-	- -	-	Fructose negative, anaerobic growth negative		
C. kroppenstedtii	F	+	-	-	+	+	-	+	ND	+	- -	-	Lacking mycolic acids		

C. lipoohiloflavum	O	+	-	-	-	+	+	-	-	-	-	-	-	Yellow
C. macginleyi	F	+	+	-	-	-	+	+	-	+	V	-	-	
C. matruchotii	F	-	+	-	V	+	-	+	+	+	-	-	-	"Whip handle" (on Gram staining)
C. minutissimum	F	-	-	-	-	+	+	+	+	V	V	-	-	Tyrosine positive
C. mucifaciens	O	-	-	-	-	+	+	+	-	V	-	-	-	Very mucoid colonies
C. propinquum	O	-	+	-	-	V	V	-	-	-	-	-	-	Tyrosine positive
C. pseudodiphtheriticum	O	-	+	+	-	+	V	-	-	-	-	-	-	
C. pseudotuberculosis	F	-	V	+	-	-	V	+	+	V	-	-	REV	
C. riegelii	F	-	-	+	-	V	V	-	(+)	-	-	-	-	
C. sanguinis	F	-	-	-	-	+	+	(+)	-	-	-	-	-	Yellowish
C. singulare	F	-	-	+	-	+	+	+	+	+	-	-	-	Tyrosine positive
C. striatum	F	-	+	-	-	+	+	+	-	V	-	-	V	Tyrosine positive
C. sundsvallense	F	-	-	+	-	V	V	+	+	+	-	-	-	Sticky colonies
C. thomssenii	F	-	-	+	-	+	+	+	+	+	-	-	-	*N*-Acetyl-ß glucosaminidase positive, sticky colonies
C. ulcerans	F	-	-	+	-	-	+	+	+	-	-	-	REV	Glycogen positive
C. urealyticum	O	+	-	+	-	+	V	-	-	-	-	-	-	
C. xerosis	F	-	V	-	-	+	+	+	+	+	-	-	-	O/129 susceptible

V, variable; parenthese, delayed or weak reaction; ND, no data; REV, reverse cAMP reaction; F, fermentation; O, oxidation; +, positive reacton; -, negative reaction.From Funke G, Bunard KA, Coryneform Gram-positive rods. In: Murray PR, Baron EJ, Pfaller MA, et al. eds. *Manual of clinical microbiology*, eds. Washington, DC: American Society for Microbiology, 1999: 328–329, with permission.

The identification of clinically significant *Corynebacterium* species (Table 51.9) can be difficult because of the heterogeneity and biochemical diversity of the group (55,56,58). *C. jeikeium* exhibits poor growth in routine biochemical tests and multiple resistance to most antimicrobial agents except vancomycin. Various

commercial identification kits have been used to identify some of the *Corynebacterium* species (55).

Direct Antigen or Nucleic Acid Detection

B. anthracis can be identified directly in specimens by fluorescent antibody stain at the CDC. Various commercial immunoassays and a DNA probe are available for the rapid detection of *Listeria* in food but these tests cannot be used on clinical specimens. A DNA probe (GenProbe) can be used for culture confirmation. Fluorescent antibody techniques have also been described for *C. diphtheriae*. Nucleic acid detection is not commercially available for these organisms.

Serologic Procedures

B. anthracis can be diagnosed by an indirect hemagglutination procedure performed at the CDC. Serology for listeriolysin (performed only in some reference or research laboratories) may be useful for diagnosis of systemic listeriosis and gastroenteritis (59,60).

Interpretation of Results

The isolation of *Listeria, Erysipelothrix, C. diphtheriae*, or *B. anthracis* is clinically relevant. The significance of any of the other numerous species of gram-positive rods depends on the type and quality of the specimen, the frequency of isolation, and the nature of the infectious disease process.

Therapeutic Considerations

The drug of choice for the treatment of anthrax is penicillin. The other *Bacillus* species demonstrate variable susceptibility to the penicillins and cephalosporins, but most appear susceptible to aminoglycosides, tetracycline, clindamycin, erythromycin, chloramphenicol, vancomycin, imipenem, and ciprofloxacin.

Most of the recommended therapy for listeriosis is based on isolated case reports. Penicillin or ampicillin plus an aminoglycoside administered over a 3- to 6-week period is probably the optimal combination. Trimethoprim-sulfamethoxazole is a second choice.

Erysipelothrix is susceptible to penicillin, clindamycin, erythromycin, fluoroquinolones, imipenem, tetracycline, and cephalosporins. Isolates are usually resistant to the aminoglycosides, vancomycin, and sulfonamides.

Patients with diphtheria should be given antitoxin and started on erythromycin. Clindamycin or rifampin may also be used. Susceptibility of the other *Corynebacterium* species must be determined by testing. *C. jeikeium* is susceptible to vancomycin and resistant to most other antimicrobial agents. *A. haemolyticum* is susceptible to the ß-lactams.

Summary and Conclusions

Gram-positive, aerobic to facultatively anaerobic rods describe a heterogeneous group of organisms containing more than 20 genera. Most of these organisms are seldom if ever isolated from clinical specimens and are found in the environment (e.g., soil, food) or on the skin and mucous membranes of humans and animals. Because of taxonomic and identification difficulties, many of these species are just now being recognized as important agents of opportunistic infections. *B. anthracis, L. monocytogenes, E. rhusiopathiae*, and *C. diphtheriae* should always be considered clinically significant when isolated.

In general, most of these bacteria are readily isolated on laboratory media containing blood but are difficult to identify definitively. Susceptibility to antimicrobial agents is variable and must be determined for many of these species.

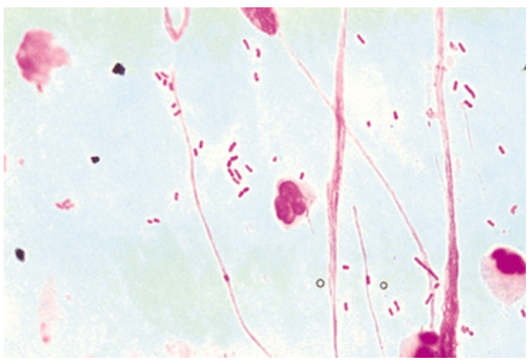


FIGURE 51.11. Gram stain of *Klebsiella pneumoniae* in sputum.

ENTEROBACTERIACEAE

Part of "51 - Bacteriology"

Microbiology

The family Enterobacteriaceae encompasses 30 genera and more than 100 species, CDC groups, and biogroups (Table 51.10).

Many of the new species are rarely or never found in human specimens. Others are isolated from specimens but seldom cause disease. Currently, two *Salmonella* species are recognized, *S. enterica* (six subspecies) and *S. bongori* that include all of the subspecies and serotypes (61). Regardless of these changes most clinical microbiology laboratories will continue to report the genus and serotype (e.g., *Salmonella* serogroup B).

TABLE 51.10. NOMENCLATURE AND CLASSIFICATION OF THE FAMILY ENTEROBACTERIACEAE

Genus	No. of Species
Budvicia	1
Buttiauxella	7
Cedecea	5
Citrobacter	11
Edwardsiella	4
Enterobacter	15
Escherichia	5
Ewingella	1
Hafnia	1
Klebsiella	7
Kluyvera	4
Leclercia	1
Leminorella	2
Moellerella	1
Morganella	1
Obesumbacterium	1
Pragia	1
Pantoea	1
Photohalodus	2
Proteus	4
Providencia	5
Rahnella	1
Salmonella	7 subgroups
Serratia	10
Shigella	4
Tatumella	1
Trabuisiella	1
Xenorhabdus	1
Yersinia	11
Yokenella	1

The members of this family are generally straight, gram-negative rods that are catalase positive and oxidase negative, reduce nitrate to nitrite, and are facultatively anaerobic. They are motile by peritrichous flagella or nonmotile, do not form spores, and usually grow on MacConkey agar. They demonstrate a vast heterogeneity in their habitats and pathogenic potential for plants, animals, and humans.

Spectrum of Disease

The members of the Enterobacteriaceae are among the most important bacteria medically and have been associated with infections in every major organ system. Their ubiquitous presence in the environment, in foodstuffs, and on animals and humans; their ability to develop antimicrobial resistance; and the increasing numbers of susceptible patients who are compromised or have prosthetic implants have ensured the involvement of these bacteria in both community- and hospital-acquired infections. Some genera and species such as *Salmonella, Shigella, Yersinia*, and some strains of *Escherichia coli* are primary enteric pathogens causing dysentery, gastroenteritis, enteric fever, hemolytic anemic syndrome, and hemorrhagic colitis. The diarrheic *E. coli* is comprised of at least four groups; the Shiga toxin producing *E. coli* or enterohemorrhagic *E. coli*, enterotoxigenic *E. coli*, enteropathogenic *E. coli*, and enteroinvasive *E. coli* based on clinical features and pathogenic mechanisms (61).

Epidemiology

The members of the family Enterobacteriaceae are widely distributed in soil and water, on plants, and in food and colonize the GI tract of animals and humans. They also contaminate many types of medical supplies and devices. Community-acquired GI disease occurs via consumption of contaminated water or foodstuffs. Bacteria causing nosocomial infections (usually found in the GI or upper respiratory tract of the patient) are transmitted on the hands of hospital personnel or introduced into the patient via contaminated solutions or devices. Numerous techniques are available to track specific strains in food-related outbreaks or in the hospital environment, including serologic typing, biotyping, bacteriocins, antibiograms, and nucleic acid analysis.

Diagnostic Procedures

Microscopy

On Gram stain, the “enterics” exhibit a characteristic morphology consisting of straight, thick, barrel-shaped rods that may exhibit bipolar staining and are often pleomorphic (Fig. 51.11). *Yersinia pestis* may have a “safety pin” appearance.

Culture Techniques

All members of this group grow luxuriantly within 24 hours in air at 35°C on routine laboratory media such as blood, MacConkey (or eosin methylene blue), or chocolate agar. On blood agar, colonies are usually 1 to 3 mm in diameter, dull, gray, opaque, sometimes mucoid (e.g., *Klebsiella*), sometimes hemolytic (e.g., *E. coli*), and sometimes swarming (e.g., *Proteus*). More selective agars, as well as enrichment broths, are used to isolate enteric pathogens from feces. The characteristic appearances of colonies on these media are listed in Table 51.11. Newer plating media for the isolation of *Salmonella* are available (62).

Sorbitol-MacConkey and cefsulodin-irgasan-novobiocin agars are used to isolate *E. coli* O157 and *Y. enterocolica*, respectively. Other selective media for the isolation of *E. coli* 0157 strains are available (61,63,64). It is recommended that media for the isolation of *Yersinia* be incubated at 22 to 25°C for 48 hours.

TABLE 51.11. APPEARANCE OF COMMONLY ISOLATED ENTEROBACTERIACEAE ON VARIOUS ENTERIC MEDIA

	MacConkey Agar	Eosin-Methylene Blue Agar	Hektoen Enteric Agar	Xylose-Lyslne Deoxycholate Agar	Salmonella-Shigella Agar	Deoxycholate Citrate Agar	Bismuth Sulfite Agar	Brilliant Green Agar
Escherichia coli								
Lac +	Flat; red or dark pink; surrounded by zone of precipitated bile	Red black with metallic sheen[a]	Yellow-orange	Yellow	Pink	Deep red-pink	Usually do not grow	Usually do not gorw
Lac -	Colorless	Colorless	Yellow-orange or green	Yellow	Colorless	Colorless	Usually do not grow	Usually do not grow
Klebsiella	Pink; mucoid	Purple	Yellow-orange	Yellow	Pink	Pink	Usually do not grow	Usually do not grow
Enterobacter	Pink; not usually as mucoid as *Klebsiella*	Purple	Yellow-Orange	Yellow	Pink	Pink	Usually do not grow	Usually do not grow
Citrobacter, Serratia, Hafnia, Providencia,	May appear colorless after 24 hr or slightly pink in 24–48 hr	Lavender or colorless	Colorless	Red, yellow, or colorless with or without black centers	Colorless	Colorless	Usually do not grow	Usually do not grow
Proteus, Morganella, Edwardsiella	Colorless[b]	Colorless	Colorless	Red, yellow, or colorless with or without black centers	Colorless	Colorless	Usually do not grow	Usually do not grow
Salmonella	Colorless	Colorless	Green or blue-green	Pink to red with black center	Colorless with black center	Colorless	Green-black	Pink white opaque: surrounded by brilliant red medium
Shigella	Colorless	Colorless	Green or blue-green	Colorless	Colorless	Colorless	Usually do not grow	Usually do not grow
Yersinia	Colorless to peach	Colorless or purple[c]	Salmon	Yellow	Colorless	Colorless	Usually do not grow	Usually do not grow

[a] Not all strains produce a metallic sheen; conversely, other species of entetic bacilli (e.g., *Yersinia enterocolitica*) may produce a sheen.

[b] *Proteus mirabilis, Proteus vulgaris*, and *Proteus penneri* may swarm.

[c] *Yersinia enterocolitica*, a nonlactose fermenter that ferments sucrose, produces colorless colonies on Levine EMB agar and purple colonies on the modified Holt-Harris Teague formula, which contains sucrose.

From Howard BJ, Klass J II, Rubin SI, et al. eds. *Clinical and pathogenic microbiology*. St. Louis: Mosby, 1987:295, with permission.

Identification Techniques

With respect to identification, the clinical laboratory has three levels of need: (a) a screen for enteric pathogens, (b) a presumptive identification of isolates from significant but not life-threatening infections, and (c) definitive identification.

Two tests (urea and triple sugar iron or lysine iron agar) have been used to screen stool cultures for potential pathogens (Table 51.12). The urease test aids in the detection of *Proteus* and urease-positive *Providencia*. Commercial latex agglutination kits are available to identify colonies of *E. coli* 0157, *Salmonella*, and *Shigella* or to detect their growth in enrichment broth.

TABLE 51.12. TYPICAL REACTION PATTERNS OF ENTERIC PATHOGENS ON TRIPLE-SUGAR IRON AND UREA AGAR

Genus	Urea	TSI (Slant/Butt)	H_2S	Oxidase
Salmonella	–	K/A	+ or -	–
Salmonella (rare)	–	A/A	+	–
Shigella	–	K/A	–	–
Edwardsiella	–	K/A	+	–
Yersinia	+	A(K)/A	–	–
Aeromonas	–	A(K)/A	–	+
Plesiomonas	–	A(K)/A	–	+

TSI, triple-sugar iron; A, acid; K, alkaline; +, positive; -, negative; (), occasional reactions.

Clinically relevant isolates from nonlife-threatening infections can be presumptively identified by colony morphology and a few rapid tests. The identification of the most frequently isolated bacteria is based on lactose fermentation and colony morphology on MacConkey agar, swarming on blood agar, spot indole, motility, and ornithine decarboxylase (Table 51.13). Members of the Enterobacteriaceae family are definitely identified by inoculation of numerous conventional biochemical tests, use of commercial manual kits, or by automated systems. All *Salmonella, Shigella*, and *E. coli* 0157 isolates should be identified both biochemically and serologically. Biochemicals for the identification of *Yersinia* should be incubated at 22 to 25°C.

TABLE 51.13. PRESUMPTIVE IDENTIFICATION OF COMMON ISOLATES OF ENTEROBACTERIACEAE

Characteristic	*Escherichia coli*	*Klebsiella*	*Enterobacter*	*Proteus mirabilis*	*Proteus vulgaris*
Colony on MacConkey agar					
Flat, dry	+	NA	NA	NA	NA
Mucoid	NA	+	+	NA	NA
Lactose fermenting	+	+	+	–	–
Swarming on blood agar	NA	NA	NA	+	+
Spot indole	+	–	–	–	+
Ornithine decarboxylase	NA	–	+	+	–
Motility	NA	–	+	NA	NA

+, positive; -, negative; NA, not applicable.

Direct Antigen or Nucleic Acid Techniques

Currently, there are no commercially available systems for the direct detection of Enterobacteriaceae in clinical specimens. However, several tests are available for the identification of colonies or the detection of virulence factors. As mentioned previously, *Salmonella, Shigella*, and *E. coli* 0157 can be identified by latex agglutination or serotyped by antisera. Also, colonies of *E. coli* can be screened for production of toxins by EIA. Nucleic acid probes are used to identify enterotoxigenic, enteroinvasive and enteropathogenic *E. coli*. At present, EIA and probes are used primarily for epidemiologic studies by reference centers. Both DNA hybridization and monoclonal antibody systems are being evaluated for the detection of *Salmonella* in food. Immunoassays for the direct detection of *E. coli* 0157 (65,66) or Shiga toxin (61) in stool are commercially available.

Serologic Procedures

The Widal test is used to help diagnose *S. typhi* infections by measuring an antibody response to the O and H antigens. However, numerous factors can produce both false-positive and false-negative reactions, and the test should not be used alone to diagnose typhoid fever.

Serodiagnosis of *Y. enterocolitica* is useful only when one or two serogroups cause the majority of the infections. Serodiagnosis is difficult in the United States because numerous O serogroups cause enteric infections. The criteria for positive and negative tests are also poorly defined.

Interpretation of Results

The isolation of *Salmonella, Shigella*, or *Yersinia* from any site is clinically relevant. The isolation of any member of the Enterobacteriaceae from a normally sterile body fluid or site is significant until proven otherwise. The significance of recovery of these organisms from other specimens (e.g., sputum, urine) must be interpreted considering the patient's underlying condition, prosthesis (including catheters), intubation, type of specimen, and frequency of isolation.

Therapeutic Considerations

Antimicrobial resistance has always been and continues to be a problem with the members of the Enterobacteriaceae. The emerging resistance to the newer cephalosporins and fluoroquinolones may pose problems for adequate therapy. Because antimicrobial resistance is unpredictable and varies within geographic areas, all clinically relevant isolates should be tested. However, *Salmonella* isolates from uncomplicated gastroenteritis are an exception because antimicrobial therapy is not recommended.

Summary and Conclusions

The members of the family Enterobacteriaceae are ubiquitous in nature, causing a significant number of GI, urinary tract, and nosocomial infections. They are readily isolated on routine media and are identified by numerous manual and commercial methods. Antimicrobial therapy must be based on susceptibility testing of the isolated bacterium.

AEROMONAS AND PLESIOMONAS

Part of "51 - Bacteriology"

Microbiology

Aeromonas and Plesiomonas are included in the family Vibrionaceae based primarily on polar flagella and the oxidase reaction. It has been proposed that aeromonads be placed in the family Aeromonadaceae, whereas *Plesiomonas* probably is related to the genus *Proteus* (67). Both are gram-negative, motile, catalase- and oxidase-positive rods that are facultatively anaerobes. The genus *Aeromonas* now contains 16 species: *A. hydrophila, A. caviae, A. sobria, A. media, A. veronii, A. schubertii, A. eucrenophila, A. salmonicida, A. bestiarum, A. jandaei, A. encheleia, A. trota, A. allosaccharophila, A. popoffii*, and two unnamed species.

Spectrum of Disease

Aeromonas causes four types of infections in humans, including diarrhea, cellulitis or wound infections (usually after exposure to water or soil), bacteremia (primarily in the immunocompromised host), and a variety of soft-tissue and deep-seated infections (68,69). Only three species, *A. hydrophila, A. caviae*, and *A. sobria* are commonly associated with clinical disease. *Plesiomonas shigelloides* primarily causes GI disease but has been associated with cases of meningitis, bacteremia, wound infections, septic arthritis, and cholecystitis (70).

Epidemiology

Both *Aeromonas* and *Plesiomonas* are found in soil and water, where they cause infections in frogs, reptiles, snakes, birds, and fish. Transmission to humans probably occurs via contaminated water or foodstuffs or contamination of an injury. Transmission via water or food is rarely documented.

Diagnostic Techniques

Microscopy

There are no available microscopic techniques to detect these bacteria directly in clinical specimens. The morphology on Gram smear is not specific for either genus.

Culture Techniques

Both genera grow well on blood and MacConkey agar and most enteric screening agars. Strains of *A. hydrophila* and *A. sobria* may be ß-hemolytic on blood agar. Selective agar is available for recovery from stool specimens.

Identification Methods

A motile, oxidase- and indole-positive organism can be tentatively identified as *Aeromonas*/*Plesiomonas*/*Vibrio*. Vibrios grow in 6% NaCl, whereas *Aeromonas* and *Plesiomonas* do not. Commercial systems will generally identify *Aeromonas* and *Plesiomonas* to the genus level. Key characteristics are listed in Table 51.14. There is no commercial test for the direct detection of antigen or nucleic acid in clinical specimens. Tests are not available for serologic diagnosis.

TABLE 51.14. DIFFERENTIAL CHARACTERISTICS OF AEROMONAS AND PLESIOMONAS SPECIES

Organism	DNase	Esculin Hydrolysis	Acid from Mannitol	Voges-Proskauer	Arginine Dehydrolase	Lysine Decarboxylase	Ornithine Decarboxylase
A. hydrophilia	+	+	+	+	+	+	-
A. caviae	+	+	+	-	+	-	-
A. sobria	+	-	+	+	+	+	-
A. schubertii	+	-	-	+	+	+	-
A. veronii	+	+	+	+	-	+	+
A. jandaei	ND	-	+	+	+	+	-
A. trota	ND	-	+	-	+	+	-
P. shigelloides	-	-	-	-	+	+	+

+, positive; -, negative; ND, not determined.

Interpretation of Results

The isolation of either *Aeromonas* or *Plesiomonas* from specimens other than stool should be considered significant. Isolation from stool is significant in children. In adults, only *A. hydrophila* and *A. veronii* seem to be significant.

Therapeutic Considerations

Aeromonas is resistant to ampicillin and other penicillins, including some of the ureidopenicillins. Isolates are generally susceptible to the third-generation cephalosporins, quinolones, trimethoprim-sulfamethoxazole, and the aminoglycosides. The susceptibility of *Plesiomonas* has not been evaluated extensively, but isolates appear susceptible to most classes of antimicrobics except aminoglycosides and tetracycline. Because both genera can easily acquire resistance to antimicrobial agents, therapy should be based on the results of *in vitro* susceptibility tests.

Summary and Conclusions

The genera *Aeromonas* and *Plesiomonas* are common inhabitants of the aquatic environment. Infections are seen in both normal and immunocompromised hosts, usually after exposure to or consumption of contaminated water or foodstuffs. Therapy is based on the result of *in vitro* susceptibility testing.

VIBRIO

Part of "51 - Bacteriology"

Microbiology

The general *Vibrio, Photobacterium, Aeromonas*, and *Plesiomonas* are in the family Vibrionaceae. Two genera (*Listonella* and *Shewanella*) have been proposed for inclusion in this family. Although more than 35 *Vibrio* species are recognized, only 12 have been associated with infections in humans. In general, *Vibrio* species share the following characteristics: gram-negative, curved or straight rod, oxidase positive, reduce nitrate to nitrite, sodium required for or stimulates growth, susceptible to vibriostatic compound 0/129, motile by polar flagella, and facultatively anaerobic. *Photobacterium* is not associated with human infections.

Spectrum of Disease

V. cholerae serogroup 01 is the etiologic agent of epidemic cholera. The other pathogenic species are usually associated with food-borne gastroenteritis, soft-tissue infections or systemic infections such as meningitis, septicemia, cholecystitis, cellulitis, and a variety of other wound infections. *V. parahaemolyticus* gastroenteritis is the most frequently encountered noncholera *Vibrio* infection. *V. vulnificus* is an especially virulent species associated with primary sepsis (often in patients with preexisting liver disease) or with serious soft-tissue infections in normal hosts.

Epidemiology

Vibrio species inhabit aquatic environments worldwide. Although found primarily in marine ecosystems, some species can live in fresh water. Disease is transmitted via contaminated food (especially shellfish) or water and exposure of wounds to contaminated water. In epidemic cholera, food and water are usually contaminated by feces of patients and carriers. Gastroenteritis from the noncholera vibrio occurs in sporadic or common source outbreaks after consumption of raw, improperly prepared, or recontaminated seafood. Wounds are infected from exposure to contaminated water. Most infections occur in coastal locations, but more inland cases are being reported because of travel and shipment of contaminated food.

Diagnostic Procedures

Microscopy

Direct examination of stools by wet mount for rapid, darting motility or by Gram smear for gram-negative curved rods may be helpful in an outbreak situation but is not useful for identification of a sporadic case.

Culture Techniques

Routine primary isolation media such as blood and MacConkey agar are adequate for recovery of vibrios from extraintestinal specimens. Stool specimens are inoculated to thiosulfate-citrate-bile salts-sucrose (TCBS) agar. Some pathogenic species may grow poorly or not at all on this medium (71). All cultures should be incubated at 35°C in air or CO_2 where most species form 1 to 2 mm colonies after 24 hours. Alkaline peptone broth can be used for the enrichment of fecal specimens.

Vibrio species may exhibit ß hemolysis on blood agar, are nonlactose fermenters (except *V. vulnificus*) on MacConkey agar, and are differentiated by sucrose fermentation on TCBS. Nonsucrose-fermenting strains produce green colonies, whereas sucrose-fermenting strains form yellow colonies. Oxidase testing should be done from nonselective media.

TABLE 51.15. CHARACTERISTICS USEFUL FOR DIFFERENTIATING *VIBRIO* FROM OTHER OXIDASE-POSITIVE GENERA

Characteristic	*Vibrio*	*Pseudomonas*	*Aeromonas*	*Plesiomonas*
Growth on TCBS	+	-	-	-
Oxidase	+[a]	+	+	+
Fermentative	+	-	+	+
Growth stimulated by NaCl	+	-	-	-
Susceptibility to 0/129 (150 µg)	+	-	-	v
Susceptibility to ampicillin (10 µg)	+	-	-	-

TCBS, thiosulfate-citrate-bile salts; +, positive; -, negative; v, variable.

[a] *V. metschnikovii* is oxidase negative.

Identification Techniques

Initially, *Vibrio* species are separated from other genera by the oxidase reaction, fermentative metabolism, requirement of NaCl for growth, and susceptibility to 2,4-diamino-6, 7-diisopropylpteridine (0/129) (Table 51.15). Identification can be made

using routine biochemical tests if the NaCl concentration is increased to 1% (wt/vol) (Table 51.16). Commercial kits identify the more common *Vibrio* species (if the inoculum is prepared in saline) but cannot be used to identify the less common isolates. However, misidentification of *Aeromonas* as *Vibrio* has occurred (72).

TABLE 51.16. TESTS FOR DIFFERENTIATION OF MEMBERS OF THE VIBRIONACEAE FROM HUMANS

Tests	*V. cholerae* 01 and non-0	*V. mimicus*	*V. parahaemolyticus*	*V. vulnificus*	*V. alginolyticus*	*V. cincinnatiensis*	*V. fluvialis*	*V. furnissii*	*V. damseta*	*V. hollisae*	*V. metschnflovii*	*Aeromonas* Spp	*Plesiomonas* Spp
Oxidase	+	+	+	+	+	+	+	+	+	+	-	+	+
NO_3-NO_2 + 1% NaCl	+	+	+	+	+	+	+	+	+	+	-	+	+
Indole + 1% NaCl	+	+	+	+	+/-	-	-/+	-	-/+	+	-/+	+/-	+
Voges-Proskauer + 1% NaCl	+/-	-	-	-	+	-	-	-	+	-	+	+/-	-
Urease	-	-	+/-	-	-	-	-	-	-	-	-	-	-
Lysine decarboxylase + 1% NaCl	+	+	+	+	+	+/-	-	-	+/-	-	-/+	-/+	+
Ornithine decarboxylase + 1% NaCl	+	+	+	+/-	+/-	-	-	-	-	-	-	-	+
Arginine dihydrolase + 1% NaCl	-	-	-	-	-	-	+	+	+	-	+/-	+/-	+
Fermentation of													
Sucrose	+	-	-	-/+	+	+	+	+	-	-	+	+/-	-
Lactose	-	-/+	-	+/-	-	-	-	-	-	-	+/-	-/+	-/+
L-Arabinose	-	-	+/-	-	-	+	+	+	-	+	-	+/-	-
Gas from glucose	-	-	-	-	-	-	-	+	-/+	-	-	-/+	-
Growth in nutrient broth													
0% NaCl	+	+	-	-	-	-	-	-	-	-	-	+	+
3% NaCl	+	+	+	+	+	+	+	+	+	+	+	+	+
6% NaCl	+/-	+/-	+	+/-	+	+	+/-	+/-	+	+/-	+	-	-
8% NaCl	-	-	+	-	+	-	-	-	-	-	-	-	-
10% NaCl	-	-	-	-	+/-	-	-	-	-	-	-	-	-
Susceptibility to 0/129													
10 µg	S	S	R	S	R	R	R	R	S	R	S	R	RS
150 µg	S	S	S	S	S	S	S	S	S	S	S	R	S
Growth on TCBS	Y	G	G	G/Y	Y	Y	Y	Y	G	G/-	Y	-	-

+ most strains positive; -, most strains negative; +/- or -/+, variable reaction (predominant reaction shown as the numerator; (), delayed reaction; S, susceptible; R, resistant, Y, yellow colonies; G, green colonies; TCBS, thiosulfate-citrate-bile salts.

From Janda JM, Powers C, Bryant RG, et al. Current perspectives on the epidemiology and pathogenesis of clinically significant *Vibrio* spp. *Clin Microbiol Rev* 1988;1;259, with permission.

Direct Antigen or Nucleic Acid Detection

Nucleic acid probes have been used in epidemiologic studies but are not commercially available and have not been evaluated as a direct diagnostic test. However, immunoassays for the direct detection of *V. cholerae* 01 in stool specimens are commercially available (71).

Serologic Procedures

V. cholerae is identified as serogroup 01 or non-01 based on somatic antigens tested by a slide agglutination procedure. A retrospective diagnosis of cholera can be made by a serologic test if the patient has not been vaccinated for cholera.

Interpretation of Results

The isolation of *V. cholerae* is a reportable event because of the potential public health consequences. The isolation of other *Vibrio* species is clinically relevant, especially in the case of *V. vulnificus*, which has devastating consequences for the patient.

Therapeutic Considerations

Fluid replacement is the primary treatment for epidemic cholera and is sometimes necessary in gastroenteritis caused by other *Vibrio* species. Antimicrobial agents are generally not necessary.
Soft-tissue or systemic infections are usually treated with doxycycline ± ceftazidime. Most strains are also susceptible to chloramphenicol, aminoglycosides, and fluoroquinolones. Antimicrobial susceptibility testing can be performed without added NaCl.

Summary and Conclusions

V. cholerae, V. parahaemolyticus, and *V. vulnificus* are the most clinically important species of the pathogenic vibrios. All vibrios inhabit an aquatic environment, primarily a marine ecosystem. Cholera is of worldwide importance, whereas other noncholera vibrios cause significant gastroenteritis and soft-tissue and systemic infections. *Vibrio* species are transmitted to humans through fecally contaminated food and water (*V. cholerae*), contaminated shellfish, or exposure to contaminated water. These bacteria are easily isolated on routine media, but TCBS agar is recommended for stool specimens. Identification is performed by routine biochemical tests containing 1% NaCl or in some cases by commercial kits. The key to the isolation of *Vibrio* is recognition of the colony on the primary isolation plate.

CAMPYLOBACTER, *ARCOBACTER*, AND *HELICOBACTER*

Part of "51 - Bacteriology"

Microbiology

The taxonomy and nomenclature of the genus *Campylobacter* is changing rapidly as old species are renamed and new species identified. Currently, two genera, *Campylobacter* and *Arcobacter*, are included in the family Campylobacteraceae, which includes 18 species and subspecies in the genus *Campylobacter* and four species in the genus *Arcobacter* (73). Two *Campylobacter* species were reclassified as *Arcobacter*, two *Wolinella* species were reclassified in the genus *Campylobacter*, and two *Campylobacter* species were placed in the genus *Helicobacter*. The genus *Helicobacter* contains 19 species, seven of which are associated with humans.
Campylobacter species are slender, curved, gram-negative rods (0.5 to 5 µm in length) that may have a comma, S, gull-wing, or spiral shape and exhibit a characteristic darting motility (polar flagella) on wet mount. They are oxidase positive, microaerophilic, and capnophilic (5% O_2, 10% CO_2, and 85% N_2) and are inactive toward carbohydrates. *Arcobacter* species are similar except that they show variable growth at 37° and 42° C. *Helicobacter* species are spiral or curved, motile, gram-negative rods (1.5 to 10 µm in length). They are oxidase positive, inactive toward carbohydrates, microaerophilic, and require a humid atmosphere with reduced oxygen (5% to 10%) and increased carbon dioxide (5% to 12%).

Spectrum of Disease

Campylobacter species are associated with two categories of disease: GI and extraintestinal. *C. fetus* subsp *fetus* typically causes extraintestinal infections such as bacteremia, cholecystitis, meningitis, septic abortion, and septic arthritis. Infection is often associated with underlying disorders such as hepatic disease, immunodeficiency, alcoholism, and diabetes. *Arcobacter* and other *Campylobacter* species less commonly cause infection. *Helicobactor pylori* is associated with ulcers of the gastric or duodenal mucosa and cancer of the GI tract. *C. jejuni* and *C. coli* are the two species most commonly associated with gastroenteritis in the United States. Extraintestinal infection is most commonly bacteremia but can include arthritis, meningitis, peritonitis, among other diseases (74). *C. jejuni* is an antecedent cause of Guillain-Barré syndrome.

Epidemiology

Campylobacter species are found in the GI tract of wild and domesticated animal species, including cattle, sheep, swine, goats, dogs, cats, rodents, and all fowl. This results in the contamination of food (especially meat), milk, and water that is consumed by humans. Transmission also occurs from direct contact with infected pets (dogs and cats), person-to-person contact, or the sexual practices of homosexual men. Cases occur year-round but with a peak incidence in the summer and fall. *Helicobacter* inhabit the stomach or the lower GI tract of humans. Person-to-person transmission occurs via oral-oral or fecal-oral routes, although transmission from dogs or cats may occur also. To date, *H. pylori* has not been recovered from the environment.

Diagnostic Techniques

Microscopy

A presumptive diagnosis of *Campylobacter* enteritis is made by the examination of a stool specimen by Gram stain or wet mount using dark-field or phase-contrast microscopy. *Campylobacter* species will have a very rapid, darting motility on wet mount or have a slender, curved morphology on Gram stain (Fig. 51.12). Safranin should remain on the smear for 2 to 3 minutes to enhance staining or carbolfuchsin should be used as the counterstain. *Helicobacter* species can be visualized in tissue collected at endoscopy and stained by hematoxylin and eosin or Giemsa stain.

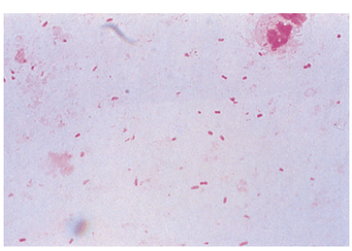


FIGURE 51.12. Gram stain of *Campylobacter jejuni* in stool. Note the curved and S-shaped forms.

Culture Techniques

Campylobacter and *Arcobacter* species grow on most blood-containing media (e.g., blood or chocolate agar). Isolation from stool usually requires direct inoculation to a selective agar or filtration (0.45 to 0.65 µm) and then inoculation to nonselective medium. The latter method is optimal for the recovery of *C. jejuni* and atypical enteric campylobacters. However, selective media containing cephalothin should not be used. Enrichment broths are not necessary for isolation from stool specimens. Most *Campylobacter* species grow at 35° to 37°C, but *C. jejuni, C. coli*, and *C. lari* will also grow at 42°C. Colonies are 1 to 2 mm in diameter, smooth, convex, and translucent after 24 to 48 hours. Enteric *Campylobacter* species produce two colony types: round and raised or flat and watery with an irregular edge. Cultures for *C. jejuni* are routinely incubated for 72 hours but may be held as long as 5 days for the uncommon *Campylobacter*. Microaerophilic conditions (5% O_2, 10% CO_2, 85% N_2) are required

for growth. Some *Campylobacter* species require hydrogen for initial isolation. *H. pylori* and other *Helicobacter* species grow best on blood (especially horse blood)-containing media in a moist atmosphere containing 5% O_2, 10% CO_2, and 85% N_2 (75). Some species require increased hydrogen. Both selective and nonselective media should be used for optimal recovery. Enteric *Helicobacter* species can be cultivated from feces on the media used for *Campylobacter*. Plates need to be incubated for as long as 7 days.

Identification Techniques

Campylobacter species from stool are presumptively identified by their typical cellular morphology, oxidase-positive reaction, and a rapid, darting motility in broth. Alternatively, suspect colonies can be tested by using a commercial immunoassay or a DNA probe. *H. pylori* is presumptively identified by demonstrating urease, catalase, and oxidase reactions.

Conventional biochemical and physiologic tests and, in some cases, molecular methods are necessary for the definitive identification of *Campylobacter* species (Table 51.17). Guidelines for setting up these tests can be found in reference 76. Phenotypic tests do not reliably differentiate between *Campylobacter*, *Arcobacter*, and enteric *Helicobacter* species.

TABLE 51.17. PHENOTYPIC PROPERTIES OF *CAMPYLOBACTER*, *ARCOBACTER*, AND *HELICOBACTER* SPECIES

									Growth at			Growth in or on			Susceptibility to	
Organism	**Catalase**	**Nitrate Reduction**	**Nitrite Reduction**	**H_2 Required**	**Urease**	**H_2S (TSI)**	**Hippurate Hydrolysis**	**Indoxyl Acetate Hydrolysis**	**15°C**	**25°C**	**42°C**	**3.5% NaCl**	**1% Glycine**	**MacConkey Agar**	**Nalidixic Acid**	**Cephalothin**
C. jejuni	+[a]	+	-	-	-	-	+	+	-	-	+	-	+	+	V	R
C. jejuni subsp *doylei*	V	-	-	-	-	-	V	+	-	-	-	-	+	-	S	S
C. coli	+	+	-	-	-	-	-	+	-	-	+	-	+	+	S	R
C. fetus	+	+	-	-	-	-	-	-	-	+	-	-	+	+	V	S
C. fetus subsp *venerealis*	+	+	-	-	-	-	-	-	-	+	-	-	-	+	R	S
C. lari	+	+	-	-	V	-	-	-	-	-	+	-	+	+	R	R
C. upsaliensis	W	+	-	-	-	-	-	+	-	-	+	-	V	-	S	S
C. hyointestinalis	+	+	-	V	-	+	-	-	-	+	+	-	+	+	R	S
C. sputorum biovar sputorum	-	+	+	-	-	+	-	-	-	-	+	-	+	+	S	S
C. sputorum biovar bubulus	-	+	+	-	-	+	-	-	-	-	+	+	+	-	R	S
C. sputorum biovar fecalis	+	+	+	-	-	+	-	-	-	-	+	-	+	+	R	S
C. helveticus	-	+	ND	-	ND	-	-	+	-	-	+	V	V	ND	S	S
C. mucosalis	-	+	+	+	-	+	-	-	-	-	+	-	+	+	R	S
C. concisus	-	+	+	+	-	+	-	-	-	-	+	-	+	+	R	R
C. curvus	-	+	+	+	-	+	-	+	-	-	+	-	+	ND	S	ND
C. rectus	-	+	+	+	-	+	-	+	-	-	W	-	+	ND	S	ND
C. showae	+	+	ND	+	-	+	-	+	-	-	+	-	V	ND	R	S
A. cryaerophilus group 1A	+	V	-	-	-	-	-	+	+	+	-	-	-	-	V	R
A. cryaerophilus group 1B	+	V	ND	-	-	-	-	+	+	+	-	-	-	+	S	V
A. butzleri	W	+	-	-	-	-	-	+	+	+	V	V	+	+	S	R
A. nitrofigilis	+	+	-	-	-	-	-	-	+	+	-	+	-	-	S	S
A. skirrowii	+	+	ND	-	-	-	-	+	+	+	V	V	V	-	S	S
H. cinaedi	+	+	ND	ND	-	-	-	-	-	-	-	ND	+	-	S	V
H. fennelliae	+	-	ND	ND	-	-	-	+	-	-	-	ND	+	-	S	S
H. pullorum	+	+	ND	ND	-	-	-	-	-	-	+	ND	ND	ND	R	S
H. pylori	+	-	ND	ND	+	-	-	-	-	-	-	ND	-	-	R	S
H. westmedii	+	+	ND	ND	-	-	-	ND	-	-	-	ND	ND	ND	S	R

TSI, triple sugar iron; W, weak reaction; V, variable reaction; ND, not determined; S, susceptible; R, resistant.

Adapted from Nachamkin I. *Campylobacter* and *Arcobacter*. In: Murray PR, Baron EJ, Pfaller MA, et al. eds. *Manual of clinical microbiology*, 7th ed. Washigton, DC: American Society for Microbiology, 1999: 717.

Direct Antigen or Nucleic Acid Detection

A nucleic acid probe for *Campylobacter* is available to identify suspected colonies or growth in enrichment cultures. Probes are not commercially available for the direct detection of *Campylobacter* or *Helicobacter* in stool specimens. However, an immunoassay for the detection of *H. pylori* antigens in fecal specimens is commercially available.

Serologic Procedures

Currently, there are no commercial systems available for the detection of antibodies to *Campylobacter* infection. Commercial EIAs are used to screen patients for exposure to *H. pylori* and have sensitivities and specificities of 86% to 100% and 76% to 98%, respectively (77,78).

Interpretation of Results

The isolation of *C. jejuni, C. coli*, C. fetus subsp *fetus*, or *H. pylori* is considered clinically significant. The isolation of the other *Campylobacter, Arcobacter*, and *Helicobacter* species should be evaluated with respect to the source and symptoms of the patient.

Therapeutic Considerations

Treatment of gastroenteritis is effective in reducing the duration of symptoms and elimination of the organism, but most cases of gastroenteritis are self-limited and do not require antimicrobial therapy. Septic or bacteremic patients require therapy. *C. jejuni* is variably susceptible to a number of antimicrobial agents (e.g., macrolides, tetracyclines, aminoglycosides, quinolones, and clindamycin); erythromycin is the drug of choice for gastroenteritis. Treatment of serious extraintestinal infections should be based on susceptibility testing. *H. pylori* is generally susceptible to amoxicillin, macrolides, nitrofurans, metronidazole, and bismuth salts. Multidrug regimens for as long as 14 days are often used.

Summary and Conclusions

Campylobacter species are isolated from a variety of animals are most often transmitted to humans via contaminated food (e.g., meat and milk) and cause significant human disease such as gastroenteritis and disseminated infections. A microaerophilic environment is necessary for their isolation on laboratory media. Selective media or filtration is required for isolation from stool specimens. A limited number of tests are available for the differentiation of the various *Campylobacter* species. Erythromycin is the drug of choice in GI infections. *Campylobacter* species are among the most frequently isolated bacterial agents of diarrheal disease. *H. pylori* is associated with ulcers of the gastric or duodenal mucosa.

LEGIONELLA

Part of "51 - Bacteriology"

Microbiology

Legionella, the only genus in the family Legionellaceae, contains more than 41 species and subspecies. Most of the species are environmental isolates not associated with human infections.

Legionella species are aerobic, gram-negative rods that are 0.5 µm in width and 1.5 to 5 µm in length. The organisms appear coccobacillary in clinical specimens but may be filamentous when grown on some culture media. They stain poorly with safranin, are nutritionally fastidious, and do not grow on routine laboratory media. L-Cysteine is essential for initial growth on culture media, whereas iron stimulates growth. The cellular fatty acids of *Legionella* are primarily branched chain, and the bacteria contain large amounts of ubiquinones. *Legionella* are catalase positive, urease negative, and weakly oxidase positive. They liquefy gelatin, do not reduce nitrates, and do not metabolize carbohydrates. Some strains produce a pigment that fluoresces yellow-green, blue-white, or red-pink under long-wave ultraviolet light.

Spectrum of Disease

There are four categories of *Legionella* infections: (a) subclinical disease, (b) nonpneumonic infection, (c) pneumonia, and (d) extrapulmonary disease (79). The high prevalence of antibodies (5% to 10%) in the general population supports the concept that many individuals are exposed to *Legionella* species. Nonpneumonic infection is generally self-limited. Extrapulmonary infections are not common but have included bacteremia, sinusitis, pyelonephritis, pericarditis, hepatic and perirectal abscess, endocarditis, peritonitis, and wound infections. Risk factors for the development of pneumonia include cigarette smoking, advanced age, chronic lung disease, and immunosuppression. Most infections are caused by *L. pneumophila* or *L. micadei*.

Epidemiology

The natural habitat of *Legionella* is aquatic environments, including surface and potable water. Almost all water distribution systems can be contaminated with *Legionella* and serve as reservoirs for human exposure. The mode of transmission from potable water to humans is via aerosolization or aspiration of contaminated water. Dissemination from the GI tract after ingestion has been proposed but not documented. Epidemic outbreaks are usually associated with institutions such as hospitals, hotels, and resorts.

Diagnostic Techniques

Microscopy

Legionella stains faintly with a routine Gram stain. Visualization is enhanced by staining with carbol-fuchsin as the counterstain or staining only with crystal violet and Gram's iodine without decolorization. The organisms are more difficult to stain in tissue, where the recommended procedure is the Warthin-Starry or Dieterle stain. *L. micadei* may appear acid fast in tissue.

Culture Techniques

The preferred culture method is to inoculate the specimen to two agar plates, one of which is buffered charcoal yeast extract (BCYE) agar without inhibitory agents. A second medium, BCYE with antimicrobial agents, is recommended for specimens that are likely to be contaminated with other microbial flora. Acid treatment of potentially contaminated specimens may increase the yield (80). *Legionella* has been recovered from blood culture (e.g., biphasic, Bactec, and lysis centrifugation) but may require blind subculture from broth media. BCYE agar is incubated at 35° to 37°C in air or CO_2 (not more than 5%) for as long as 5 days. Pinpoint growth may be detected in 2 to 3 days. After 5 to 7 days of incubation, colonies are 3 to 4 mm in diameter, gray, convex, and iridescent.

Identification Methods

Suspected colonies on BCYE agar are subcultured to another BCYE plate and to blood agar or preferably to a BCYE plate without added cysteine. Growth on BCYE but not on the other medium is presumptive evidence of a *Legionella* species. *Francisella tularensis* and *Bordetella pertussis* grow on BCYE and may not grow on cysteine-deficient medium.

The isolate is identified at the genus level by commercial immunologic reagents or nucleic acid probes. *Legionella* species can be separated into groups based on a limited number of phenotypic characteristics but require genetic analysis for definitive identification. Because therapy is similar for all species, identification to the genus level is adequate for most clinical laboratories.

Direct Antigen or Nucleic Acid Detection

Kits for the direct detection of antigen by direct fluorescent antibody (DFA) or immunoassay are commercially available. The sensitivities of the DFA and immunoassay for urinary antigen are 25% to 80% and more than 70%, respectively (79). Specificity for the two methods is 96% to 100%.

Serologic Procedures

Several serologic tests are available for epidemiologic studies and diagnosis of *Legionella* infection. However, 2 to 8 weeks are sometimes needed to demonstrate an antibody rise. Cross-reactions with other gram-negative organisms are infrequently observed. The sensitivity and specificity of serologic diagnosis are 80% and 96% to 99%, respectively (79).

Interpretation of Results

Because *Legionella* species are not part of the normal human flora, the isolation of any species is clinically significant. Culture is the mainstay of diagnosis, with a sensitivity of 51% to 80% from respiratory secretions (81). Direct detection methods are helpful when positive but cannot be used to rule out infection by *Legionella*. A fourfold rise to a 1:128 antibody titer is presumptive evidence of a recent infection. A single high titer is suggestive of recent infection but is not diagnostic.

Therapeutic Considerations

A macrolide or fluoroquinolone remains the drug of choice for the treatment of *Legionella* infections. Rifampin may be added in some cases because the organism is primarily intracellular. Anecdotal reports suggest that imipenem, trimethoprim-sulfamethoxazole, or tetracycline may be efficacious. Standardized methods for *in vitro* susceptibility testing of *Legionella* are not available.

Summary and Conclusions

Legionella species are ubiquitous inhabitants of natural and man-made aquatic systems. Transmission is thought to occur from inhalation of aerosolized bacteria. Infections are usually asymptomatic or present as a flulike illness or pneumonia. Laboratory diagnosis is made by isolation on BCYE agar, direct detection of antigen in clinical specimens, or retrospectively by serology. Identification of the isolate to the genus level is sufficient in most instances. A macrolide or fluoroquinolone is the drug of choice for treatment.

HAEMOPHILUS

Part of "51 - Bacteriology"

Microbiology

The genus *Haemophilus* is one of three genera in the family Pasteurellaceae. Currently, the genus includes eight species associated with humans and five species with animals. *Haemophilus* species are facultatively anaerobic, nonmotile, pleomorphic, coccoid to coccobacillary, gram-negative rods that require hemin (X factor) or nicotinamide adenine dinucleotide (NAD, V factor) or both for growth. They are usually oxidase and catalase positive, and CO_2 enhances the growth of some species.

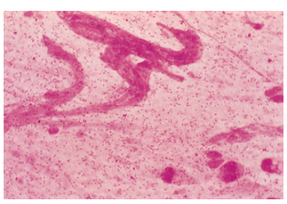


FIGURE 51.13. Gram stain of *Haemophilus influenzae* in sputum.

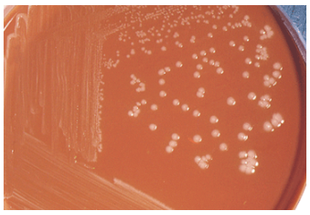


FIGURE 51.14. Colony of Haemophilus influenzae on chocolate agar.

Spectrum of Disease

H. influenzae consists of six capsular serotypes (a to f) and unencapsulated strains. Most *Haemophilus* infections in humans are caused by *H. influenzae*, and approximately 95% of cases of invasive disease were due to serotype b in the prevaccine era. *H. influenzae* infections occur primarily in children and include meningitis, bacteremia, epiglottitis, pneumonia and empyema, septic arthritis, cellulitis, pericarditis, osteomyelitis, and a variety of less frequently encountered infections. Unencapsulated strains are associated with infections such as chronic bronchitis, otitis media, sinusitis, and conjunctivitis. After introduction of the vaccine, invasive infection declined 95% in children (82).

The other *Haemophilus* species are uncommon causes of infection. *H. ducreyi* is the etiologic agent of the sexually transmitted disease chancroid (soft chancre).

Epidemiology

Haemophilus species are residents in the upper respiratory tract (nasopharynx, pharynx, and oral cavity) and the urogenital tract. Unencapsulated strains of *H. influenzae* are isolated from 50% of the population, whereas serotype b is carried by less than 5% of children and rarely by adults. Transmission of *H. influenzae* is thought to occur via respiratory droplets and is enhanced in crowded conditions (e.g., day care centers). *H. ducreyi* is transmitted by sexual contact. The highest incidence of chancroid is in Asia, Africa, and Central America.

Diagnostic Techniques

Microscopy

The presence of small, gram-negative coccobacilli in sterile body fluids (e.g., CSF, synovial fluids) or in respiratory secretions is suggestive of *Haemophilus* (Fig. 51.13).

Gram-stained smears of exudate from genital lesions revealing large numbers of gram-negative coccobacilli occurring in parallel chains (schools of fish) is suggestive of *H. ducreyi*, but smears are usually negative.

Culture Techniques

Specimens should be plated on enriched chocolate agar and incubated in 5% CO_2 at 35° to 37°C for 2 to 3 days. *Haemophilus* may be isolated on blood agar as satelliting colonies around staphylococci or other organisms.

Although *H. ducreyi* may be isolated on enriched chocolate agar containing vancomycin (3 µg/mL), other media (e.g., gonococcal base with 5% Fildes reagent, 5% horse blood with 3 µg/mL vancomycin) and multiple cultures are recommended for optimal recovery (82). Media should be incubated at 33°C in 5% CO_2 in a moist environment for 72 hours.

Haemophilus colonies on chocolate agar are usually small (0.5 to 1 mm), grayish, round, translucent, and smooth (Fig. 51.14). *H. haemolyticus* and *H. parahaemolyticus* may be ß-hemolytic on horse blood. Colonies of *H. ducreyi* are small, flat, smooth, and yellow-gray and can be pushed across the agar. *H. aphrophilus* and *H. paraphrophilus* produce larger colonies (1 to 1.5 mm) that are opaque, granular, and yellowish.

Identification Methods

Most laboratories presumptively identify *Haemophilus* species based on their growth around the disks containing hemin or NAD or by using the commercially available porphyrin test. Additional testing is necessary for a definitive identification (Table 51.18). *H. influenzae* and *H. parainfluenzae* can be further differentiated into biotypes (82). Commercial systems are available for the identification and biotyping of some *Haemophilus* species. A DNA probe is available for the culture identification of *H. influenzae*.

TABLE 51.18. DIFFERENTIAL CHARACTERISTICS OF THE GENUS *HAEMOPHILUS*

Species	X-Factor Requirement[a]	V-Factor Requirement	Indole	Urease	Hemolysis	Oxidase	Catalase	Glucose	Sucrose	Lactose
H. influenzae[b]	+	+	V[c]	V	-	+	+	+	-	-
H. haemolyticus	+	+	V	-	+	+	+	+	-	-
H. parainfluenzae	-	+	V	V	-	+	V	+	+	-
H. segnis	-	+	-	-	-	-	V	W[d]	W	-
H. parahaemolyticus	-	+	-	+	+	+	+	+	+	-
H. paraphrophilus	-	+	-	-	-	+	-	+	+	+
H. aphrophilus	-	-	-	-	-	-	-	+	+	+
H. ducreyi	+	-	-	-	-	+	-	-	-	-

[a] As determined by prophyrin test.

[b] Includes biogroup *aegypticus*.

[c] Variable

[d] Weak.

Modified from Albritton WL *Haemophilus influenzae* infections. In: Balows A, Hausler WJ Jr, Ohashi M, et al. eds. *Laboratory diagnosis of infectious disease principles and practice, vol 1*. New York: Springer-Verlag, 1983:302–311.

Direct Antigen and Nucleic Acid Detection

Particle agglutination kits are available for the direct detection of *H. influenzae* type b capsular antigen in clinical specimens but are seldom used by laboratories. The systems have the highest sensitivity and specificity when performed on CSF rather than on urine or serum. The sensitivity of the test on CSF specimens varies from 70% to 90%.

Serologic Procedures

Serologic procedures are not generally useful for assisting in the diagnosis of infection by *Haemophilus* species.

Interpretation of Results

The isolation of *H. ducreyi*, other *Haemophilus* species from normally sterile body fluids or deep tissues, or *H. influenzae* from the eye is clinically relevant. The significance of the isolation of *H. influenzae* (especially serotype b) or other species from respiratory secretions depends on the evaluation of the patient.

Therapeutic Considerations

Empiric treatment for life-threatening infections caused by *H. influenzae* type b consists of third-generation cephalosporins. ß-Lactamase activity and susceptibility testing should be performed on isolates from serious infections.

Summary and Conclusions

Haemophilus species are a component of the normal flora of the upper respiratory tract and, in part, the genital tract. *H. influenzae* (especially serotype b) causes the majority of human infections, which range from invasive life-threatening infections (e.g., meningitis) to less serious problems such as bronchitis. The other species are infrequently implicated but can cause similar types of infections, especially endocarditis. *H. ducreyi* is the etiologic agent of chancroid. The requirement of hemin and NAD for growth helps determine the identity of some species. The increasing antimicrobial resistance of *H. influenzae* has complicated therapy; therefore, routine susceptibility testing of clinically significant isolates is necessary.

NONFERMENTATIVE GRAM-NEGATIVE BACILLI

Part of "51 - Bacteriology"

Microbiology

The group of nonfermentative gram-negative bacilli (NFB) is one of convenience for the clinical microbiologist and describes bacteria that are gram-negative rods and relatively inert biochemically. On the oxidation-reduction medium of Hugh and Leifson, they may produce an oxidative or alkaline reaction or no reaction. Most species are strict aerobes, are catalase positive (*Eikenella* is catalase negative and facultatively anaerobic), and are oxidase positive (*Acinetobacter* and CDC group Ve-1,Ve-2 are negative and *Burkholderia cepacia* and *Stenotrophomonas maltophilia* are weakly positive). *Kingella* and CDC group EF-4a are weak fermenters and may be mistaken for NFB. Most species grow on MacConkey agar (except *Eikenella, Moraxella*, some *Flavobacterium* species) at 35°C in air. Some species grow better at 25° to 30°C. The arrangement of the flagella may be useful for the identification of some species.

TABLE 51.19. TAXONOMY OF NONFERMENTATIVE GRAM-NEGATIVE BACILLI

Current Name	Previous Name
Acidovorax delafieldii	*Pseudomonas delafieldii*
Alcaligenes faecalis	*Alcaligenes odorans*
Brevundimonas dimanuta	*Pseudomonas diminuta*
Brevundimonas vesicularis	*Pseudomonas vesicularis*
Burkholderia cepacia	*Pseudomonas cepacia*
Burkholderia pseudomallei	*Pseudomonas pseudomallei*
Burkholderia mallei	*Pseudomonas mallei*
Burkholderia gladioli	*Pseudomonas gladioli*
Chryseobacterium meningosepticum	*Flavobacterium meningosepticun*
Comamonas acidovorans	*Pseudomonas acidovorans*
Comamonas testosteroni	*Pseudomonas testosteroni*
Myroides odoratus	*Flavobacterium odoratum*
Ochrobactrum anthropi	CDC group Vd
Oligella urethralis	*Moraxella urethralis*
Oligella urealytica	CDC group IVe
Ralstonia pickettii	*Pseudomonas pickettii*
Sphingobacterium sp	*Flavobacterium* sp
Stenotrophomonas maltophilia	*Xanthomonas maltophilia*

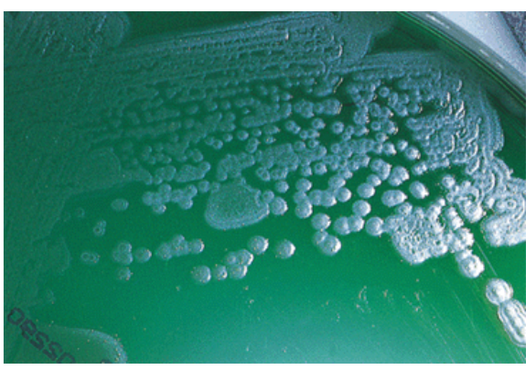


FIGURE 51.15. Colony of *Pseudomonas aeruginosa* on Mueller-Hinton agar. Note the blue-green diffusible pigment.

Members of the genus *Pseudomonas* are now classified into nine additional genera: *Burkholderia, Stenotrophomonas, Comamonas, Acidovorax, Brevundimonas, Shewanella, Methylobacterium, Sphingomonas*, and *Ralstonia* (83). The classification of the NFB is constantly being revised. Table 51.19 lists the current name and the former name for selected species. The reader is referred

to the reference section for a more thorough discussion of the classification of NFB (84).

Spectrum of Disease

The NFB comprise approximately 15% to 20% of all the gram-negative rods that are isolated in the clinical laboratory. *Pseudomonas aeruginosa* is the most frequently recovered NFB and a major cause of nosocomial infections, especially in the immunocompromised person. Infection with *P. aeruginosa* is often life-threatening and can present as bacteremia, pneumonia, skin and soft-tissue infection, urinary tract infection, or GI infection. Burn and cystic fibrosis patients are highly susceptible to infection. Community-acquired infections include folliculitis associated with whirlpools; otitis externa in divers, swimmers, and diabetics; osteomyelitis associated with deep puncture wounds; ocular infections associated with wearing contact lenses; and endocarditis and vertebral osteomyelitis in injection drug users. The next most frequently isolated species are *Acinetobacter, B. cepacia*, and *S. maltophilia*. Infections by the other species are rare, and most isolations are not clinically significant but represent colonization. Infections, when they occur, are usually hospital-acquired necrotizing pneumonia, urinary tract infection, wound infection, bacteremia, and endocarditis. These organisms are not associated with GI disease and are seldom isolated from feces.

Acinetobacter is recovered primarily as a colonizer but has caused suppurative infections in nearly every organ system. It is primarily a nosocomial pathogen, with most infections caused by *A. baumannii*. Nosocomial infections include respiratory infection, catheter site infection, and urinary tract infection.

B. pseudomallei causes melioidosis, a rare infection in North America. Most cases occur in people returning from travel to Southeast Asia or northern Australia. The disease often presents with acute pneumonia but may occur as a localized infection. A high mortality rate is associated with septicemia and dissemination.

B. mallei causes glanders, a disease of horses, mules, and donkeys that may be transmitted to humans. It is rarely seen today.

B. cepacia and *S. maltophilia* are important nosocomial pathogens. Although less virulent than *P. aeruginosa*, these organisms cause similar types of infections in the hospitalized patient. *B. cepacia* infection in patients with cystic fibrosis or chronic granulomatous disease produces significant morbidity and mortality. Cystic fibrosis patients experience a rapid decrease in lung function, frequent bacteremia, and death (85).

Alcaligenes xylosoxidans has been reported to cause individual cases of meningitis, pneumonia, peritonitis, urinary tract infection, bacteremia, and endocarditis (prosthetic valve). *Oligella* species have been recovered primarily from the urinary tract and may cause bacteremia. *Chryseomonas* species cause infrequent cases of neonatal meningitis, bacteremia, endocarditis, wound infections, and respiratory tract infections. *Weeksella zoohelcum* is isolated from bites and scratches of dogs and cats and *W. virosa* from urinary tract infections. *Ochrobactrum* causes catheter-related septicemia. *Methylobacterium* species has been associated with bacteremia, skin ulcers, and peritonitis in patients on continuous ambulatory peritoneal dialysis. *Moraxella* (especially *M. lacunata*) causes primarily eye infections. *Eikenella* has been recovered from a variety of infections, including human bites, head and neck infections, skin infections in injection drug users, and a variety of other deep-seated or systemic infections.

Epidemiology

Most NFB can be isolated from soil, water, vegetation (including fruits and vegetables), or nearly any moist location. Minimal nutritional requirements and their antimicrobial resistance, particularly *P. aeruginosa*, have led to their emergence as nosocomial pathogens. Hospital reservoirs include respiratory equipment, disinfectants, sinks, solutions, food, whereas disease-related reservoirs outside the hospital include whirlpools, swimming pools, cosmetics, illicit injectable drugs, hot tubs, and contact lens solution. *P. aeruginosa* colonizes the skin, oropharynx, and GI tract in 2% to 10% of healthy persons. The rate of colonization is usually higher in hospitalized patients. Transmission occurs via contaminated solutions, equipment, food, and probably on the hands of hospital personnel.

B. pseudomallei inhabits soil and water but is restricted primarily to Southeast Asia and other tropical or subtropical parts of the world. Transmission is primarily through abrasions or cuts but may occur by inhalation or ingestion.

Eikenella and *Moraxella* are common inhabitants of the mucous membranes of the upper respiratory tract. *Eikenella* is also part of the GI flora. Infections arise from breaks in the mucous membranes.

Diagnostic Procedures

Microscopy

All the NFB are gram-negative rods. In general, *Pseudomonas* species appear as thin rods, whereas *Acinetobacter* and *Moraxella* are coccobacillary. The morphology alone cannot be used for a definitive identification. *Acinetobacter* and *Moraxella* may resemble *Neisseria*.

TABLE 51.20. CHARACTERISTICS OF GENERA OF NONFERMENTATIVE BACTERIA

Genus	Catalase	Oxidase	Growth on MacConkey Agar	Motility	Utilization of Glucose
Pseudomonas	+	+	Good	Motile by means of polar flagella	Oxidative
Achromobacter	+	+	Good	Motile by means of peritrichous flagella	Oxidative
Alcaligenes	+	+	Good	Motile by menas of peritrichous flagella	Inactive
Acinetobacter	+	-	Good	Nonmotile	Oxidative or inactive
Chryseobacterium	+	+	Variable	Nonmotile	Oxidative
Moraxella	+/-	+	Variable	Nonmotile	Inactive
Eikenella	-	+	Negative	Nonmotile	Inactive
Kingella	-	+	Variable	Nonmotile	Delayed fermentative
EF-4	+	+	Variable	Nonmotile	Delayed fermentative

+, positive; -, negative; +/-, most strains positive.

Modified from Oberhofer TR, Howard RJ. Nonfermentative gram-negative bacteria. In: Howard BJ, Klass J M, Rubin SJ, et al. eds. *Clinical and pathogenic microbiology*. St. Louis: Mosby, 1987:329–358.

Culture Techniques

Most NFB are recovered on blood or MacConkey agars after 24 hours at 35°C in air or CO_2. *Eikenella, Methylobacterium, Weeksella*,

some *Moraxella*, *Chryseobacterium* species and CDC groups do not grow on MacConkey agar, and some NFB grow better at 25° to 30°C. The use of selective media is recommended for the recovery of *B. cepacia* from cystic fibrosis patients.

P. aeruginosa produces a flat colony with a feathered edge, may be ß-hemolytic, has a grapelike odor, and may produce a variety of pigments [pyocyanin (blue-green), pyoverdin (yellow-green), pyorubin (red), or pyomelanin (brown)] (Fig. 51.15). *B. pseudomallei*, *P. stutzeri*, *P. luteola*, and *P. oryzihabitans* form dry, wrinkled colonies. *Alcaligenes* may produce a fruity odor, whereas *B. pseudomallei* and some strains of *P. aeruginosa* may have a pungent, earthy odor. The *Chyrseobacterium*, *Myroides*, and *Flavobacterium* species generally produce a yellow pigment; *Acinetobacter* colonies resemble the Enterobacteriaceae; *Moraxella* colonies are tiny (less than 1 mm) after 24 hours; and *Eikenella* colonies may pit or corrode the agar.

Identification Methods

The initial characterization of NFB is based on the presence of pigment, oxidase reaction, catalase reaction, growth on MacConkey agar, motility, and whether the organism uses glucose fermentatively, oxidatively, or not at all (Table 51.20).

The differential characteristics of *P. aeruginosa* and the other NFB are presented in Table 51.21,Table 51.22,Table 51.23, and Table 51.24. If pyocyanin is present and the isolate is oxidase positive, it is identified as *P. aeruginosa* without any additional tests. *B. cepacia*, *S. maltophilia*, and *Pseudomonas* species are motile with polar flagella, grow on MacConkey agar, are oxidase positive (*B. cepacia*, *S. maltophilia*, *P. luteola*, and *P. orzihabitans* may be negative), glucose oxidizers, and catalase positive.

TABLE 51.21. CHARACTERISTICS OF *PSEUDOMONAS* SPECIES FOUND IN CLINICAL SPECIMENS

Test	*P. aeruginosa* (*n* = 201)	*P. fluorescens* (*n* = 155)	*P. putida* (*n* = 16)	*P. stutzeri* (*n* = 28)	*P. mendocina* (*n* = 4)	*P. pseudoalcaligenes* (*n* = 34)	*P. alcaligenes* (*n* = 26)	*Pseudomonas sp* CDC group 1[6] (*n* = 31)	*P. luteola* (*n* = 34)	*P. oryzihabitans* (*n* = 36)
Oxidase	99	97	100	100	100	100	96	100	0	0
Growth										
MacConkey	100	100	100	100	100	100	96	97	100	100
Cetrimide	94	89	81 (6)	4	75 (25)	56 (18)	15	13 (6)	0	25 (28)
6.5% NaCl	65	43	100	80 (16)	100	62 (6)	41	14	74	62
42°C	100	0	0	69	100	94	0	48	94	33
Nitrate reduction	98	19	0	100	100	100	54	100	62	6
Gas from nitrate	93	3	0	100	100	0	0	100	0	0
Pyoverdin	65	96	93	0	0	0	0	0	0	0
Arginine dihydrolase	100	97	100	0	100	78	12	33	100	14
Lysine decarboxylase	0	0	0	0	0	0	0	0	0	7
Ornithine decarboxylase	0	0	0	0	0	0	0	0	0	3
Indole	0	0	0	0	0	0	0	0	0	0
Litmus milk	89 pep[b]	95 pep	62 k[b]	57 k	25 (75) k	38 k	46 k	39 k	44 k	57 k
Hydrolysis										
Urea	48 (9)	21 (31)	13 (44)	33 (22)	50	3 (6)	0	3 (7)	26 (38)	77
Gelatin	82	100	0	0	0	0	0	4	61	17
Acetamide	100	6 (12)	0	0	0	ND	ND	ND	ND	ND
Esculin	0	0	0	0	0	0	0	0	100	0
Acid from[c]										
Glucose	97	100	100	96	100	9	0	0	100	100
Fructose	ND	ND	ND	ND	ND	79 (21)	0	ND	ND	ND
Starch	ND	ND	ND	100	0	ND	ND	ND	ND	ND
Xylose	90	100	100	93 (7)	75 (25)	18 (12)	0	0	100	100
Lactose	<1	24	25 (13)	0	0	0	0	0	3 (24)	14 (22)
Sucrose	0	48	0	0	0	0	0	0	12	25
Maltose	<1	2	31	100	0	0	0	0	100	97
Mannitol	70	53	25	89	0	0	0	0	76 (18)	100
Simmons citrate	95	93	94(6)	82 (14)	100	26 (9)	57 (8)	42 (6)	100	97
No. of flagella	1	>1	>1	1	1	1	1	1	>1	1

[a] Percentage of strains positives; percentage in parentheses represents strains with delayed reactions.

[b] Type of reaction on litmus milk.

[c] Oxidative-fermentation basal medium with 1% carbohydrate.

PEP, peptonization; k, alkaline; ND, no data.

Data from Kiska DL, Gilligan PH. *Pseudomonas*. In: Murray R, Baron EJ, Pfaller MA, et al., eds. *Mannual of clinical microbiology*, 7th ed., Washington, DC: American Society for Microbiology, 1999: 518-519.

Alcaligenes are motile by peritrichous flagella, are oxidase and catalase positive, grow on MacConkey agar, do not oxidize glucose, and are relatively inert. Species of *Acinetobacter* are nonmotile, catalase positive, oxidase negative, and nitrate negative, grow on MacConkey agar, and are nonfermentative. *Chyrseobacterium*, *Myroides*, *Flavobacterium* species, some *Weeksella*, and some *Sphingobacterium* species are nonmotile, are catalase and oxidase positive, exhibit variable growth on MacConkey agar, produce a yellow pigment, and may be oxidative or inactive toward glucose. The *Moraxella* species are nonmotile, inactive toward glucose, oxidase and catalase positive, and exhibit variable growth on MacConkey agar. *Eikenella corrodens* is oxidase positive, catalase negative, and inactive toward glucose, does not grow on MacConkey agar, is nonmotile, may pit the agar, and has a bleachlike odor.

Several commercial systems with variable accuracy are available for the identification of NFB. Another commercially available system analyzes cell-wall fatty acids by gas-liquid chromatography.

Direct Antigen or Nucleic Acid Detection

No commercial system is available for the direct detection of NFB in clinical specimens, although research methods are available for *B. pseudomallei* and *B. cepacia*.

Serologic Procedures

A fourfold rise in titer with the hemagglutination, complement-fixation, or agglutination test aids in the diagnosis of melioidosis. Titers greater than 1:8 (complement fixation) or greater than 1:160 (agglutination) are highly suggestive of infection during the acute illness. These tests are performed only in reference laboratories because of the low incidence of the disease. Serologic tests for the diagnosis of infections caused by the other NFB are not routinely available.

Interpretation of Results

The isolation of *B. pseudomallei* or *B. mallei* is always clinically significant. Because *P. aeruginosa* is primarily a nosocomial pathogen and is so devastating in the compromised host, its presence in clinical specimens should be considered relevant until proven otherwise. The colonization of the skin and mucous membranes in the compromised patient often precedes systemic infection. *P. aeruginosa* and *B. cepacia* are significant isolates in cystic fibrosis patients. The other NFB are often colonizers, and the clinical significance of their isolation depends on the underlying condition of the patient, the source of the specimen, the number of other bacterial species isolated, and the need for epidemiologic data.

TABLE 51.22. CHARACTERISTICS OF *BURKHOLDERIA* SPECIES AND *RALSTONIA PICKETTII* FROM CLINICAL SPECIMENS

	% Positive Strains						
					R. pickettii		
Test	*B. pseudomallei* (*n* = 70)	*B. mallei* (*n* = 8)	*B. cepacia* (*n* = 159)	*B. gladioli* (*n* = 58)	Biovar 1 (*n* = 70)	Biovar 2 (*n* = 54)	Biovar 3 (*n* = 31)
Oxidase	100	25	86	47	100	100	100
Growth							
MacConkey	100	88	100	97	99	100	100
Cetrimide	0	0	44	3	1	0	0
42°C	100	0	83	9	83	3	84
Nitrate reduction	100	100	57	43	100	100	13
Gas from nitrate	100	0	0	0	86	100	0
Arginine dihydrolase	100	100	0	2	6	0	3
Lysine decarboxylase	0	0	80	0	0	0	0
Ornithine decarboxylase	0	0	48	0	0	0	0
Hemolysis	0	0	0	0	0	0	0
Hydrolysis							
Urea	13	12	60	30	100	100	81
Citrate	77	0	94	93	99	100	100
Gelatin	79	0	20	12	12	38	30
Esculin	59	0	63	0	0	0	0
Acid from							
Glucose[a]	100	100	100	98	100	100	100
Xylose	86	12	100	98	100	100	100
Lacrose	99	12	99	9	100	0	100
Sucrose	66	0	86	0	0	0	0
Maltose	99	0	99	0	100	0	100
Mannitol	94	62	100	91	0	0	100
Motile	100	0	100	100	100	100	100
No. of flagella	>2	0	>2	>2	1–2	1–2	1–2

[a] Oxidation-fermentation basal medium with 1% carbohydrates.

Data from Gilligan PH, Whittier S. *Burkholderia*, *Stenotrophomonas*, *Ralstonia*, *Brevundimonas*, *Comamonas*, and *Acidovorax*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999: 526–538.

Therapeutic Considerations

The antimicrobial susceptibility of *P. aeruginosa* and many of the other genera is not predictable. Appropriate therapy depends on the results of susceptibility testing. However, interpretive breakpoints for many of the NFB have not been determined. Empiric therapy involves one, two, or three agents, depending on the severity of the disease and the susceptibility pattern of local isolates. In general, therapy will include an aminoglycoside and an antipseudomonal penicillin or parenteral cephalosporin active against *P. aeruginosa*. The currently recommended treatment for the septicemic form of melioidosis is trimethoprim-sulfamethoxazole, ceftazidime, amoxicillin-clavulanate, or imipenem. Trimethoprim-sulfamethoxazole can be used for *B. cepacia* or *S. maltophilia* infections.

Alcaligenes is usually susceptible to trimethoprim-sulfamethoxazole and variably susceptible to most other antimicrobial agents. *Moraxella* and *Eikenella* are usually susceptible to penicillin and ampicillin, although isolates containing β-lactamase have been reported. *Eikenella* is resistant to clindamycin, cephalexin, erythromycin, and metronidazole.

Summary and Conclusions

The NFB are ubiquitous in soil, water, and most environments. *Eikenella* and *Moraxella* are part of the normal flora of the mucous membranes in humans and cause endogenously acquired infections. The other NFB, particularly *P. aeruginosa*, are important nosocomial pathogens found in numerous hospital reservoirs (e.g., contaminated solutions, equipment, food, and sinks) and infect nearly every organ system. *B. pseudomallei* and *B. mallei* are the etiologic agents of melioidosis and glanders, respectively. In general, the NFB are easily isolated and grow on most routine media, but they are difficult to identify by either conventional tests

or commercial systems. Pyocyanin pigment is a definitive characteristic of *P. aeruginosa*. The susceptibility of the NFB to antimicrobial agents must be determined by testing. *Eikenella* and *Moraxella* are usually susceptible to penicillin or ampicillin.

TABLE 51.23. CHARACTERISTICS OF *ACIDOVORAX*, *BREVUNDIMONAS*, *comamonas*, AND *STENOTROPHOMONAS* SPECIES FOUND IN CLINICAL SPECIMENS

	% Strains Positive							
Test	*A. delafieldii* (*n* = 2)	*A. facilis* (*n* = 2)	*A. temperans* (*n* = 2)	*B. diminuta* (*n* = 68)	*B. vesicularis* (*n* = 94)	*C. acidovorans* (*n* = 69)	*C. testosteroni* (*n* = 1)	*S. maltophilia* (*n* = 228)
Oxidase	100	100	100	100	98	100	100	0
Growth								
MacConkey	100	0	100	100	43	100	100	100
Cetrimide	0	0	0	0	0	4	0	2
6.0% NaCl	0	0	0	21	23	6	0	22
42°C	50	0	100	38	19	29	100	48
Nitrate reduction	100	100	100	3	5	99	100	39
Gas from nitrate	0	0	100	0	0	0	0	0
Pigment	Yellow; soluble	None	Yellow; soluble	Brown-tan; soluble	52%; yellow-orange; insoluble	26%; fluorescent	Tan; soluble	Brown-tan; soluble
Arginine dihydrolase	100	100	0	0	0	0	0	0
Lysine decarboxylase	0	0	0	0	0	0	0	93
Ornithine decarboxylase	0	0	0	0	0	0	0	0
Indole	0	0	0	0	0	0	0	0
Hemolysis	0	0	0	0	0	0	0	1
Hydrolysis								
Urea	100	100	50	13	2	0	0	3
Citrate	100	0	0	1	1	94	100	34
Gelatin	0	100	0	68	25	11	0	93
Esculin	0	0	0	5	88	0	0	39
Acid from								
Glucose[a]	100	100	100	21	87	0	0	85
Xylose	85	100	0	0	27	0	0	35
Lacrose	0	0	0	0	0	0	0	60
Sucrose	0	0	0	0	0	0	0	63
Maltose	0	0	0	0	94	0	0	100
Mannitol	50	100	50	0	0	100	0	0
H_2S[b]	100	100	100	34	49	57	0	95
Motility	100	100	100	100	100	100	100	100
No. of flagella	1–2	1–2	1–2	1–2	1–2	>2	>2	>2

[a] Oxidation-fermentation basal medium with 1% carbohydrate.

[b] Lead acetate paper.

Data from Gilligan PH, Whittier S. *Burkholderia, Stenotrophomonas, Ralstonia, Brevundimonas, Comamonas*, and *Acidovorax*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999: 526–538.

MISCELLANEOUS FACULTATIVELY ANAEROBIC GRAM-NEGATIVE BACILLI

Part of "51 - Bacteriology"

Microbiology

The genera and species discussed in this section are usually slow growing; may not grow on MacConkey, blood, or chocolate agar; and often grow better in a CO_2-enriched atmosphere. They include *Gardnerella, Cardiobacterium, Chromobacterium, Pasteurella, Actinobacillus, Kingella, Capnocytophaga, Streptobacillus, Suttonella*, and CDC group EF-4.

G. vaginalis was previously classified as *Haemophilus vaginalis* or *Corynebacterium vaginale*. The cell wall is similar to but much thinner than gram-positive organisms. *G. vaginalis* is a small (0.5 × 1.5 µm), pleomorphic, fastidious, facultatively anaerobic, nonmotile, oxidase- and catalase-negative, gram-negative to gram-variable rod that does not grow on MacConkey agar.

Cardiobacterium hominis and *Suttonella indologenes* are a facultatively anaerobic, slow-growing, fastidious, nonmotile, oxidase-positive, catalase-negative, pale-staining gram-negative rods. They do not grow on MacConkey agar, and growth is enhanced by humidity and CO_2.

The genus *Chromobacterium* contains two species, but only *C. violaceum* is isolated from clinical specimens. It is a facultatively anaerobic, motile, oxidase- and catalase-positive, gram-negative rod that usually produces a violet pigment. Unlike the other organisms in this section, *C. violaceum* grows on most routine laboratory media in ambient air.

The genus *Pasteurella* contains 20 recognized species and is similar to *Actinobacillus* based on moles percent G + C of their chromosomal DNA and phenotypic characteristics. *P. multocida* is the most common human pathogen of the genus. The other species are only rarely isolated from clinical specimens. *P. multocida* is a facultatively anaerobic, nonmotile, catalase- and oxidase-positive, gram-negative rod that may exhibit bipolar staining and does not grow on MacConkey agar.

The genus *Actinobacillus* contains six species, *A. ureae, A. lignieresii, A. equuli, A. suis, A. hominis*, and *A. pleuropneumoniae*. The *Actinobacillus* species are primarily animal pathogens. *A. actinomycetemcomitans* is closely related to *H. aphrophilus* and *H. segris*. These latter three species do not belong in the genus *Haemophilus* or *Actinobacillus* and probably will be classified in a new genus in the future. *A. actinomycetemcomitans* is a facultatively

TABLE 51.24. CHARACTERISTICS OF *EIKENELLA CORRODENS*, *MORAXELLA*, *PSYCHROBACTER*, AND *OLIGELLA* SPECiEs

Characteristics	*E. corrodens*	*M. atlantae*	*M. lacunata*	*M. nonliquefaciens*	*M. osloensis*	*P. phenylpyruvica*	*Q. urethralis*
Urease	-	-	-	-	-	+	-
Phenylalanine deaminase	-	-	-/+	-	-	+	-/+
Nitrate reduction	+	-	+	+	+/-	+/-	-
Growth on MacConkey	-	+/-	-	-	-/+	+	+/-
Ornithine decarboxylase	+	NA	NA	NA	NA	NA	NA
Growth in 3% NaCl	NA	NA	+	NA	-	+	+

+, >90% of strains positive; –, >90% of strains negative; +/-, most strains positive; -/+, most strains negative; NA, not available.

anaerobic, fastidious, slow-growing, small (0.5 × 1.5 μm), nonmotile, catalase-positive, gram-negative coccobacillary rod that does not grow on MacConkey agar and grows best in humidity and CO_2.
The genus *Kingella* consists of three species: *K. kingae, K. denitrificans*, and *K. oralis*. *Kingella* species are facultatively anaerobic, nonmotile, catalase-negative, oxidase-positive, gram-negative rods that do not grow on MacConkey agar.
Seven species of *Capnocytophaga* are recognized: *C. ochracea, C. gingivalis, C. sputigena, C. canimorsus, C. cynodegmi, C. haemolytica*, and *C. granulosa*. They are facultatively anaerobic, fastidious, capnophilic, oxidase-variable, catalase-negative, gram-negative fusiform rods that may exhibit gliding motility and do not grow on MacConkey agar. *C. canimorsus* is a weakly oxidase-positive, catalase-positive, nonmotile, fermentative, long, thin, gram-negative fusiform rod that does not grow on MacConkey agar and grows poorly on other laboratory media (86).
CDC group EF-4 is an oxidase- and catalase-positive, nonmotile, fermentative (for glucose), gram-negative rod that grows poorly on MacConkey agar.

Spectrum of Disease

G. vaginalis is associated with BV (nonspecific vaginitis), which is characterized by a malodorous vaginal discharge and overgrowth with various anaerobic bacteria. Extragenital infections include bacteremia, vaginal abscesses, neonatal bacteremia, and soft-tissue infections. *C. hominis* and *S. indologenes* have been implicated in cases of endocarditis and bacteremia (87,88). *C. violaceum* is associated with urinary tract infections, localized abscesses, and fulminant sepsis with abscesses in multiple organs.
P. multocida is associated with three broad categories of disease: (a) localized infection after an animal bite or scratch (usually from a cat), (b) chronic pulmonary disease, particularly in patients with underlying pulmonary disease, and (c) bacteremia with or without metastatic foci. The latter group includes meningitis; brain, liver, and renal abscesses; otitis; mastoiditis; sinusitis; endocarditis; septic arthritis; and peritonitis. Spontaneous peritonitis and septicemia occur most often in patients with cirrhosis.
Actinobacillus is associated with juvenile periodontitis; endocarditis (particularly damaged or prosthetic heart valves); and infrequent reports of meningitis, brain abscess, osteomyelitis, urinary tract infection, pericarditis, and pulmonary infections.
Kingella infections include endocarditis, meningitis, bacteremia, septic arthritis, osteomyelitis, intervertebral disk infections, and infections of the eye. *K. kingae* infections occur primarily in children and young adults. *K. denitrificans* causes endocarditis in adults.
Capnocytophaga is associated with juvenile periodontitis, bacteremia, endocarditis, osteomyelitis, septic arthritis, sinusitis, pulmonary infections, conjunctivitis, and keratitis. Most infections occur in immunocompromised patients with severe granulocytopenia and oral ulcerations. *C. canimorsus* and *C. cynodegma* are associated with animal bites.
C. canimorsus causes two patterns of disease: (a) shock and disseminated intravascular coagulation seen in splenectomized patients, and (b) a milder form of disease in patients with a spleen. Infections include septicemia, meningitis, endocarditis, pneumonia, necrotizing wound infections, brain abscess, cellulitis, septic arthritis, corneal ulceration, and glomerulonephritis. Predisposing factors include splenectomy and a history of alcoholism.
CDC group EF-4 is associated with dogs and cats and cause wound infections and rarely bacteremia after bites, scratches, or contamination of a preexisting wound.

Epidemiology

G. vaginalis is part of the endogenous flora of the human genital tract and is associated with bacterial vaginosis (BV), the most common cause of vaginal discharge in sexually active women. BV is associated with preterm birth, premature rupture of membranes, and chorioamnionitis (89). Sexual transmission of the disease is not proven, but a high proportion of the male sex partners of women with the syndrome carry *G. vaginalis* on their urethral mucosa.
C. hominis, S. indologenes, A. actinomycetemcomitans, Kingella species, and *Capnocytophaga* species are part of the oropharyngeal flora in humans. Most infections result from endogenous spread of the bacterium, often after dental manipulations or oral disease such as periodontitis.
C. violaceum inhabits soil and water, particularly in warmer climates such as the southeastern United States, Southeast Asia, or South America. Infections generally occur after contamination of wounds with soil or water, or rarely after ingestion of contaminated water.
P. multocida, C. canimorsus, C. cynodegma, and CDC group EF-4 are found in the upper respiratory tract of dogs and cats. *C. canimorsus, C. cynodegma*, and EF-4a infections in humans generally follow dog bites or exposure to dog secretions. *C. canimorsus, C. cynodegma*, and EF-4b are also associated with cat bites and scratches. The reservoir for *P. multocida* is the upper respiratory tract of cats, dogs, other mammals, birds, and probably humans. The frequency of carriage varies with the species but is most common in cats. Transmission to humans occurs after bites, scratches, or exposure to secretions.

Diagnostic Procedures

Microscopy

BV can be presumptively diagnosed by Gram stain of vaginal secretions that demonstrate numerous clue cells (vaginal epithelial cells covered by gram-variable coccobacilli) associated with a mixed flora of large numbers of gram-negative to gram-variable coccobacilli in the absence of lactobacilli (90). The syndrome is best diagnosed by clinical criteria, a Gram stain or wet mount preparation of discharge, vaginal pH, and 10% KOH preparation of discharge rather than by culture of *G. vaginalis*.
The organisms discussed in this section are gram-negative rods that cannot be definitely identified by direct examination of the clinical specimen. *C. hominis* is a pale-staining pleomorphic (sometimes teardrop-shaped) rod; *P. multocida* is a small coccobacillus that may exhibit bipolar staining; *A. actinomycetemcomitans*

is a small coccobacillus; *Kingella* are small coccobacilli arranged in pairs or chains; *Capnocytophaga* are thin fusiform rods; and *C. canimorsus* is a filamentous, sometimes curved fusiform rod. Both *C. hominis* and *Kingella* have a tendency to retain the primary stain.

Culture Techniques

G. vaginalis is most easily isolated on human blood bilayer Tween agar (HBT), vaginalis agar (V-agar), or Columbia CNA. Media are incubated at 35° to 37°C in a humid atmosphere containing 5% CO_2 for 2 days. After 48 hours, colonies are approximately 0.5 mm in diameter and gray and exhibit diffuse ß hemolysis on human but not on sheep blood. On routine sheep blood agar, colonies are pinpoint to barely visible.

C. hominis produces pinpoint colonies on chocolate or blood agar after 2 to 3 days incubation in a moist, 5% CO_2 atmosphere. A slight green to brown color may develop around the colony.

C. violaceum grows well on blood and MacConkey agar in 24 hours at 35° to 37°C in air. Most colonies exhibit a characteristic violet pigment (produced optimally at 22°C) and often produce an odor of cyanide (almond).

Actinobacillus and *Capnocytophaga* grow best on blood or chocolate agar incubated in high humidity and in a 5% to 10% CO_2-enriched atmosphere. After 2 to 3 days of incubation, colonies of *Actinobacillus* are small (0.5 mm), punctate, and adherent with a mixed morphology. On continued incubation, a starlike, opaque center may be observed under 100× magnification (Fig. 51.16). *Capnocytophaga* produces yellow nonhemolytic colonies with fingerlike projections and an almond odor after 2 to 3 days of incubation. Some strains will grow only anaerobically on initial isolation. Colonies may pit the agar.

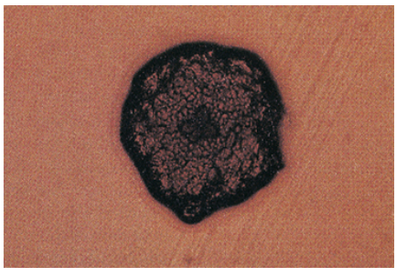


FIGURE 51.16. Colony of *A. actinomycetemcomitans* on heart infusion agar. Note the central star-shaped structure. (From Baron EJ, Finegold SM. *Bailey and Scott's diagnostic microbiology*, St. Louis: Mosby, 1990, with permission.)

Kingella grows in air or CO_2 (may enhance growth) but grows poorly under anaerobic conditions. Colonies are 0.5 to 1 mm in diameter after 48 hours of incubation at 35°C and may pit the agar or produce a "fried-egg" appearance. *K. denitrificans* grows on Martin-Lewis agar and may be confused with *N. gonorrhoeae*.

Pasteurella grows well on nonselective media such as blood or chocolate agar at 35°C in air or CO_2. Colonies are usually smooth, gray and 1 to 2 mm after 24 hours of incubation, have a musty odor, and may produce a green to brown discoloration on blood agar.

C. canimorsus grows slowly (2 to 4 days) on enriched chocolate agar incubated at 35°C in a humid, CO_2-enriched atmosphere. Growth is poor under anaerobic conditions.

EF-4 grow well on nonselective media (e.g., blood or chocolate agar) within 24 hours of incubation at 35°C. Colonies are small and yellow-orange and produce a popcornlike odor.

Identification Methods

G. vaginalis can be identified presumptively on the basis of negative catalase and oxidase tests, ß-hemolytic colonies on HBT agar, and a typical Gram-stain morphology. Confirmatory tests include starch and hippurate hydrolysis, α- and ß-glucosidase tests, or use of some commercial systems for identification of *Corynebacterium*.

C. hominis and *S. indologenes* are oxidase positive and catalase negative and are differentiated from similar organisms (e.g., *Eikenella, Kingella*, and *H. aphrophilus*) by the production of indole (Table 51.25). Serum may have to be added to carbohydrate tests to ensure growth.

TABLE 51.25. DIFFERENTIATION OF SELECTED GENERA AND SPECIES OF GRAM-NEGATIVE BACILII

Organism	Fusiform Shape	Catalase	Oxidase	Indole	Motility	Nitrate Reduction	Glucose Fermentation	Growth on MacConkey
Cardiobacterium	-	-	+	+	-	-	+	v
Chromobacterium violaceum	-	+	v	-[a]	+	+	+	+
Capnocytophaga	+	v	v	-	+	-	+	-
Actinobacillus actinomycetemcomitans	-	+	v	-	-	+	+	-
Pasteurella multocida	-	+	+	+	-	+	+	-
Eikenella corrodens	-	-	+	-	-	+	-	-
Kingella	-	-	+	v	-	v	+	v
Haemophilus aphrophilus	-	-	v	-	-	+	+	-
CDC Group EF4a	-	+	+	-	-	+	+	v

+, >90% positive; -, 90% negative; v, variable reaction; [a]Some nonpigmented stains are positive.

C. violaceum is oxidase and catalase positive and motile, reduces nitrate, and ferments glucose. It can be distinguished from other genera such as *Aeromonas, Plesiomonas*, and *Vibrio* by negative tests for indole, ornithine, and lysine decarboxylase reactions and lack of fermentation of mannitol and maltose.

Actinobacillus may be differentiated from similar genera with a minimum number of tests, but the definitive identification of a suspected isolate requires a more comprehensive battery.

Kingella species are catalase negative, oxidase positive, and nonmotile (Table 51.25). *K. kingae* is distinguished from the other two species by ß hemolysis on sheep blood agar and the absence of indole production, nitrate reduction, and growth on Martin-Lewis agar.

Pasteurella species are generally catalase negative, oxidase positive, and nonmotile (Table 51.25). Other characteristic reactions for *P. multocida* include a positive ornithine decarboxylase and indole reaction.

Characteristics useful for the identification of the genus *Capnocytophaga* are listed in Table 51.25. Commercial identification systems are also available. *C. canimorsus* often requires the addition of rabbit serum to obtain adequate growth in the biochemical tests. It is weakly oxidase (positive) and catalase positive; indole, urease, and nitrate-reduction negative; and arginine dihydrolase positive.

CDC group EF-4 strains are oxidase and catalase positive, reduce nitrate, and are negative for urease and indole production.

Direct Antigen to Nucleic Acid Detection

No commercial systems are available for the direct detection of any of these bacteria in clinical specimens.

Serologic Procedures

No routine serologic tests are available for the diagnosis of infections caused by this group of organisms.

Interpretation of Results

With the exception of *G. vaginalis,* these bacteria are seldom isolated from clinical specimens. Therefore, their isolation is clinically significant until proven otherwise. The isolation of any members of the HACEK group (*Haemophilus, Actinobacillus, Cardiobacterium, Eikenella, Kingella*) from blood is highly suspicious of endocarditis, especially if the patient has oral trauma or poor dentition. The presence of large numbers of *G. vaginalis* organisms in vaginal secretions is suggestive of BV but does not prove that *G. vaginalis* causes the syndrome.

Therapeutic Considerations

Standardized susceptibility tests are often difficult to perform because of the growth requirements of some of these bacteria. Oral metronidazole or clindamycin treatment is the treatment of choice for BV.

Cardiobacterium and *Suttonella* are susceptible to ß-lactam agents, tetracycline, chloramphenicol, and aminoglycosides but variably resistant to clindamycin, erythromycin, oxacillin, and vancomycin.

C. violaceum should be tested for susceptibility to antimicrobial agents. Most isolates are susceptible to tetracycline, chloramphenicol, erythromycin, and trimethoprim-sulfamethoxazole with variable resistance to aminoglycosides. They are often resistant to the ß-lactam antibiotics and clindamycin.

P. multocida is susceptible to penicillin, cephalothin, trimethoprim-sulfamethoxazole, and tetracycline. Limited experience is available for the treatment of infections with the newer antimicrobial agents.

Actinobacillus is generally susceptible to chloramphenicol, tetracycline, carbenicillin, and streptomycin, and is variably susceptible to penicillin. *In vitro* tests indicate susceptibility to the newer cephalosporins, rifampin, trimethoprim-sulfamethoxazole, and fluoroquinolones. Combined penicillin-aminoglycoside therapy has been recommended for endocarditis.

Kingella are generally susceptible to the ß-lactams (*K. denitrificans* may be moderately resistant to penicillin), cephalothin, aminoglycosides, tetracycline, chloramphenicol, and sulfonamides.

Capnocytophaga are susceptible to penicillin, extended-spectrum cephalosporins, chloramphenicol, tetracycline, erythromycin, clindamycin, quinolones, and are resistant to vancomycin, aminoglycosides, metronidazole, and trimethoprim-sulfamethoxazole.

CDC group EF-4a organisms appear susceptible to the ß-lactams, chloramphenicol, and tetracycline based on limited studies.

Summary and Conclusions

The miscellaneous facultatively anaerobic gram-negative bacilli discussed in this section are usually inhabitants of the upper respiratory tract of humans (e.g., *Actinobacillus, Kingella, Cardiobacterium*, and *Capnocytophaga*), the genital tract (e.g., *Gardnerella*), the upper respiratory tract of animals (e.g., *Pasteurella, C. canimorsus*, and CDC group EF-4), or soil and water (e.g., *Chromobacterium*). Infections usually follow the spread of the bacterium from a colonized site or after exposure to animal bites, scratches, or secretions. These species cause a variety of diseases ranging from localized soft-tissue infections to systemic life-threatening infections. Most of the organisms are fastidious, do not grow on MacConkey agar, and require increased humidity and CO_2-enriched atmosphere for growth. *Chromobacterium* grows well on most routine media. *C. violaceum* is resistant to penicillin and *A. actinomycetemcomitans* is moderately resistant; the other species are susceptible to penicillin.

MISCELLANEOUS GRAM-NEGATIVE BACILLI WITH SPECIAL GROWTH REQUIREMENTS

Part of "51 - Bacteriology"

Microbiology

The microorganisms discussed in this section are not usually isolated on routine laboratory media (e.g., blood and MacConkey agars) after 24 hours of incubation. In general, these bacteria are associated with a specific clinical syndrome of which the laboratory needs to be aware. Many of these organisms pose a considerable challenge regarding their isolation and identification.

The genus *Francisella* contains two species, *F. tularensis* and

F. philomiragia. The latter organism rarely causes disease in humans. *F. tularensis* is composed of three biovars: type A (biovar *tularensis*), type B (biovar *palaearctica*), and biovar *novicida*. Type A is the most virulent. *F. tularensis* is an aerobic, small (0.2 × 0.2 to 0.7 μm), nonmotile, oxidase-negative, weakly catalase positive, gram-negative rod that generally requires enriched media with added cystine or cysteine for growth.

There are six recognized species of the genus *Brucella*, four of which are pathogenic for humans (*B. abortus, B. melitensis, B. suis*, and *B. canis*). *B. neotomae* (found in the desert wood rat) and *B. ovis* (a pathogen of sheep) have not been shown to cause disease in humans. A new species, *B. maris*, has been isolated from marine mammals but not humans. DNA hybridization studies suggest that all species are very closely related. *Brucella* species are nonmotile, catalase-positive, usually oxidase-positive, gram-negative coccobacilli or short rods (0.5 × 0.6 to 1.5 μm) that are usually arranged singly but may occur in short chains or groups. The organisms are strict aerobes, often requiring a CO_2-enriched atmosphere on primary isolation. The addition of glucose (1.5% to 2.5%), serum, thiamine, nicotinamide, and magnesium to the medium promotes growth.

The genus *Bordetella* consists of seven species: *B. pertussis, B. parapertussis, B. bronchiseptica, B. avium, B. hinzii, B. holmesii*, and *B. tematum*. *B. pertussis* is the etiologic agent of whooping cough. *B. parapertussis* may be a variant of *B. pertussis*, although some studies suggest it is more closely related to *B. bronchiseptica*. *B. bronchiseptica* is an animal pathogen and rarely causes symptomatic disease in humans. *B. avium* and *B. hinzii* are avian pathogens. The other two species rarely cause bacteremia and wound and ear infections. *Bordetella* species are strictly aerobic, small (0.3 × 0.5 to 2.0 μm), gram-negative coccobacilli that often exhibit bipolar staining and are arranged singly or in pairs. The two human strains and *B. holmisii* are nonmotile. All species require nicotinic acid or nicotinamide and grow best at 35° to 37°C.

Calymmatobacterium granulomatis is the only species in this unassigned genus. It is a nonmotile, encapsulated, gram-negative coccobacillus (0.5 × 1 to 2 μm) that appears to be antigenically related to *Klebsiella rhinoscleromatis*. It is isolated in chicken embryos, egg yolk slants (Dulaney slants), or special fluid medium.

Streptobacillus moniliformis is one agent of rat-bite fever. It is a highly pleomorphic, catalase- and oxidase-negative facultatively anaerobic, gram-negative rod (0.1 to 0.7 × 1 to 5 μm). Single cells, chains, or filaments (10 to 150 μm in length) may show beaded central swelling. The organism grows only on medium containing blood, serum, or ascitic fluid.

The genus *Bartonella* contains 11 species, five of which cause human infections: *B. bacilliformis, B. quintana, B. henselae, B. elizabethae*, and *B. clarridgeiae*. The other six species are found in small mammals, fish, birds, and other animals (91). The genus *Afipia* consists of three species (*A. felis, A. broomeae*, and *A. clevelandensis*) and five genospecies (92). *Bartonella* species are small (0.6 to 1.0 μm), slightly curved, gram-negative rods that are oxidase-negative and relatively inert biochemically. All species require hemin or blood for growth. *Afipia* species are oxidase and urease positive and motile and do not require blood for growth.

Spectrum of Disease

F. tularensis is the etiologic agent of tularemia, which can present in the following clinical forms: ulceroglandular (80% of all cases), oculoglandular, oropharyngeal, glandular (no lesion detected), typhoidal, pleuropulmonary, GI, or combinations of these.

Brucellosis is primarily a genitourinary tract infection in animals and affects the reticuloendothelial system in humans. The disease ranges from subclinical to acute infections with the potential to develop relapses or become a chronic infection. Complications may affect almost any organ system. *Brucella* infections are often difficult to diagnose because of their chronicity and varied clinical manifestations.

Whooping cough (pertussis) is caused by *B. pertussis* and is characterized by a series of short expiratory coughs followed by an inspiratory gasp that results in the whoop. Complications are primarily owing to secondary infections. *B. parapertussis* or rarely *B. bronchiseptica* may cause a pulmonary infection or opportunistic infections in compromised hosts.

Granuloma inguinale, donovanosis, and granuloma venereum are synonyms for the disease caused by *C. granulomatis*. The disease begins as single or multiple subcutaneous nodules that progress to open granulomatous lesions that are usually painless. The genitalia are involved in approximately 90% of cases, the rectum in 5% to 10 %, and other sites in 1% to 5%.

S. moniliformis is an agent of rat-bite fever that begins with an abrupt onset of fever, chills, rash, and headache. Complications include arthritis, pneumonia, endocarditis, myocarditis, meningitis, and nephritis.

The disease caused by *B. bacilliformis* can present in two forms: Oroya fever (systemic form) and verruga peruana (cutaneous form). The organism grows within and on the surface of erythrocytes as well as in the vascular endothelium. *B. henselae* and *B. quintana* (and infrequently *B. elizabethae*) are associated with bacteremia, endocarditis, bacillary angiomatosis, and peliosis (91). Infections occur in the immunocompetent individual but are more common in immunocompromised persons, especially human immunodeficiency virus–infected. Cat-scratch disease is the most common infection caused by *Bartonella* species (91). Most, if not all, cases are caused by *B. henselae*, but some may be caused by *B. clarridgeiae* or *A. felis*.

Epidemiology

Tularemia occurs exclusively in the northern hemisphere. *F. tularensis* is found in wild mammals (e.g., rabbits, muskrats, beavers, deer, bear), domestic animals (e.g., sheep, cats, cattle, dogs), birds, some amphibians and fish, insects (e.g., ticks, deerflies, mosquitoes), and water from streams and wells. The disease is transmitted most often by contact with infected animals or bites of infected insects and less frequently by aerosols, consumption of contaminated water or improperly cooked meat, and bites of animals. Tick-borne cases occur primarily in the spring and summer months, whereas rabbit and muskrat associated cases increase in the winter.

Brucellosis is largely a disease of domestic animals throughout the world, especially in the Mediterranean countries of Europe,

Africa, Asia, and Central and South America where there are no cattle vaccination programs. The organisms, normal flora of the genital and urinary tracts of animals, are transmitted to humans by direct contact with infected animal tissue, inhalation of aerosols, or consumption of contaminated meat or dairy products. Brucellosis is an occupational disease of persons who handle livestock (e.g., veterinarians, ranchers, slaughterhouse workers).

Pertussis continues to cause significant morbidity and mortality in populations that are not vaccinated. Humans are the only known reservoir of *B. pertussis*. In nonvaccinated populations, the disease occurs primarily in children 1 to 5 years of age. Currently in the United States, approximately 50% of the cases occur in children younger than 1 year of age and 30% occur in adults.

C. granulomatis is endemic in New Guinea, India, Australia, the Caribbean, and other tropical countries. It is transmitted sexually but may be contracted by other direct contact. The disease is only mildly contagious and has a higher prevalence among individuals of low socioeconomic status and poor personal hygiene.

Wild and laboratory rats and other rodents are the primary reservoir for *S. moniliformis*, although the bacterium has been found in turkeys and milk. Transmission occurs after the bite of a rodent, handling infected tissue, or occasionally after consumption of contaminated food (e.g., milk).

B. bacilliformis is restricted to the habitat of the sandfly vector (*Phlebotomus*) in the Andes Mountains of Peru, Ecuador, and Colombia. Asymptomatic human carriers serve as a reservoir. *B. quintana* and *B. henselae* are found worldwide and transmission is associated with exposure to *Pediculus humanus* and cats, respectively (92,93). *B. clarridgeiae* is also associated with cat exposure (91).

Diagnostic Procedures

Microscopy

Gram-stained smears of material containing *B. pertussis, F. tularensis*, or *Brucella* species may appear negative because the organisms stain poorly and are small. Material should be stained with fluorescent antibody reagents for the detection of *F. tularensis* and *B. pertussis*. These stains are best performed by reference laboratories.

C. granulomatis is usually observed in a Giemsa- or Wright-stained smear of a scraping of the lesion as clusters of safety pin–shaped rods in the cytoplasm of macrophages (Donovan bodies). The organism appears blue to deep purple with a pink capsule.

S. moniliformis may be directly seen in Giemsa-, Wayson-, Gram-, or acridine orange–stained smears of clinical specimens as extremely pleomorphic gram-negative rods with long filaments, chains, and swollen cells.

The diagnosis of bartonellosis can be made by demonstrating pleomorphic, rod or ring-shaped bacteria in Wright- or Giemsa-stained blood smears.

Culture Techniques

Cystine-glucose blood agar is recommended for the isolation of *F. tularensis*, but growth is also observed on enriched chocolate, Thayer-Martin, and charcoal yeast extract agars. Occasional strains will grow on sheep blood agar. Growth has also occurred in blood culture bottles inoculated with blood or pleural fluid (94). Colonies are pinpoint, gray, and smooth after incubation for 24 to 48 hours at 35°C. *F. tularensis* is highly infectious. Gloves should be worn and cultures worked up in a biological safety hood.

Brucella species are commonly isolated from blood and bone marrow specimens plated on sheep blood, chocolate, and charcoal yeast extract (95) agars as well as selective medium for the isolation of *N. gonorrhoeae*. Commercial blood culture broths will also support growth in 5 to 21 days. Solid media should be incubated in 5% to 10% CO_2 at 35° to 37°C for 3 weeks. After 3 to 4 days, colonies are pinpoint and gray on blood agar. All suspected cultures should be handled in a biological safety cabinet.

B. pertussis was originally isolated on Bordet-Gengou medium, but this medium has been generally replaced by charcoal blood agars or Regan-Lowe medium. Both selective and nonselective media should be used. The organism produces small colonies (less than 1 mm) that are shiny and anthracite in color. Cultures should be incubated 3 to 7 days at 35° to 37°C in a moist environment without added CO_2.

C. granulomatis can be cultured but requires specialized media or eggs. The laboratory diagnosis is usually based on smear results.

Isolation of *S. moniliformis* requires the use of medium containing blood, serum, or ascitic fluid. It has been recovered on sheep blood and chocolate agars. Cultures are incubated at 35° to 37°C in 5% to 10% CO_2 and high humidity for 1 to 2 weeks. Colonies are round, gray, and glistening and remain viable for 3 to 7 days.

Bartonella has been detected in commercial blood culture systems by acridine orange staining of a smear after 7 days incubation, but more often blood is processed by lysis centrifugation and plated on enriched media (chocolate or blood agar) and incubated at 35° to 37°C in 5% to 10% CO_2 and high humidity for as long as 30 days (91). Freshly prepared heart infusion agar containing 5% to 10% rabbit or horse blood is optimal. *B. bacilliformis* and *Afipia* cultures should be incubated at 25° to 30°C.

Identification Methods

The biochemical identification of *F. tularensis* is not recommended because of the highly infectious nature of the organism. Presumptive identification is based on growth characteristics (absence of growth on routine media), morphology (small coccobacillus), oxidase negativity, and catalase positivity (weak). Specific identification is performed by fluorescent antibody or agglutination procedures and most often performed in a reference laboratory.

Brucella species are usually urease, oxidase, and catalase positive and nonmotile and reduce nitrates. Presumptive species identification is based on urease activity (*B. canis* and *B. suis* are positive in less than 30 minutes, *B. abortus* in 1 to 2 hours, and *B. melitensis* is variable), H_2S production (*B. abortus* and *B. canis* are positive), and CO_2 requirement (*B. abortus* usually requires CO_2 for growth). Definitive identification is accomplished by additional tests and agglutination in specific antiserum.

A fluorescent antibody test or agglutination with specific antisera for *B. pertussis* and *B. parapertussis* can be used for the rapid presumptive identification of suspicious colonies. Some antisera may cross-react with *Haemophilus* species.
Differentiation among the *Bordetella* species is accomplished by determination of the growth rate (fresh isolates of *B. pertussis* take 3 to 5 days, other species 1 to 2 days); urease activity (*B. pertussis* is negative); growth on blood agar (fresh isolates of *B. pertussis* do not grow); and growth on MacConkey agar (*B. bronchiseptica* is positive). Isolates resembling *B. bronchiseptica* must be differentiated from *Alcaligenes* species.
S. moniliformis is presumptively identified by its growth characteristics and Gram-stain morphology. Biochemical tests must be supplemented with serum. *S. moniliformis* does not liquify gelatin, produce indole, or reduce nitrates; nor does it contain catalase, oxidase, or urease activity.
Presumptive identification of *Bartonella* species is based on their characteristic cellular morphology, growth requirements, “twitching” motility, and negative catalase and oxidase reactions.

Direct Antigen or Nucleic Acid Detection

Nasopharyngeal specimens can be examined for *B. pertussis* and *B. parapertussis* by the DFA test. In experienced laboratories, the sensitivity and specificity of the DFA test are 60% to 65% and 90% to 99%, respectively (96,97). The detection of *B. pertussis* in nasopharyngeal specimens by DNA amplification is more sensitive and specific but is not widely available.

Serologic Procedures

Tularemia is diagnosed by a standard agglutination test indicating a fourfold rise in titer or a single titer of 1:160 or greater (indicates current or past infection). The test cross-reacts with *Brucella* at low titers.
Several tests have been employed for the serologic diagnosis of brucellosis. Antibodies against all *Brucella* species, except *B. canis*, are detected in the standard agglutination test using *B. abortus* antigens. A fourfold rise in titer or a single titer of 1:160 or greater with a compatible illness is diagnostic of brucellosis. Other serologic methods such as EIA may be useful for differentiating acute from chronic infection. Populations exposed to *Brucella* (e.g., veterinarians, slaughterhouse workers) often have baseline titers of 1:80 or 1:160.
Serologic tests for *B. pertussis* exposure have been used primarily for epidemiologic purposes but may be helpful for the diagnosis of acute infection when the patient presents late in the course of the disease.
Patients with rat-bite fever usually have an agglutination titer of 1:80 or greater or demonstrate a fourfold rise in titer. The test is performed only at national reference laboratories. No serologic tests are commercially available for the diagnosis of granuloma inguinale or bartonellosis. Serologic tests are available for the diagnosis of infections caused by *B. bacilliformis, B. quintana*, and *B. henselae*. Cross-reactions have occurred with *Coxiella* and *Chlamydia* and between *B. quintana* and *B. henselae*.

Interpretation of Results

The isolation or direct identification of any of the aforementioned bacterial species (except *B. parapertussis* and *B. bronchiseptica*) from patients with a compatible illness is clinically significant.

Therapeutic Considerations

The organisms discussed in this section have special growth requirements that make *in vitro* susceptibility testing difficult. Therefore, most recommendations for therapy are based on clinical response. Streptomycin is the drug of choice for the treatment of tularemia, although some data suggest that gentamicin may be an alternative. The use of tetracycline or chloramphenicol is associated with a higher relapse rate. Limited data suggest that the fluoroquinolones, imipenem, or rifampin may be useful.
Because brucellosis is difficult to eradicate, some investigators recommend the use of tetracycline plus streptomycin or gentamicin. Aminoglycosides, rifampin, trimethoprim-sulfamethoxazole, chloramphenicol, and fluoroquinolones have also been used, but single-drug therapy is associated with a 10% to 40% relapse rate. Therapy should be continued for 3 to 6 weeks.
B. pertussis infections can be treated with erythromycin, trimethoprim-sulfamethoxazole, or chloramphenicol. Few clinical data are available on the efficacy of other antimicrobials.
C. granulomatis is susceptible to tetracycline, sulfonamide, trimethoprim-sulfamethoxazole, chloramphenicol, and gentamicin. The drug of choice for the treatment of rat-bite fever is penicillin. Other alternatives include chloramphenicol, tetracycline, and streptomycin. Most infections caused by *B. quintana* or *B. henselae* respond to treatment with erythromycin or doxycycline. Alternative drugs include other macrolides and fluoroquinolones. *B. bacilliformis* infections are usually treated with chloramphenicol.

Summary and Conclusions

The diseases discussed in this section are infrequently observed, and the laboratory diagnosis is difficult because of the special growth requirements of the etiologic agents. The clinician must alert the laboratory to the possibility of infection by one of these agents so that appropriate culture procedures will be used. The laboratory must also exclude these bacteria when unusual isolates are seen on Gram smear or isolated.

ANAEROBIC BACTERIA

Part of "51 - Bacteriology"

Microbiology

Anaerobic microorganisms are a genetically and phenotypically heterogeneous group that includes both aerotolerant and anaerobic strains. Many of the genera have or will undergo significant taxonomic revision. For example, Collins et al. (98) have proposed five new genera of spore-forming rods. Changes in the taxonomy are expected to continue, but fortunately for the clinical microbiology laboratory, relatively few species cause human disease. A list of the medically important genera of anaerobic bacteria based on Gram-stain morphology is provided in Table 51.26.

TABLE 51.26. GENERA OF MEDICALLY IMPORTANT ANAEROBIC BACTERIA

Gram-negative rods
Bacteroides
Fusobacterium
Prevotella
Porphyromonas
Bilophilia
Wolinella
Sutterella
Oribaculum
Johnsonella
Hallella
Campylobacter
Catonella
Dialister
Gram-negative cocci
Veillonella
Gram-positive rods
Clostridium
Actinomyces
Propionibacterium
Bifidobacterium
Eubacterium
Lactobacillus
Mobiluncus
Gram-positive cocci
Peptostreptococcus
Streptococcus
Staphylococcus
Gemella

The anaerobic gram-negative bacilli are straight, curved, or helical and motile or nonmotile. They are among the most common components of the normal flora of the human mucous membranes. Among these genera, *Bacteroides, Fusobacterium, Prevotella*, and *Porphyromonas* are isolated most frequently from infections.
Of the gram-negative cocci, only three of the seven *Veillonella* species are occasionally isolated from human infections. *Megasphaera* and *Acidaminococcus* are seldom associated with infections in humans.
The genera of anaerobic gram-positive bacilli consist of *Clostridium* (spore formers) and seven genera of non-spore formers (*Actinomyces, Propionibacterium, Bifidobacterium, Eubacterium, Atopobium, Mobiluncus*, and *Lactobacillus*). *Mobiluncus* species stain gram variable to gram negative. Some species are aerotolerant and vary in their morphology with the type of medium and growth conditions, which cause confusion with the aerobic bacilli and streptococci. In general, the aerotolerant anaerobes grow better anaerobically (produce larger colonies) than aerobically, and the aerotolerant clostridia (catalase negative) rarely produce spores aerobically. *Bacillus* species are catalase positive and rarely produce spores anaerobically.
The anaerobic gram-positive cocci are the second most common group of anaerobes encountered in clinical infections. Major taxonomic changes (Table 51.27) have occurred recently with these genera and several new species have been described. All pathogenic *Peptococcus* species (except *P. niger*) were transferred to the genus *Peptostreptococcus*. Other anaerobic cocci were transferred to the genus *Streptococcus*, and *P. saccharolyticus* was transferred to the genus *Staphylococcus*. Some of the many nomenclature changes that have occurred in recent years are listed in Table 51.27.

TABLE 51.27. RECENT TAXONOMIC CHANGES OF ANAEROBIC BACTERIA

Current Nomenclature	Previous Nomenclature
Bacteroides distasonis	*Porphyromonas*
Bacteroides furcosus	*Porphyromonas*
Bacteroides putredinis	*Rikenella*
Campylobacter gracilis	*Bacteroides gracilis*
Campylobacter showae	*New species*
Campylobacter cuvus	*Wollinella cuvus*
Campylobacter rectus	*Wollinella rectus*
Dialister pneumosintes	*Bacteroides penumosintes*
Johnsonella ignava	*Clostridium*
Porphyromonas catoniae	*Oribaculum catoniae*
Prevotella dentalis	*Mitsuokella dentalis*
Prevotella enoeca	New species
Prevotella pallens	New species
Prevotella tannerae	New species
Sutterella wadsworthensis	New genus and species

Spectrum of Disease

Anaerobic bacteria are recovered from all types of infections in any organ of the body (99). The most frequently encountered anaerobic infections or toxin-related diseases include pleuropulmonary infection (e.g., pneumonia, abscess, empyema), intraabdominal infection (e.g., abscess, appendicitis, cholecystitis, peritonitis), antibiotic-associated colitis, and female genital tract infection (e.g., abscess, pelvic inflammatory disease, endometritis, BV). Other types of infections include head and neck infections (e.g., brain abscess, meningitis, sinusitis, otitis media, dental and periodontal infection), soft-tissue infections (cellulitis, myonecrosis), bone and joint infections, bacteremia, endocarditis, bite wounds, botulism (including infant and wound), tetanus, and food poisoning.

Epidemiology

Anaerobic infections caused by non–spore-forming organisms generally arise from the host's indigenous flora after a break in the skin or mucosal barrier (e.g., trauma, surgery, implantation of a device). The resulting tissue damage and the presence of microaerophiles or facultative anaerobes act to produce the appropriate environment for the growth of anaerobes and explains why most deep-seated abscesses and necrotizing lesions contain multiple species of bacteria.
Infections or intoxications caused by spore-forming bacilli (*Clostridium*) are usually transmitted via improperly prepared or stored food (e.g., botulism, *C. perfringens* food poisoning), contamination of damaged tissue with soil or bowel contents (e.g., clostridial myonecrosis, tetanus), or contaminated objects or medical personnel in hospitals (e.g., antibiotic-associated colitis caused by *C. difficile*).

Diagnostic Procedures

This section provides a limited overview of the procedures that may aid in the isolation or preliminary identification of anaerobic bacteria. The extent to which anaerobes are identified depends on the available facilities, the technical competence of the personnel, and the clinical need for the information (e.g., the detection of the agent or toxin responsible for the disease may indicate a therapy change or the need for an antimicrobial susceptibility test). The selection, collection, and transport of specimens are crucial to clinically relevant laboratory results. The workup of inappropriate or poorly transported specimens is a waste of effort and money. The presence of anaerobes in a specimen is suggested when the material has a foul odor or black discoloration, contains sulfur granules, or fluoresces brick-red under long-wave ultraviolet light (366 nm). All laboratories should be able to isolate and presumptively identify anaerobic isolates based on cellular and colony morphology and a few essential characteristics. More definitive identification of common isolates can be accomplished with the use of commercial kits, and other significant isolates can be submitted to a reference laboratory for identification.

Microscopy

The Gram stain can provide useful clinical information on the presence of anaerobes and suggest additional media to set up based on the characteristic morphology of some species. The *Bacteroides fragilis* group may be uniform or pleomorphic with a safety-pin appearance (Fig. 51.17), whereas the pigmented species are often coccobacillary (*Haemophilus*-like) and can be mistaken for gram-negative cocci (Fig. 51.18). *Fusobacterium nucleatum* is usually very thin with tapered ends (spindle shaped), often appearing end to end in pairs (Fig. 51.19). *F. mortiferum* and *F. necrophorum* are filamentous and pleomorphic, with swollen areas and large round bodies. *Actinomyces* appear as thin, branching gram-positive rods with beaded staining. *Bifidobacterium* species often exhibit clubbing or terminal bifurcation. Peptostreptococci appear as clusters and chains of irregularly stained, somewhat pleomorphic, gram-positive cocci. *P. magnus* is usually large, and *P. tetradius* occurs in tetrads. *C. perfringens* is a gram-positive rod (may appear gram-negative) with square ends and usually does not contain a spore (Fig. 51.20). The presence of clostridia and absence of inflammatory cells in a Gram smear of a wound is highly suggestive of clostridial myonecrosis.

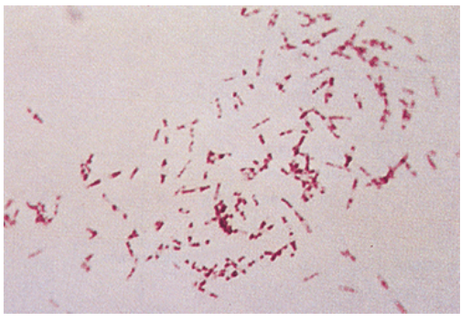


FIGURE 51.17. Gram stain of *Bacteroides fragilis*.

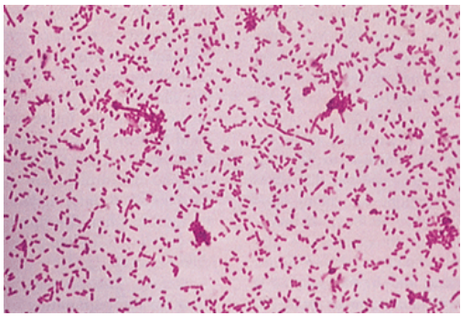


FIGURE 51.18. Gram stain of pigmented *Prevotella* species.

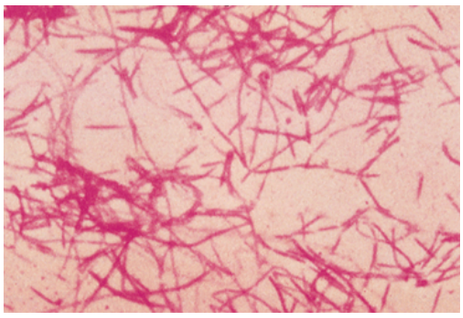


FIGURE 51.19. Gram stain of *Fusobacterium nucleatum*. Fusiform shape. (From Baron EJ, Finegold SM. *Bailey and Scott's diagnostic microbiology*. St. Louis: Mosby, 1990, with permission.)

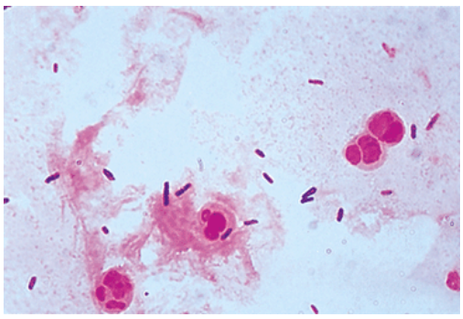


FIGURE 51.20. Gram stain of body fluid containing *Clostridium perfringens*.

Culture Techniques

A combination of enriched, prereduced, nonselective (e.g., *Brucella* sheep blood agar), selective (e.g., *Bacteroides* bile esculin, kanamycin-vancomycin–laked blood, phenylethyl alcohol agars), and enriched thioglycollate media should be used for the recovery and presumptive identification of anaerobes. Primary plates are incubated under anaerobic conditions at 35°C for 48

hours before examination and are reincubated for 5 to 7 days. In serious infections or suspected actinomycosis, the broth medium should be held for 2 to 4 weeks. During the examination of primary plates and subculture of anaerobic isolates, a holding jar should be used to minimize the exposure of the isolates to oxygen. Colony morphology is examined with a dissecting microscope or a hand-held lens, and each type Gram-stained and subcultured for incubation in CO_2 and anaerobically. Gram-positive cocci that are capnophiles may not grow in CO_2 on initial subculture and require a second subculture to demonstrate growth. The thioglycollate medium should be Gram stained and subcultured if morphologic types appear on smear that are not seen on the primary plates. The cellular and colony morphology on the primary plates can be suggestive of some groups of organisms (Table 51.28). Cycloserine-cefoxitin-fructose agar is one medium that is recommended for the isolation of *C. difficile* from stool specimens.

TABLE 51.28. MICROSCOPIC AND COLONIAL MORPHOLOGY OF SELECTED ANAEROBES

Organism	Medlium[a]	Microscopic Morphology	Colonial Morphology
Bacteriodes fragilis group	BBE	gnb, pleomorphic	>1 mm ± esculin hydrolysis
Bacteroides spp	KVLB	gnb	Nonpigmented
Porphyromonas or *Prevotella* spp	KVLB	cb, gnb	Brown to black or fluorescing brick-red
Bacteriodes ureolyticus	BA	gnb	Pitting
Bacteriodes fragilis group	BA	gnb, pleomorphic	Circular, grayish-white translucent to opaque
Pigmenting species	BA	cb, gnb	Circular, darker in center with brown edges
Fusobacterium nucleatum	BA	Fusiform gnb with pointed ends	Opalescent and speckled, fluoresces chartreuse
Fusobacterium necrophorum	BA	pleomorphic gnb, many contain round bodies	Convex to umbonate, fluoresces chartreuse
Actinomyces israelii	BA	branching gpb	Molar tooth
Clostridium perfringens	BA	gpb, square ends	Double zone of ß-hemolysis
Clostridium difficile	CCFA	gpb	Yellow colonies that may fluoresce
Clostridium septicum	BA	gpb ± spores	Swarming, ß-hemolysis medusa head
Clostridium tetani	BA	gpb, “drumstick”	May swarm, narrow zone ß-hemolysis
Peptostreptococcus	BA	gpc	Grow only on BA after 48 hr
Veillonella parvula	BA	gnc	Fluoresce red

BBE, *Bacteroides* bile esculin-agar; KVLB, kanamycin-vancomycin laked blood agar; BA, *Brucella* blood agar: CCFA, cycloserine-cefoxitin-fructose agar; gnb, gram-negative bacillus; cb, coccobacillus; gpb, Gram-positive bacillus; gpc, Gram-positive coccus.

Identification Methods

The preliminary or presumptive identification of anaerobes can be accomplished by using some special tests such as colonial fluorescence, susceptibility to special-potency antibiotic and polyanethol-sulfonate disks, and catalase, urease, and nitrate reduction (Table 51.29) (100,101 and 102). Dowell and Lombard (103) devised a modified conventional plate system (presumpto method) that can be used for the preliminary identification of anaerobes. The definitive identification of anaerobes by reference laboratories includes conventional biochemical tests and gas-liquid chromatographic analysis of fatty acid metabolites and is beyond the scope of this discussion. For the nonreference laboratory, several commercial systems are available that will correctly identify 70% to 90% of the more commonly encountered anaerobes and offer a rapid alternative to the conventional method.

Direct Antigen or Nucleic Acid Detection

Gas-liquid chromatographic analysis of specimens can presumptively detect the presence of *Bacteroides, Prevotella, Porphyromonas*, or *Fusobacterium* species but is not widely used. Currently, there are no commercial nucleic acid probes for the detection or identification of anaerobes.

The laboratory diagnosis of antibiotic-associated colitis (AAC) is made by demonstrating the organism, glutamate dehydrogenase, or toxins in the stool specimen of a symptomatic patient. The recovery of *C. difficile* by culture is reliable test for AAC if the isolate is tested for toxin production. Commercial EIA tests for the detection of toxin A or toxin A and B are also available but vary in their accuracy (104). The detection of *C. difficile* toxins by cytopathic effect in tissue culture remains the gold standard. An EIA test that detects glutamate dehydrogenase is rapid and comparable with toxin detection by EIA or cytotoxin neutralization assays (105). A latex test for glutamate dehydrogenase lacks sensitivity (104).

The detection of botulinal neurotoxin in the patient's vomitus, serum, feces, or consumed food should be performed by the State Health Department or CDC.

Serologic Procedures

Serologic procedures are not important tests for the diagnosis of anaerobic infections.

TABLE 51.29. PRELIMINARY IDENTIFICATION OF ANAEROBES

Group/Species	Antibiotic Disk Pattern			Characteristics
	K (1 mg)	Co (10 µg)	Va (5 µg)	
Bacteriodes fragilis group	R	R	R	Bile, R; catalase, V
Pigmenting species	R	V	V	Tan to black colonies[a]
				Brick-red fluorescence[b]
Prevotella spp	R	V	R	Brown to black colonies[a]
Porphyromonas spp	R	R	S	
Other *Bacteriodes* spp	R	V	R	
Bacteroides ureolyticus	S	S	R	Colonies pit agar or require formate/fumarate for growth; nitrate, +, urease +
Bilophila spp	S	S	R	Urease, +; motile -; catalase +; nitrate +
Fusobacterium spp	S	S	R	Colonies generally larger and more opaque than *B. ureolyticus* group
F. nucleatum	S	S	R	Slender cells with pointed end. Indole, +; greens agar; yellow-green fluorescence; three possible colony types
F. necrophorum	S	S	R	Indole, +; lipase, +[a] Umbonate colony[b]
Gram-negative cocci	S	S	R	Three genera; requires GLC, if not *Veillonella*
Veillonella spp	S	S	R	Nitrate, + Small coccus
Gram-positive cocci	V	R	S	
Peptostreptococcus anaerobius	R	R	S	Sodium polyanethol sulfonate, S
Peptostreptococcus asaccharolyticus	S	R	S	Indole, +
Clostridium spp	V	R	S	Spores present: may appear gram-negative
C. perfringens	S	R	S	Double-zone ß-hemolysis. Cells boxcar shape; spore test, V; reverse CAMP test, +; Nagler test, +
Non-spore-forming bacilli	V	R	S	
Propionibacterium acnes	S	R	S	Indole, +[a], catalase, +[a]; nitrate, +[b]; cells may show short branching
Eubacterium lentum	S	R	S	Cells and colonies small; arginine required for growth in broth: nitrate, +

[a] Not all strains positive; if negative, more tests required.

[b] Not all strains positive.

R, resistant; S, susceptible; V, variable; +, positive; -, negative; GLC, gas-liquid chromatography.

Interpretation of Results

The significance of the isolation of anaerobes that are a part of the indigenous flora depends on the quality of the specimen submitted and the subsequent Gram stain and must be interpreted in light of the clinical picture. The presence of botulinal neurotoxin is diagnostic of botulism.

Tests used for the laboratory diagnosis of AAC should be interpreted in conjunction with the clinical setting of the patient. The use of two tests can often improve the detection of affected patients.

Therapeutic Considerations

For most institutions, antimicrobial susceptibility testing of anaerobic isolates should be limited to circumstances related to the seriousness of the infection, lack of response to empiric therapy, or lack of established data with a specific agent or to monitor susceptibility patterns of local isolates. The susceptibility pattern of specific anaerobes cannot always be predicted because of the increasing resistance to ß-lactam drugs and the differing degrees of effectiveness of the newer agents.

Metronidazole (inactive against *Propionibacterium* and *Actinomyces*), chloramphenicol, imipenem, and the ß-lactam plus ß-lactamase inhibitor drugs are active against most anaerobes. Agents that are usually active include clindamycin (increasing resistance among *B. fragilis* group and clostridia), cefoxitin (increasing resistance among *B. fragilis* group and poor activity against clostridia), and antipseudomonal penicillins. Penicillin, cephalosporins (other than cefoxitin and cefotetan), tetracycline, macrolides, and vancomycin (active against gram-positive anaerobes) have variable activity. The aminoglycosides, monolactams, and most quinolones have poor activity against anaerobes.

Summary and Conclusions

Anaerobic bacteria are common, indigenous organisms found on the mucosal surfaces of humans that cause a variety of infections in any tissue or organ system. The most important aspect of anaerobic microbiology is to ensure the appropriate selection, collection, transport, and inoculation of the specimen. The extent of isolate identification depends on the seriousness of the infection, the clinical need for information, the technical expertise of the laboratory personnel, and the availability of resources.

SPIROCHETES

Part of "51 - Bacteriology"

Microbiology

The order Spirochaetales consists of two families. The family Spirochaetaceae contains the genera *Spirochaeta, Cristispira, Treponema, Brevinema*, and *Borrelia*, and the family Leptospiraceae contains two genera: *Leptospira* and *Leptonema*. Spirochetes of medical importance are in the genera *Treponema, Borrelia*, and

Leptospira. The pathogenic species are restricted to a living host and are heterogeneous in their physiology and habitat. Common features include spiral morphology and unique flexuous motility related to axial filaments located in the periplasmic space between the outer membrane and protoplasmic cylinder. In general, the spirochetes are slender, flexible, helix-shaped bacteria that exhibit a corkscrewlike motility and can be loosely grouped into genera by their morphology (Fig. 51.21).

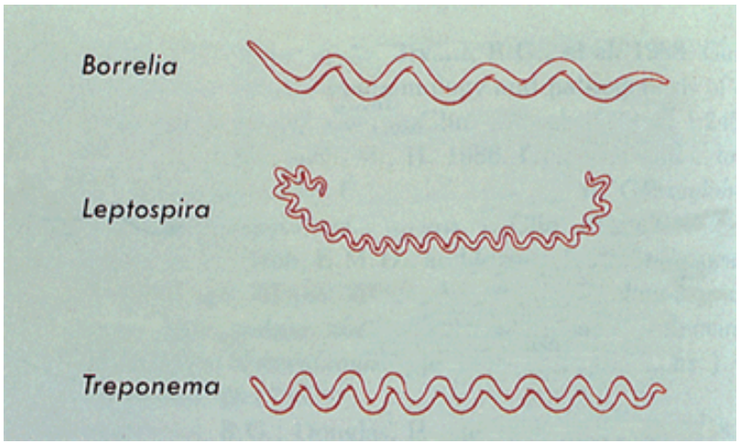


FIGURE 51.21. Schematic designation of spirochetes based on general morphology. (From Baron EJ, Finegold SM. *Bailey and Scott's diagnostic microbiology*. St. Louis: Mosby, 1990, with permission.)

The genus *Treponema* currently contains 12 species. The pathogenic species are 6 to 20 µm in length and 0.1 to 0.2 µm in width with regularly spaced coils (six to 14 per cell) and a rotational back-and-forth (corkscrew) motility with some bending and snapping.
The treponemes pathogenic for humans can be continuously cultured only by animal passage and consist of two species: *T. carateum* (causes pinta) and *T. pallidum*, which contains three subspecies: *pallidum* (causes syphilis), *pertenue* (causes yaws), and *endemicum* (causes nonvenereal endemic syphilis or bejel).
The genus *Borrelia* contains 36 species. This classification may not be correct and awaits further clarification by genetic relatedness techniques. *B. burgdorferi* (Lyme disease) has been divided into two additional species, *B. garinii* and *B. afzelii* (106). *Borrelia* are 0.2 to 0.5 × 5 to 25 µm, have three to 10 loose coils, and grow well from 30° to 35°C in specialized media.
The genus *Leptospira* contains two new species: *L. biflexa*, which is a nonpathogenic, free-living spirochete with more than 60 serovars, and *L. interrogans*, which causes leptospirosis in animals and humans with more than 218 serovars (107). Molecular techniques suggest grouping the *L. interrogans* serovars into seven named and five unnamed species (108). *Leptospira* are 6 to 20 × 0.1 to 0.2 µm with very tightly wound coils that may be hooked at one or both ends. They are aerobic and require particular nutrients (e.g., long-chain fatty acids) and a specific pH (7.0 to 7.8) for growth at 28 to 30°C.

Spectrum of Disease

T. pallidum subsp *pallidum* is the etiologic agent of syphilis, which presents in three stages: primary (a chancre at the site of entry), secondary (flulike illness with generalized rash), and tertiary (lesions in the central nervous, cardiovascular, or any other organ system). Infection of the fetus (congenital syphilis) can occur in untreated women.
Yaws, pinta, and endemic syphilis are distinguished from syphilis based on their clinical manifestations and geographic area of occurrence. Yaws occurs primarily in rural populations of tropical countries; pinta is found in remote areas of Central and South America; and endemic syphilis primarily affects children in the arid areas of Africa and the Middle East.
B. recurrentis is the etiologic agent of louse-borne (epidemic) relapsing fever, whereas other *Borrelia* species cause tick-borne (endemic) relapsing fever. *B. burgdorferi, B. garnii*, and *B. afzelii* are associated with Lyme disease.
Leptospirosis is primarily a disease of animals that is occasionally transmitted to humans. The clinical manifestations of the infections range from a mild febrile illness to a fulminating fatal disease.

Epidemiology

Syphilis occurs worldwide and is generally transmitted through sexual contact or transplacentally. Infections may also be acquired through nonsexual contact with lesions or blood. The highest prevalence of the disease is associated with economically deprived individuals engaged in prostitution or sexual activity with multiple partners.
Louse-borne (epidemic) relapsing fever (*B. recurrentis*) is transmitted by the human body louse (*Pediculus humanus*). The infective organisms penetrate intact skin or mucous membranes. Humans are the only hosts for *B. recurrentis* and *B. drottonii*. Disease usually occurs from overcrowding such as during famine or wars. The disease is endemic in Central and East Africa and in Bolivia and Peru.
Tick-borne (endemic) relapsing fever (*Borrelia* spp) is transmitted to humans by infected soft-shelled ticks of the genus *Ornithodorus*. Infection occurs when saliva or coxal secretions are released by feeding ticks. The principal reservoirs are rodents and small animals as well as the ticks that pass the spirochete transovarially.
Lyme disease (*B. burgdorferi, B. garnii*, and *B. alzelii*) occurs worldwide and is transmitted primarily by the *Ixodes* species of ticks. The principal reservoirs are rodents and large mammals (e.g., deer). The disease can be spread by the adult, larval, and nymphal stage of the tick.
Leptospirosis is a zoonosis that occurs worldwide and affects many different species of animals. In the United States, the most common sources are dogs, livestock, rodents, wild mammals, and cats. Humans are indirectly infected by contact with urine or with contaminated water or soil or by direct contact with animals. Leptospirosis is an occupational hazard among farmers, meatpackers, sewer workers, and veterinarians and is a recreational hazard for campers and swimmers.

Diagnostic Procedures

Microscopy

In primary, secondary, and early congenital syphilis, the treponemes may be detected in serous fluid from the lesion by dark-field

microscopic examination (Fig. 51.22) or by performance of a DFA test. The dark-field examination is not appropriate for oral or rectal lesions because of the presence of morphologically similar nonpathogenic treponemes. The DFA test or silver stain can be used to detect treponemes in tissue.

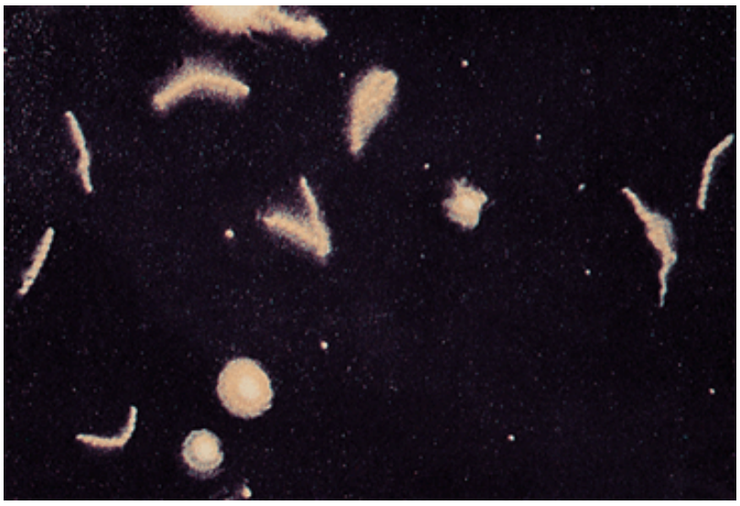


FIGURE 51.22. Appearance of *Treponema pallidum* in a dark-field preparation. (From Baron EJ, Finegold SM. *Bailey and Scott's diagnostic microbiology*. St. Louis: Mosby, 1990, with permission.)

The diagnosis of borreliosis is made in 70% of cases by examination of blood collected during a febrile episode by dark-field microscopy, Giemsa- or Wright-stained thick and thin smears (Fig. 51.23), or acridine orange–stained smears. Rarely, spirochetes are detected in the CSF of patients with central nervous system symptoms. Because of the scarcity of spirochetes in clinical specimens, direct examination is not recommended for the diagnosis of Lyme disease.

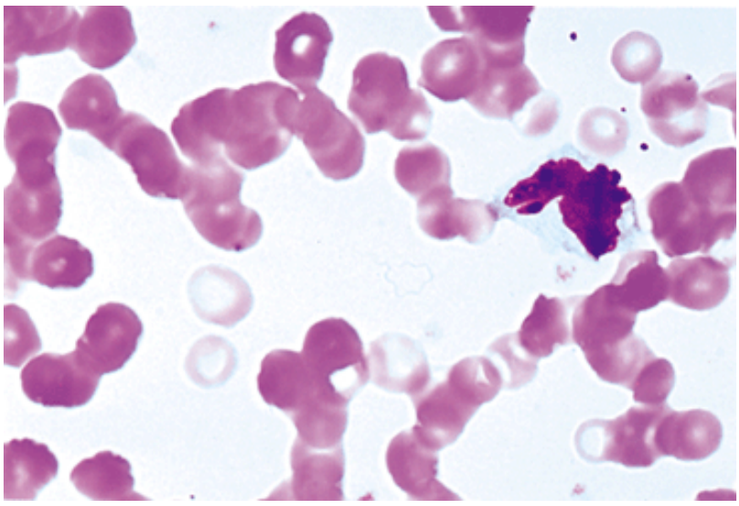


FIGURE 51.23. Wright-stained peripheral blood smear demonstrating *Borrelia*.

The direct examination of blood, CSF, or urine by dark-field examination has low sensitivity for samples with few leptospires, as has the DFA test or silver stain on tissue.

Culture Techniques

Isolation of *Borrelia* species in the routine clinical laboratory by *in vitro* cultivation is not recommended because of the low yield. Although improvements in the cultivation of *B. burgdorferi* from clinical specimens have occurred, the yield remains low and probably should not be attempted by the routine clinical laboratory.

Leptospira is cultivated in two types of media: those supplemented with bovine serum albumin and polysorbate 80 (Ellinghausen, McCullough, Johnson, Harris [EMJH] medium) or those enriched with rabbit serum (Fletcher or Stuart) or PLM-5TM, a proprietary medium similar to EMJH (108). The EMJH-based formulations are probably the best. Approximately two to three drops of blood (not citrated) or CSF collected during the first week of illness are added to each of three to five tubes of medium. One or two drops of undiluted and 1:10 dilution of urine collected after the first week is added to each of three to five tubes of medium with and without 200 µg/mL of 5-fluorouracil. The cultures are incubated in the dark for 6 weeks at 30°C. Tubes are examined weekly for growth in the form of a ring 1 to 3 cm below the surface, and organisms can be identified by dark-field microscopy.

Identification Methods

Diagnosis of syphilis depends on the observation of the organisms in lesions by dark-field examination or DFA test or by the serologic detection of antibodies.

Relapsing fever is diagnosed by the demonstration of organisms in peripheral blood by dark-field examination or by Giemsa- or Wright-stained smears. Lyme disease is recognized by serologic tests.

The isolation of *Leptospirra* in culture or the detection of antibodies by serologic tests aids in the laboratory diagnosis. Identification of *Leptosprira* species should be performed by a reference laboratory.

Direct Antigen or Nucleic Acid Detection

The DFA test is commercially available for the direct detection of *T. pallidum* in clinical specimens. Polymerase chain reaction tests for the diagnosis of Lyme disease, leptospirosis, and syphilis have been developed but are not generally available from commercial sources.

Serologic Procedures

Serologic tests for the diagnosis of syphilis measure two types of antibodies. The nontreponemal (nonspecific) tests measure both immunoglobulin G and M antilipid antibodies formed in response to lipoidal antigen released from damaged host cells and from the treponemes themselves. The treponemal (specific) tests measure antibodies directed against cellular components of the organism.

There are four standard nontreponemal tests: (a) Venereal Disease Research Laboratory (VDRL) slide, (b) rapid plasma reagin (RPR), (c) unheated serum reagin, and (d) toluidine red unheated serum test. All the nontreponemal tests have approximately the same sensitivity and specificity. Of these tests, the RPR is the most widely used in the United States, and the VDRL is the only accepted nontreponemal test for detection of antibodies

in the CSF of patients with suspected neurosyphilis. The nontreponemal tests are used primarily as screening tests and to follow patients on therapy. The sensitivity of the test varies with the stage of the disease (Table 51.30). In the early primary stage, the nontreponemal test usually becomes positive 1 to 4 weeks after appearance of the chancre (109). In congenital syphilis, a rising titer during a 6-month period is diagnostic, whereas passively transferred antibodies should not be detected after 3 months. The nontreponemal test becomes nonreactive in primary syphilis approximately 1 year after therapy and in secondary syphilis 2 years after treatment. A low titer may persist after 2 years in approximately 50% of patients treated in the latent or late stage of the infection (110). The specificity of these tests is approximately 98% overall but varies with the population. Causes of biological false positives include a variety of infectious diseases, injection drug use, pregnancy, and autoimmune disease.

TABLE 51.30. Reactivity of Serologic Tests in untreated syphilis

	Stage of Disease		
Test	**Primary**	**Secondary**	**Latent**
Nontreponemal			
(VDRL, RPR)	70–90[a]	100	70–90
Treponemal			
FTA-Abs	70–100	100	96–100
MHA-TP	70–90	100	96–100

[a] Percentage reactive.

VDRL, Venereal Disease Research Laboratory; RPR, rapid plasma reangin; FTA-Abs, fluorescent treponemal antibody absorbed; MHA-TP, microhem-agglutination assay-*Treponema pallidum*.

There are two principal treponemal tests: (a) the fluorescent treponemal antibody-absorbed and (b) the microhemagglutination assay for *T. pallidum*. These tests are used primarily to confirm a reactive nontreponemal test or to diagnose late-stage syphilis and usually remain positive for life. The range of sensitivity of the tests is shown in Table 51.30, and the specificity is 98% to 100%.

Several newer EIA tests for the diagnosis of syphilis are commercially available (109). They include nontreponemal, treponemal, and formats for the detection of immunoglobulin M in cases of congenital syphilis. In general, the sensitivity and specificity of these tests are similar to those of the older methods. The Western blot format is performed by the Syphilis Diagnostic Immunology Activity Laboratory at the CDC for diagnosis of adult and congenital syphilis.

Several different serologic tests are available for the serodiagnosis of relapsing fever. Unfortunately, these tests have limited success because of the antigenic variability observed with these organisms and the lack of test specificity.

The methods used to diagnose Lyme disease are the indirect fluorescent antibody (IFA) and EIA tests: the latter test appears to be easier to use and may be more sensitive. Both tests can be used to detect antibodies in serum, CSF, or synovial fluid. The problem with both tests is the low sensitivity early in the infection (within 3 weeks) and the lack of specificity. The Western immunoblot method is considered a necessary confirmatory test for positive EIA or IFA screening tests (111). Cross-reactivity among the *Borrelia* species and *T. pallidum* also occurs, but syphilis can be excluded with a nontreponemal test.

Diagnosis of leptospirosis is usually made by serologic tests that include the macroscopic agglutination, indirect hemagglutination, EIA, and dipstick tests for the detection of antibodies (108). Antibodies usually appear after 1 week and peak after 3 to 4 weeks. Titers may remain elevated for years. The microscopic agglutination test is sensitive and serovar specific. A rising titer or a single titer of 1:800 or greater with compatible clinical findings is indicative of the disease.

Interpretation of Results

The laboratory diagnosis of syphilis is based on the microscopic demonstration of spirochetes with the appropriate morphology in clinical specimens and a reactive serologic test. Problems in diagnosis occur if nonpathogenic spirochetes are present in the lesion or if the person has a biological false-positive test. The laboratory results need to be evaluated in light of the patient's symptoms.

A positive microscopic preparation is diagnostic for relapsing fever. The serodiagnosis of Lyme disease remains a problem mainly because of the lack of standardization of reagents and cutoff values. Inconsistent interlaboratory results occur.

Leptospirosis is diagnosed by the isolation of the organism or serodiagnosis. The possibility exists for a false-negative serologic test if the antigen pool used does not contain antigens to the infecting serovar.

Therapeutic Considerations

Penicillin remains the drug of choice for immunocompetent patients with syphilis. Alternatives include doxycycline, tetracycline, and ceftriaxone. In the immunocompromised patient, higher doses and longer duration of treatment may be required.

Relapsing fever can be treated with doxycycline or erythromycin. Penicillin is associated with slower clearance and more relapses. Lyme disease is treated with penicillin, doxycycline, amoxicillin, cefuroxime, or clarithromycin. The choice of therapy depends on the stage of the disease.

There is some controversy as to the efficacy of therapy for leptospirosis, but based on clinical and experimental data, the disease responds to treatment with doxycycline or intravenous penicillin.

Summary and Conclusions

Pathogenic spirochetes belong to three genera: *Treponema, Borrelia*, and *Leptospira*, which cause syphilis, relapsing fever and Lyme disease, and leptospirosis, respectively. Humans are the only reservoir for *T. pallidum*, whereas species of the other two genera are found primarily in rodents and larger mammals. Syphilis is transmitted through sexual contact or transplacentally; borreliosis via arthropod vectors (ticks); and leptospirosis via contaminated water and soil. With the exception of *Leptospira*, the other organisms cannot be isolated easily in the clinical laboratory. The diagnosis depends on demonstration of the spirochete in clinical specimens (e.g., dark-field microscopy,

DFA testing, and Giemsa or Wright stains) or antibodies by serologic tests. In general, spirochetal diseases can be treated with penicillin or tetracycline.

REFERENCES

1. Sneath PHA, Mair NS, Sharpe ME, et al., eds. *Bergey's manual of systematic bacteriology*. Baltimore: Williams & Wilkins, 1986:2.
2. Kloos WE, Ballard DN, George CG, et al. Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and the new species *Macrococcus equipericicus* sp. nov., *Macrococcus bovis* sp. nov., and *Macrococcus carouselicus* sp. nov. *Int J Syst Bacteriol* 1998;48:859–877.
3. Kloos WE, Bannerman TL. *Staphylococcus* and *Micrococcus*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:264–282.
4. Parsonnet J, Hickman RK, Eardley DD, et al. Induction of human interleukin-1 by toxic shock syndrome toxin-1. *J Infect Dis* 1985;151:514–522.
5. Gatterman SG, Crossley KB. Urinary tract infections. In: Crossley KB, Archer GL, eds. *The staphylococci in human disease*. New York: Churchill-Livingstone, 1997:493–508.
6. Talen DA, Staatz D, Staatz A, et al. *Staphylococcus intermedius* in canine gingiva and canine inflicted human wound infections: laboratory characterization of a newly recognized zoonotic pathogen. *J Clin Microbiol* 1989;27:78–81.
7. Muller E, Takeda S, Shiro H, et al. Occurrence of capsular polysaccaride/adhesion among clinical isolates of coagulase-negative staphylococci. *J Infect Dis* 1993;168:1211–1218.
8. Archer GL. *Staphylococcus epidermidis* and other coagulase-negative staphylococci. In: Mandell GL, Douglas RG Jr, Bennett JE, eds. *Principles and practice of infectious disease*, 3rd ed. New York: Churchill-Livingstone, 1990:1511–1518.
9. Pfaller MA, Herwaldt LA. Laboratory, clinical, and epidemiological aspects of coagulase-negative staphylococci. *Clin Microbiol Rev* 1988;1:281–299.
10. Koneman EW, Allen SD, Janda WM, et al., eds. *The gram-positive cocci: part I: staphylococci and related genera*. Philadelphia: Lippincott–Raven, 1997:539–576.
11. Fournier JM, Boutonnier A, Bouver A. *Staphylococcus aureus* strains which are not identified by rapid agglutination methods are of capsular serotype 5. *J Clin Microbiol* 1989;27:1372–1374.
12. Davis TE, Fuller DD. Direct identification of bacterial isolates in blood cultures by using a DNA probe. *J Clin Microbiol* 1991;29:2193–2196.
13. Wenzel RP, Edmond MB. Vancomycin-resistant *Staphylococcus aureus*: infection control considerations. *Clin Infect Dis* 1998;27:245–251.
14. Garrett DO, Jochimsen E, Murfitt K, et al. The emergence of decreased susceptibility to vancomycin in *Staphylococcus epidermidis*. *Infect Control Hosp Epidemiol* 1999;20:167–170.
15. Strausbaugh L. Vancomycin-intermediate *Staphylococcus epidermidis*: curio or omen? *Infect Control Hosp Epidemiol* 1999;20:163–165.
16. Centers for Disease Control and Prevention. Interim guidelines for prevention and control of staphylococcal infection associated with reduced susceptibility to vancomycin. *MMWR Morb Mortal Wkly Rep* 1997;46:626–628, 635–636.
17. Tenover FC, Lancaster MV, Hill BC, et al. Characterization of staphylococci with reduced susceptibilities to vancomycin and other glycopeptides. *J Clin Microbiol* 1998;36:1020–1027.
18. Ruoff KL, Whiley RA, Beighton D. *Streptococcus*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:283–296.
19. Facklam RR, Sahm DF, Teixeira LM. *Enterococcus*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:297–305.
20. Teixeira LM, Carvallio MGS, Merquior VLC, et al. Recent approaches on the taxonomy of the enterococci and some related microorganisms. *Adv Exp Med Biol* 1997;418:397–400.
21. Ruoff KL. *Leuconostoc, Pediococcus, Stomatococcus* and miscellaneous gram-positive cocci that grow aerobically. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:306–315.
22. Facklam R, Hollis D, Collins MD. Identification of gram-positive coccal and coccobacillary vancomycin-resistant bacteria. *J Clin Microbiol* 1989;27:724–730.
23. Stevens DL. Streptococcal toxic-shock syndrome: spectrum of diseases, pathogenesis, and new concepts in treatment. *Emerg Infect Dis* 1995;3:69–78.
24. Bert F, Lambert-Zechovsley N. Septicemia caused by *Streptococcus carnis* in a human. *J Clin Microbiol* 1997;35:777–779.
25. Bochud PY, Calandra T, Francioli P. Bacteremia due to viridans streptococci in neutropenic patients: a review. *Am J Med* 1994;97:256–264.
26. Endty HP, vanden Braak N, van Belkum A, et al. Fecal carriage of vancomycin-resistant enterococci in hospitalized patients and those living in the community in The Netherlands. *J Clin Microbiol* 1997;35:3026–3031.
27. Kellogg J. Suitability of throat culture procedures for detection of group A streptococci and as reference standards for evaluation of streptococcal antigen detection kits. *J Clin Microbiol* 1990;28:165–169.
28. Centers for Disease Control and Prevention. Prevention of perinatal group B streptococcal disease: a public health perspective. *MMWR Morb Mortal Wkly Rep* 1996;45:1–24.
29. Kircher SM, Meyer MP, Jordan JA. Comparison of a modified DNA hybridization assay with standard culture enrichment for detecting group B streptococci in obstetric patients. *J Clin Microbiol* 1996; 34:343–344.
30. Bourbeau PP, Heiter BJ, Figdore M. Use of Gen-Probe AccuProbe group B *Streptococcus* test to detect group B streptococci in broth cultures of vaginal-anorectal specimens from pregnant women: comparison with traditional culture method. *J Clin Microbiol* 1997;35:144–147.
31. Chew TA, Smith JMB. Detection of diacetyl (caramel odor) in presumptive identification of the "*Streptococcus milleri*" group. *J Clin Microbiol* 1992;30:3028–3029.
32. Murray PR. Modification of the bile solubility test for rapid identification of Streptococcus pneumoniae. *J Clin Microbiol* 1979;9:290–291.
33. Ruoff KL. *Streptococcus anginosus* ("*Streptococcus milleri*"). *Clin Microbiol Newslett* 1988;10:65–68.
34. Knapp JS. Historical perspectives and identification of *Neisseria* and related species. *Clin Microbiol Rev* 1988;1:415–431.
35. Krieg NR, Holt JR, eds. *Bergey's manual of systemic bacteriology*. Baltimore: Williams & Wilkins, 1984:1.
36. Centers for Disease Control and Prevention. *Sexually transmitted disease surveillance 1996*. Atlanta: Centers for Disease Control and Prevention, 1997.
37. Vaneechoutte M, Verschraegen G, Clacys G, et al. Respiratory tract carrier rates of *Moraxella (Branhamella) catarrhalis* in adults and children and interpretation of the isolation of *M. catarrhalis* from sputum. *J Clin Microbiol* 1990;28:2647–2680.
38. Handsfield HH. *Neisseria gonorrhoeae*. In: Mandell GL, Douglas RG Jr, Bennett JE, eds. *Principles and practice of infectious diseases*, 3rd ed. New York: Churchill-Livingstone, 1990:1613–1631.
39. Schachter J, McCormick WM, Smith RF, et al. Enzyme immunoassay for diagnosis of gonorrhea. *J Clin Microbiol* 1984;19:57–59.
40. Roongpisuthipong A, Lewis JS, Kraus SJ, et al. Gonococcal urethritis diagnosed from enzyme immunoassay of urine sediment. *Sex Transm Dis* 1988;15:192–195.
41. Hale YM, Melton ME, Lewis JS, et al. Evaluation of the PACE 2 *Neisseria gonorrhoeae* assay by three public health laboratories. *J Clin Microbiol* 1993;31:451–453.
42. Vlaspolder F, Mutsaers JAEM, Blog F, et al. Value of a DNA probe assay (Gen Probe) compared with that of culture for diagnosis of a gonococcal infection. *J Clin Microbiol* 1993;31:107–110.

43. Lewis JS, Fakile O, Foss E, et al. Direct DNA probe assay for *Neisseria gonorrhoeae* in pharyngeal and rectal specimens. *J Clin Microbiol* 1993;31:2783–2785.
44. Kehl SC, Georgakas K, Swain GR, et al. Evaluation of the Abbott LCx assay for detection of *Neisseria gonorrhoeae* in endocervical swab specimens from females. *J Clin Microbiol* 1998;36:3549–3551.
45. Smith KR, Ching S, Lee H, et al. Evaluation of ligase chain reaction for use with urine for identification of *Neisseria gonorrhoeae* in females attending a sexually transmitted disease clinic. *J Clin Microbiol* 1995;33:455–457.
46. Stary A, Ching SF, Teodorowicz L, et al. Comparison of ligase chain reaction and culture for detection of *Neisseria gonorrhoeae* in genital and extragenital specimens. *J Clin Microbiol* 1997;35:239–242.
47. Centers for Disease Control. *Sexually transmitted diseases clinical practice guidelines*. Atlanta: Centers for Disease Control, 1991.
48. Knapp JS. Antimicrobial resistance in *Neisseria gonorrhoeae* in the United States. *Clin Microbiol Newslett* 1999;21:1–7.
49. Centers for Disease Control and Prevention. 1998 Sexually transmitted diseases treatment guidelines. *MMWR Morb Mortal Wkly Rep* 1998;47:59–63.
50. Doern GV. *Branhamella catarrhalis* and *Neisseria* species. In: Balows A, Hausler JW Jr, Ohashi M, et al., eds. *Laboratory diagnosis of infectious diseases: principles and practice*. New York: Springer-Verlag, 1988;1:134–142.
51. Brown JM, McNeil MM, Desmond EP. *Nocardia, Rhodococcus, Gordona, Actinomadura, Streptomyces* and other actinomycetes of medical importance. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:370–398.
52. Koneman EW, Allen SD, Janda WM, et al., eds. *Color atlas and textbook of diagnostic microbiology*, 5th ed. Philadelphia: Lippincott–Raven, 1997:651–708.
53. Logan NA, Turnbull PCB. *Bacillus* and recently derived genera. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*. 7th ed. Washington, DC: American Society for Microbiology, 1999:357–369.
54. Bille J, Rocourt J, Swaminathan B. *Listeria, Erysipelothrix* and *Kurthia*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:346–356.
55. Funke G, Bernard KA. Coryneform gram-positive rods. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:319–345.
56. Funke G, von Graevenitz A, Clarridge III JE, et al. Clinical microbiology of coryneform bacteria. *Clin Microbiol Rev* 1997;10:125–159.
57. Bille J, Doyle MP. *Listeria* and *Erysipelothrix*. In: Balows A, Hausler WH Jr, Herrmann KL, et al., eds. *Manual of clinical microbiology*, 5th ed. Washington, DC: American Society for Microbiology, 1991:287–295.
58. Coyle MB, Lipsky BA. Coryneform bacteria in infectious diseases: clinical and laboratory aspects. *Clin Microbiol Rev* 1990;3:227–246.
59. Berche P, Reich KA, Bonnichon M, et al. Detection of antilisteriolysin O for serodiagnosis of human listeriosis. *Lancet* 1990;ii:624–627.
60. Gholizadeh Y, Poyart C, Juvin M, et al. Serodiagnosis of listeriosis based upon detection of antibodies against recombinant truncated forms of listeriolysim O. *J Clin Microbiol* 1996;34:1391–1395.
61. Bopp CA, Brenner FW, Wells JG, et al. *Escherichia, Shigella*, and *Salmonella*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:459–474.
62. Dusch H, Altwegg M. Evaluation of five new plating media for isolation of *Salmonella* species. *J Clin Microbiol* 1995;33:802–804.
63. Zadik PM, Chapman PA, Siddons CA. Use of tellurite for the selection of verocytotoxigenic *Escherichia coli* O157. *J Med Microbiol* 1993;39:155–158.
64. Scotland SM, Cheasty T, Thomas A, et al. Beta-glucuronidase activity of Vero cytotoxin producing strains of *Escherichia coli*, including serogroup 0157, isolated in the United Kingdom. *Lett Appl Microbiol* 1991;13:42–44.
65. Park CH, Hixon DL, Morrison WL, et al. Rapid diagnosis of enterohermohagic *Escherichia coli* O157:h7 directly from fecal specimens using immunofluorescence stain. *Am J Clin Pathol* 1994;101:91–94.
66. Dylla BL, Vetter EA, Hughes JG, et al. Evaluation of an immunoassay for direct detection of *Escherichia coli* O157 in stool specimens. *J Clin Microbiol* 1995;33:222–224.
67. Martinez-Murcia AJ, Benlloch S, Collins MD. Phylogenic interrelationships of member of the genera *Aeromonas* and *Plesiomonas* as determined by 165 ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. *Int J Syst Bacteriol* 1992;42:412–421.
68. Janda JM, Duffey PS. Mesophilic aeromonads in human disease: current taxonomy, laboratory identification, and infectious disease spectrum. *Rev Infect Dis* 1988;10:980–997.
69. Altwegg M, Geiss HK. *Aeromona* as a human pathogen. *Crit Rev Microbiol* 1989;16:253–286.
70. Brenden RA, Miller MA, Janda JM. Clinical disease spectrum and pathogenic factors associated with *Plesiomonas shigelloides* infections in humans. *Rev Infect Dis* 1988;10:303–316.
71. Tison DL. *Vibrio*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:497–506.
72. Abbott SL, Seli LS, Catino M Jr, et al. Misidentification of unusual *Aeromonas* species as members of the genus *Vibrio*: a continuing problem. *J Clin Microbiol* 1998;36:1103–1104.
73. Vandamme P, Falsen E, Rossau R, et al. Revision of *Campylobacter, Helicobacter*, and *Wolinella* taxonomy: emendation of genetic descriptions and proposal of *Arcobacter* gen. nov. *Int J Syst Bacteriol* 1991;41:81–103.
74. Nachamkin I. *Campylobacter* and *Arcobacter*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:716–726.
75. Hazell SL. Cultural techniques for the growth of and isolation of *Helicobacter pylori*. In: Goodwin CS, Worsley BW, eds. *Helicobacter pylori: biology and clinical practice*. Boca Raton, FL: CRC Press, 1993:279–283.
76. Isenberg H, ed. *Clinical microbiology procedures handbook*. Washington, DC: American Society for Microbiology, 1992.
77. Meijer BC, Thijs JC, Kleibeuker JH, et al. Evaluation of eight enzyme immunoassays for detection of immunoglobulin G against *Helicobacter pylori*. *J Clin Microbiol* 1997;35:292–294.
78. Wilcox MH, Dent THS, Hunter JO, et al. Accuracy of serology for the diagnosis of *Helicobacter pylori* infection—a comparison of eight kits. *J Clin Pathol* 1996;49:373–376.
79. Winn Jr WC. *Legionella*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:572–585.
80. Buesching WJ, Brust RA, Ayers LW. Enhanced primary isolation of *Legionella pneumophila* from clinical specimens by low-pH treatment. *J Clin Microbiol* 1983;17:1153–1155.
81. Winn WC Jr. Legionnaire's disease: historical perspective. *Clin Microbiol Rev* 1988;1:60–81.
82. Campos J. *Haemophilus*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:604–613.
83. Kersters K, Ludwig W, Vancanneyt M, et al. Recent changes in the classification of the pseudomonads: an overview. *Syst Appl Microbiol* 1996;19:465–477.
84. Bruckner DA, Colonna P. Nomenclature for aerobic and facultative bacteria. *Clin Infect Dis* 1997;25:1–10.
85. Gilligan PH, Whittier S. *Burkholderia, Stenotrophomonas, Ralstonia, Brevundimonas, Comamonas* and *Acidovorax*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:526–538.
86. Brenner DJ, Hollis DG, Fanning GR, et al. *C. canimorsus* sp. nov. (formerly *C. cynodegmi* sp. nov.), a cause of localized wound infection following dog bite. *J Clin Microbiol* 1989;27:231–235.
87. Wormser GP, Bottone EJ. *Cardiobacterium hominis*: review of microbiologic and clinical features. *Rev Infect Dis* 1983;5:680–691.

88. Jenn DB, Letendre PW, Iverson G. Endocarditis caused by *Kingella indologenes*. *Rev Infect Dis* 1987;9:787–789.
89. Catlin BW. *Gardnerella vaginalis*: characteristics, clinical considerations, and controversies. *Clin Microbiol Rev* 1992;5:213–237.
90. Mazzulli T, Simor AE, Low DE. Reproducibility of interpretation of gram-stained vaginal smears for the diagnosis of bacterial vaginosis. *J Clin Microbiol* 1990;28:1506–1508.
91. Welch DF, Slater LN. Bartonella and *Afipia*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:638–646.
92. Brenner DJ, Hollis DG, Moss CW, et al. Proposal of *Afipia* gen. nov., with *Afipia felis* sp. nov. (formerly the cat scratch disease bacillus), *Afipia clevelandensis* sp. nov. (formerly the Cleveland Clinic Foundation strain), *Afipia broomae* sp. nov., and three unnamed genospecies. *J Clin Microbiol* 1991;29:2450–2460.
93. Jameson P, Greene C, Regnery R, et al. Prevalence of *Bartonnella henselae* antibodies in pet cats throughout regions of North America. *J Infect Dis* 1995;172:1145–1149.
94. Provenza JM, Klotz SA, Penn RL. Isolation of *Francisella tularensis* from blood. *J Clin Microbiol* 1986;24:453–455.
95. Raad I, Rand K, Gaskins D. Buffered charcoal-yeast extract medium for the isolation of brucellae. *J Clin Microbiol* 1990;28:1671–1672.
96. Gilchrist MJR. *Bordetella*. In: Balows A, Hansler Jr WJ, Herrmann KL, et al., eds. *Manual of clinical microbiology*, 5th ed. Washington, DC: American Society for Microbiology, 1991:471–477.
97. McNicol P, Giercke SM, Gray M, et al. Evaluation and validation of a monoclonal immunofluorescent reagent for the direct detection of *Bordetella pertussis*. *J Clin Microbiol* 1995;33:2868–2871.
98. Collins MD, Lawson PA, Willems A, et al. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combination. *Int J Syst Bacteriol* 1994;44:812–826.
99. Finegold SM, George WL, eds. *Anaerobic infections in humans*. New York: Academic Press, 1989.
100. Appelbaum PC. Advances in rapid preliminary and definitive identification of anaerobes. *Clin Microbiol Newslett* 1989;11:89–93.
101. Baron EJ, Citron DM. Anaerobic identification flowchart using minimal laboratory resources. *Clin Infect Dis* 1997;25(Suppl 2):S143–S146.
102. Summanen P, Barron EJ, Citron DM, et al. *Wadsworth anaerobic bacteriology manual*, 5th ed. Belmont, CA: Star Publishing, 1992.
103. Dowell VR Jr, Lombard GL. Differential agar media for identification of anaerobic bacteria. In: Tilton RC, ed. *Rapid methods and automation in microbiology*. Washington, DC: American Society for Microbiology, 1985:258–262.
104. Groschel DH. *Clostridium difficile* infection. *Crit Rev Clin Lab Sci* 1996;33:203–245.
105. Allen SD, Emery CL, Siders JA. *Clostridium*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:654–671
106. Baranton G, Postic D, Saint-Girons I, et al. Delineation of *Borrelia burgdorferi sensu stricto, Borrelia garvinii* sp. nov., and group VS 461 associated with Lyme borreliosis. *Int J Syst Bacteriol* 1992;42:378–383.
107. Kmety E, Dikken H. *Classification of the species* Leptospira interrogans *and history of its serovars*. Groningen, The Netherlands. University Press, 1993.
108. Weyant RS, Bragg SL, Kaufmann AF. *Leptospira* and *Leptonema*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:739–745.
109. Larsen SA, Norris SJ, Pope V. *Treponema* and other host-associated spirochetes. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:759–776.
110. Fiumara NJ. Serologic responses to treatment of 128 patients with late latent syphilis. *Sex Transm Dis* 1979;6:243–246.
111. Johnson BJB, Robbins KE, Bailey RE, et al. Serodiagnosis of Lyme disease: accuracy of a two-step approach using a flagella-based ELISA and immunoblotting. *J Infect Dis* 1996;174:346–353.
112. Kiska DL, Gilligan PH. *Pseudomonas*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:518–519.

SUGGESTED READINGS

Anderson BE, Neuman MA. *Bartonella* spp. as emerging human pathogens. *Clin Microbiol Rev* 1997;10:203–219.

Bascomb S, Manafi M. Use of enzyme tests in characterization and identification of aerobic and facultatively anaerobic gram-positive cocci. *Clin Microbiol Rev* 1998;11:318–340.

Dunn BE, Cohen H, Blaser MJ. *Helicobacter pylori*. *Clin Microbiol Rev* 1997;10:720–741.

Forbes BA, Sahm DF, Weissfeld AS, eds. *Bailey and Scott's diagnostic microbiology*, 10th ed. St. Louis: Mosby, 1998.

Howard BJ, Klass J II, Rubin SJ, et al., eds. *Clinical and pathogenic microbiology*. St. Louis: Mosby, 1987.

Isenberg HD, ed. *Essential procedures for clinical microbiology*. Washington, DC: American Society for Microbiology, 1998.

Kiska DL. Staphylococcal and streptococcal toxic shock syndrome. *Clin Microbiol Newslett* 1997;19:33–37.

Koneman EW, Allen SD, Janda WM, et al., eds. *Color atlas and textbook of diagnostic microbiology*, 5th ed. Philadelphia: Lippincott–Raven, 1997.

Larsen SA, Steiner BM, Rudolph AH. Laboratory diagnosis and interpretation of tests for syphilis. *Clin Microbiol Rev* 1995;8:1–21.

Mandell GL, Douglas RG Jr, Bennett JE, eds. *Principles and practice of infectious diseases*, 3rd ed. New York: Churchill-Livingstone, 1990.

Marcon MJ. Clinical and laboratory diagnostic features of *Bordetella* spp.—pertussis and beyond. *Clin Microbiol Newslett* 1997;19:185–191.

Miller M. *A guide to specimen management in clinical microbiology*, 2nd ed. Washington, DC: American Society for Microbiology, 1998.

Murray PR, Baron, EI, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999.

Spiegel CA. Bacterial vaginosis: changes in laboratory practice. *Clin Microbiol Newslett* 1999;21:33–37.

52
Fungi and Fungal Infections

Imaad Mujeeb

Deanna A. Sutton

Annette W. Fothergill

Michael G. Rinaldi

Michael A. Pfaller

Not long ago, fungal infections of humans were considered nuisance diseases—often irritating and cosmetically unappealing and affecting mainly the skin, mucous membranes, hair, and nails. Rarely, a patient might live in or pass through an area endemic for one of the classic dimorphic systemic pathogens and present with coccidioidomycosis, histoplasmosis, or blastomycosis. Even more rarely, an infection would occur after surgery or chemotherapy for cancer. For the most part, fungal infections were not regarded as significant entities for the clinician or laboratorian.

In contemporary medicine, the situation has changed significantly. Medical mycology has come into its own. With the ongoing discovery and more frequent use of the cutting-edge techniques of medical technology, i.e., more sophisticated surgical procedures used more aggressively, the advent of organ transplantation, the wider use of cytotoxic chemotherapy for malignancy, the application of broad-spectrum antibiotics, vaccines, indwelling catheters, lines, and devices, and, perhaps of most importance, the ever-escalating incidence of human immunodeficiency virus (HIV) infection and resultant acquired immunodeficiency syndrome (AIDS), fungi have become major agents of infectious diseases, particularly in hosts with abrogated immunity (1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29 and 30).

Persons whose immune systems are dysfunctional or nonfunctional are, in a genuine sense, “living Petri dishes.” Fungi that are ordinarily nonpathogenic, harmless saprobes, most often living in nature in soil or on plants, are able to take advantage of the situation in the immunocompromised host. Such organisms gain an opportunity to be pathogens, hence, the term opportunistic infections. The fungi have become some of the major agents of opportunistic infections of humans. The relative dearth of nontoxic therapies, the absence of diagnostic tests that permit early and correct diagnosis, and the general lack of training in mycology and mycotic diseases among health care professionals have resulted in a steady and consistent increase in both the number and significance of fungal infections over the past decade. Presently, mycotic infections pose one of the major threats to the health and life of immunocompromised patients (3,4,6,8,10,11,15,18,19,20,21 and 22,24,25,27,29,30).

As in most infectious diseases, the laboratorian is called on to assist the clinician in diagnosing and treating fungal infections. In the case of infections caused by fungi, the role of a good clinical mycology laboratory is paramount. Because many mycotic infections are caused by fungi that are not well known in human medicine, it becomes essential to have access to laboratory personnel with training and skills in classic botanical mycology. Likewise, with the advent of great numbers of new and efficacious antifungal drugs, the laboratory is increasingly called on to provide *in vitro* susceptibility data to assist the clinician in choosing optimal therapeutic regimens. It is unlikely that fungal infections will decrease in number or significance; in fact, it is quite the contrary. The role of the laboratory has never been greater or more important in the field of medical mycology than it is now.

FUNGI PATHOGENIC FOR HUMANS

Part of "52 - Fungi and Fungal Infections"

Fungi are eukaryotic organisms, both macroscopic and microscopic, that make up one of the major kingdoms of organisms. Contrary to definitions provided in many sources, fungi are not lower plants and are not properly classified in the plant world. Rather, fungi are classified in their own distinct taxon, the kingdom Fungi (or kingdom Mycota). Fungi are unique and merit being placed into their own kingdom for several reasons. Fungi, as noted, are true eukaryotes, having their DNA surrounded by a membrane, possess typical membrane structure and function, and have the typical array of eukaryotic organelles (e.g., ribosomes, mitochondria).

However, fungi have cell walls containing complex polysaccharides, the main ones being glucans and mannans. They also demonstrate *N*-acetylglucosamine or chitin in their walls. Fungi lack chlorophyll and are heterotrophic. Hence, they produce a plethora of enzymes that enable them to degrade almost any substrate known in nature. In fact, the major function of fungi in life is to degrade organic carbon. It is the fungi that degrade or decompose the bulk of organic carbon on the planet and return it to the atmosphere in the form of carbon dioxide.

Fungi reproduce both sexually and asexually and exist in literally

thousands of different morphologic expressions and multitudinous life cycles (31,32). Most relevant to medical mycology is the ability of fungi to produce asexual propagules of many types, shapes, and sizes. Such structures are often the means by which humans become infected through inhalation, contact with skin and mucous membranes, or after trauma. The two main growth forms of fungi consist of single-celled, budding forms called yeasts and multicellular, filamentous forms called molds. Among medical fungi, several possess the ability to convert from one morphology to another and back again, depending on the conditions of growth and the genetics of the specific organism. Such fungi are called dimorphic because of their ability to exist in two distinct morphologic forms. Other fungi, including many species of medical interest, demonstrate polymorphism: existence in many distinct morphologies. Once again, depending on growth conditions, such polymorphic fungi may transform from one form to another and back again. Given their genetic makeup, their ability to elaborate enzymes capable of powerful degradation, and their production of huge numbers of reproductive propagules, the fungi are clearly among the most versatile and successful of living creatures. The colonization, invasion, and use of human tissue as a growth medium are easy tasks for the fungi.

Few areas in biology have elicited more controversy and emotion than the taxonomy, nomenclature, and classification of the fungi. Table 52.1 indicates one current system of fungal classification, and Table 52.2 shows contemporary thinking as to the nomenclature of the actual fungal infections (mycoses). Most recently, the techniques of molecular biology have been employed to ascertain the relationships between fungi. Results of such early investigations show that a combination of classic morphologic study and molecular genetics will likely yield the most accurate data.

TABLE 52.1. CLASSIFICATION SCHEME FOR THE KINGDOM FUNGI

Group	Group Ending
Kingdom	
Division	-mycota
Subdivision	-mycotina
Class	-mycetes
Subclass	-mycetidae
Order	-ales
Family	-aceae
Genus	-no specific ending
Species	-no specific ending
Variety	-no specific ending
Example of the use of the taxonomic system presented in classifying a clinically significant, common dermatophyte: The organism is *Trichophyton mentagrophytes* var. *interdigitale*	
SUPERKINGDOM	Eukaryota
KINGDOM	Fungi
FORM-DIVISION	Fungi Imperfecti
FORM-CLASS	Hyphomycetes
FORM-ORDER	Hyphomycetales
FORM-FAMILY	Moniliaceae
FORM-GENUS	*Trichophyton*
FORM-SPECIES	*T. mentagrophytes*
FORM-VARIETY	*T. mentagrophytes* var. *interdigitale*

MYCOSES—THE DISEASES CAUSED BY FUNGI

Part of "52 - Fungi and Fungal Infections"

The gamut of fungal infections in humans is broad and ranges from the most superficial to the ultimate in tissue invasion and destruction. In discussing fungal infections, most authorities would employ the following scheme of mycotic infection classification or a variant thereof.

Superficial mycoses: These mycoses infect the superficial layers of the skin; most often the patient demonstrates little or no pathology except the cosmetic aspects. A very common example of superficial mycosis is pityriasis versicolor (tinea versicolor), caused by the yeastlike fungus *Malassezia furfur*.

Cutaneous mycoses: Although many mycoses may manifest as skin lesions, often secondary to disseminated disease or a primary cutaneous infection by a fungus that may cause deep disease, the cutaneous mycoses as discussed here really refer to skin infections caused by a specific group of fungi called dermatophytes. Dermatophytosis is among the most common of fungal infections throughout the world and affects the hair, skin, and nails. Dermatophytoses are caused by three genera of dermatophytes, *Trichophyton*, *Microsporum*, and *Epidermophyton*. Some dermatophyte-caused mycoses are not only among the most common of fungal infections but also among the most common of human diseases, e.g., tinea pedis or athlete's foot.

Subcutaneous mycoses: These mycoses affect not only the skin but also the subcutaneous tissues and even bone. Most often, these infections occur after an accidental and/or traumatic implantation of the fungus into the subcutaneous tissue. The fungi inciting these infections are dwellers in nature and do not require a human host. However, after inoculation, these organisms can precipitate a relentless, chronic disease, with the potential for great destruction of tissue. The span of subcutaneous mycoses is long, with many different fungi responsible for various clinical manifestations. Classic subcutaneous mycoses are sporotrichosis (*Sporothrix schenckii*), chromoblastomycosis (caused by at least six different species of darkly pigmented molds), and eumycotic mycetoma (caused by a wide variety of fungi).

Systemic (deep) mycoses: The traditional systemic agents are those fungi capable of inciting disease in the deep viscera after their acquisition. The classic systemic fungal pathogens are dimorphic (occur in both yeast and mold forms): *Coccidioides immitis* (coccidioidomycosis), *Histoplasma capsulatum* (histoplasmosis), *Blastomyces dermatitidis* (blastomycosis), and *Paracoccidioides brasiliensis* (paracoccidioidomycosis). *Penicillium marneffei* is a recently recognized dimorphic fungus endemic to Southeast Asia that causes systemic infections in HIV-infected individuals (9,33,34,35,36,37 and 38). Most deep fungal infections result after dissemination of the fungus from the lungs following its inhalation from nature. These infections occur quite regularly in the normal host, although, as with most infectious diseases, they are much more rampant and severe when present in the immunocompromised individual, e.g., histoplasmosis and coccidioidomycosis in those with AIDS.

Opportunistic mycoses: This group of mycoses is, in contemporary medicine, the most significant by far. Opportunistic mycoses are those occurring in the host with abrogated or altered immunity. It is now safe to say that given the right moment, in the correct “living Petri dish,” virtually any fungus may incite infectious

disease. There are virtually no fungal organisms that cannot cause serious disease in this patient population. These mycoses, more than any other factor, have changed the face of medical mycology and demonstrated the true power of those microbes to render horrific infections. In particular, patients who have neutropenia or leukemia, bone marrow transplant recipients, solid organ transplant patients, burn victims, those with genetic defects in their immune systems, among others, are all susceptible to infection by an ever-growing number of taxonomically diverse fungi. Of course, persons with AIDS have been victims of the opportunistic fungi in escalating numbers. Recognizing that any fungus may incite disease in this patient group, the classic opportunistic mycoses still remain: candidiasis (largely owing to *Candida albicans*, but with other species being more and more frequently implicated), cryptococcosis (*Cryptococcus neoformans*, particularly in the AIDS population), aspergillosis (*Aspergillus fumigatus* and *A. flavus* are the major agents, particularly in neutropenic patients), zygomycosis (mainly owing to *Rhizopus arrhizus* in the ketoacidotic diabetic), phaeohyphomycosis [caused by darkly pigmented (dematiaceous) molds and yeasts] and hyalohyphomycosis (caused by lightly pigmented, colorless, or hyaline molds). In large part, modern medical mycology has become the science of infections caused by opportunistic fungi (3,4,6,8,10,11,15,18,25,27,29,30).

TABLE 52.2. NOMENCLATURE OF MYCOTIC INFECTIONS

Nomenclature	Definition
Adiaspiromycosis	Mycosis owing to *Chrysosporium parvum*
African histoplasmosis (histoplasmosis duboisii)	Mycosis owing to *Histoplasma capuslatum* var. *duboisii*
Aspergillosis	Mycosis owing to *Aspergillus* spp
Black peidra	Mycosis owing to *Piedraia hortae*
Blastomycosis	Mycosis caused by *Blastomyces dermatitidis*
Candidiasis/candidosis	Yeast infection caused by *Candida* spp
Chromoblastomycosis	Sclerotic cells (bodies) in the dermis and/or in scales
Coccidioidomycosis, valley fever	Mycosis owing to *Coccidioides immitis*
Cryptococcosis	Yeast infection caused by *Cryptococcus* spp
Dermatomycosis	Mycosis of the skin and/or appendages
Dermatophytic mycetoma, pseudomycetoma, mycetomalike dermatophytosis (mycetomatoid dermatophytosis)	Mycosis caused by a dermatophytic fungus which clinically and histopathologically resembles eumycotic mycetoma
Dermatophytid	Skin allergic manifestation of a fungal infection
Dermatophytosis, tinea, ringworm	Mycosis caused by a dermatophyte
Epizootic lymphangitis (histoplasmosis farciminosi)	Mycosis owing to *Histoplasma capuslatum* var. *farcininosum*
Eumycotic mycetoma	Mycosis characterized by fistulae, granules in tissues, tumefaction, and often lesions of bone
Favus	Dermatophytosis of the skin/hair of the scalp with presence of crusty, purulent scutulae
Fungemia	Presence of fungi in the blood
Geotrichosis	Mycosis caused by *Geotrichum* spp
Histoplasmosis (histoplasmosis capuslati)	Mycosis owing to *Histoplasma capsulatum* var. *capsulatum*
Hyalohphomycosis	Opportunistic mycosis caused by various moniliaceous (light-colored, nondematiaceous) molds whose basic tissue form is in the nature of hyaline hyphal elements that are septate, branched or unbranched, and occasionally toruloid
Lobomycosis	Mycosis owing to *Loboa loboi*
Mycosis (plural, mycoses)	Parasitic disease caused by a fungus
Mycotic keratitis	Mycosis of the cornea
Onychomycosis	Mycosis of the nails
Opportunistic mycosis	Mycosis occurring in a host whose immune system has been/is compromised, either through innate disease or as result of medical intervention; often caused by fungi normally considered as harmless, contaminants, saprobes, or plant pathogens
Otomycosis	Mycosis of the external auditory canal
Paracoccidioidomycosis	Mycosis owing to *Paracoccidioides brasiliensis*
Phaeohyphomycosis	Mycosis caused by various dematiaceous fungi in which dematiaceous (dark) yeastlike cells, pseudohyphaelike elements, hyphae that may be short or elongate, regular, distorted to swollen shape, or any combination of these forms may be found in tissue
Pityriasis versicolor	Dermatomycosis caused by *Malassezia furfur*
Rhinosporidiosis	Mycosis owing to *Rhinosporidium seeberi*
Sporotrichosis	Mycosis owing to *Sporothrix schenkii*
Systemic mycosis	Mycotic infection of the viscera
Thrush	*Candida albicans* infection of the mucous membranes of the oral cavity
Tinea nigra	Phaeohyphomycosis (dermatomycosis) caused by *Phaeoannellomyces werneckii*
Torulopsosis	Yeast infection caused by *Torulopsis* spp
Trichosporosis	Yeast infection caused by *Trichosporon* spp
White piedra	Trichosporosis of the hair
Yeast infection	Mycosis caused by a yeast
Zygomycosis	Mycosis caused by members of the Zygomycetes

EPIDEMIOLOGY OF MYCOSES

Fungal infections are acquired in a variety of ways but mainly via acquisition of the etiologic fungus from its site in nature to the human host by accident or happenstance. In most cases, the fungus is living in nature, humans intervene through work or play, and the fungal inoculation is acquired. This can be by inhalation (very common), nontraumatic or traumatic inoculation (common), or ingestion (rare). Many of the mycoses are obtained by inhalation of the fungus from a geographic area of the world where the fungus is either wholly endemic or found much more frequently than elsewhere. For example, *H. capsulatum* var *capsulatum* and *B. dermatitidis* are global pathogens, but they are particularly prevalent in the Ohio-Mississippi River valley of the United States. Most cases of histoplasmosis and blastomycosis are diagnosed in individuals who either live or travel in, to, or from these areas. Similarly, *C. immitis* is endemic only to the lower Sonoran life zone regions of the earth (southwestern United States, parts of Mexico, and Central and South America). Coccidioidomycosis occurs after inhalation of the fungus while in one of the endemic regions. With the above-mentioned mycoses, there is no transmission from person to person; these are not communicable, contagious diseases. Another group of human fungal pathogens, the dermatophytic fungi, however, can be transmitted from person to person. With their skin fungal infections, it is possible to implement transmission from one human to another via indirect or direct contact (e.g., acquiring athlete's foot after using a shower stall in bare feet after its previous use by an individual with tinea pedis or the spread of ringworm of the scalp through a grade-school class when children have used a comb belonging to a classmate with tinea capitis). Likewise, humans are frequently infected by dermatophytes after handling their pet cat or dog or domestic cattle that are infected with various animal dermatophytic pathogens.

As has been mentioned, humans may acquire mycoses after traumatic implantation of the fungal agent from a source in nature. The classic example is rose-handler's or gardener's disease, sporotrichosis, acquired after the handling of rose thorns or sphagnum moss. Many contemporary cases of phaeohyphomycosis occur after traumatic implantation of soil or other materials from nature.

Of growing concern, however, are mycoses obtained in the hospital, often from the same fungi that comprise a part of the normal human flora. Although the array of fungal pathogens known to cause nosocomial infection is diverse, the vast majority of these infections are owing to *Candida* species. Recent studies confirmed the predominant role of candidiasis among the invasive mycoses (1,11,15,16,17,18,19,20,21,22,23,24 and 25,39). Among the species of *Candida* causing bloodstream infection in the United States, *C. albicans* accounts for 50% to 70% of all infections (15,24). As a group, *Candida* species are a part of the microbiota of the oral and gastrointestinal tracts of humans, and, in women, it is often found as normal vaginal flora. After surgery, antibiotic therapy, chemotherapy, and placement of indwelling devices, *C. albicans*, and other *Candida* species have become major bloodstream pathogens. The most recent epidemiologic figures show *Candida* to be the fourth most prevalent cause of hospital-acquired bloodstream infection (10,19,20,21,22 and 23). In surgical, medical, and neonatal intensive care units, *Candida* is a matter for critical concern in modern medicine (11,22). In fact, without question, candidiasis is the most prevalent and significant mycosis of humans (3,4,6,7,15,16,17,18,19,20,21,22,23,24 and 25).

Finally, in the setting of the neutropenic patient, fungi from among the environmental flora, found everywhere at all times, have become major killers. For example, in bone marrow and liver transplant services, the aspergilli are of concern (11,13,25,27,40,41,42,43 and 44). The mortality rate from invasive aspergillosis is unacceptably high, approximately 80% to 100% in most published studies (11,25,40). Often, outbreaks follow periods of hospital construction or renovation when the asexual spores (conidia) of aspergilli are prevalent in the air. Another opportunistic fungus that is widely dispersed in nature is *Fusarium*. Fusaria are among the most frequent of plant pathogenic fungi but now cause a higher rate of mortality in granulocytopenic patients than aspergilli (12,14,41,44,45 and 46). Such nosocomial opportunistic pathogenic fungi are on the rise, and constant vigilance must be exercised if mortality is to be minimized in the susceptible patient populations.

DIAGNOSTIC PROCEDURES (BY DISEASE TYPE)

Part of "52 - Fungi and Fungal Infections"

Superficial Mycoses

Agents of superficial mycosis are organisms that attack the keratinized outer layers of the skin, hair, and nails (47). Infections of this sort are usually of cosmetic concern only and cause little if any discomfort to the patient.

Pityriasis Versicolor

M. furfur is responsible for the condition known as pityriasis versicolor and can be found worldwide. Pityriasis versicolor presents as a series of darkened patches in very light-skinned individuals or as nonpigmented patches on individuals with darker skin. These areas will fluoresce yellow under a Wood's lamp. KOH (potassium hydroxide) examination reveals a characteristic “spaghetti and meatballs” appearance owing to the fragments of hyphae mixed with round yeast cells. *M. furfur* has also been implicated in fungemia in neonates and some adult patients receiving intralipid intravenously (16,31,48,49 and 50). Fungemic episodes usually resolve once the indwelling catheter is removed. *M. furfur* can easily be seen on Gram stain of positive blood culture broths but will not grow on Sabouraud agar when subcultured. The simple overlaying of the culture plate with olive oil will result in abundant growth within 48 hours. Any time a yeast is suspected but is not recovered, this simple method should be employed. At present, there are no direct antigen or DNA methods nor serologic procedures useful for detection of this yeast. The characteristic appearance of yeast cells resembling small bottles (this fungus was originally termed the “bottle bacillus”) seen in blood from patients receiving parenteral nutritional support with products containing oleic acid, coupled with lack of growth on routine laboratory media, should present a high index of suspicion to both the clinician and laboratorian for *M. furfur* fungemia.

Tinea Nigra

Phaeoannellomyces werneckii, previously called *Cladosporium werneckii*, is responsible for the superficial phaeohyphomycotic condition of tinea nigra. Tinea nigra appears as a dark brown to black patch, most often on the palms of the hands and soles of the feet. Although tinea nigra is also mainly of cosmetic concern, it must be differentiated from malignant melanoma before any treatment. Often extensive surgery can be avoided with a KOH examination of the affected area.

When fungal elements are determined to be present, skin scrapings should be placed on mycologic media with antibiotics. A dematiaceous yeastlike colony should appear within 3 weeks and may become velvety with age. The degree of hyphal growth depends heavily on the media and conditions under which the colony grows. Microscopic examination reveals two-celled cylindrical yeastlike cells and, depending on the age of the colony, toruloid hyphae.

P. werneckii is identified microscopically by seeing characteristic darkly pigmented annelloconidia (conidia possessing annellides or rings), which often slide down the sides of the conidiophore (Fig. 52.1) (31,32,51). Because no additional means, e.g., serologic tests, are currently available to identify this fungus or this disease, microscopic morphology, coupled with the characteristic lesions on the patient's skin, is the only satisfactory identification procedure.

FIGURE 52.1. Microscopic appearance of *Phaeoannellomyces werneckii* (Nomarski optics, original magnification ×1,250).

White Piedra

Trichosporon asahi (beigelii) (formerly *Trichosporon cutaneum*) is the causative agent for the condition called white piedra that affects the hair of the scalp, face, and body. This fungus surrounds the hair shaft and forms a white to brown swelling along the hair strand. By running a section of hair between the thumb and forefinger, the nodules can easily be removed. This condition is related to poor hygiene. When infected, the hair should be cut or shaved and kept clean. These steps usually negate the necessity of medical treatment.

When microscopic examination reveals hyphal elements or arthroconidia and/or budding yeast cells, infected hair may be placed on mycologic media without cycloheximide. Cream-colored, dry, wrinkled colonies should form within 48 to 72 hours at room temperature.

T. asahi (beigelii) can be identified in the same manner as other yeast isolates. Sugar assimilations, potassium nitrate (KNO_3) assimilation, urease production, and morphology on corn meal should be determined. *T. asahi (beigelii)* is KNO_3 negative and urease positive and grows at 37°C.

A most important microscopic characteristic is production of both arthroconidia (rectangular cells resulting from the fragmentation of hyphal cells) and blastoconidia (budding yeast cells). In *T. asahi (beigelii)*, one can see the arthroconidia and blastoconidia concurrently (31,32,50).

Black Piedra

Another condition affecting the hair, most often that of the scalp, is black piedra. The causative agent of black piedra is *Piedraia hortae*. Black piedra presents as dark, rocklike hyphal masses that surround and penetrate the hair shaft. Examination of the hyphal masses reveals branched hyphae tightly held together with a cementlike substance. Black piedra is also a condition of poor hygiene and is easily resolved with a haircut and proper regular washings.

P. hortae can be cultured on routine mycologic media. Very slow growth may be observed at 25°C and may begin as a yeastlike colony, later becoming velvety as hyphae develop. Asci may be observed microscopically, usually ranging from 4 to 30 μm and containing as many as eight ascospores. The characteristic cementlike material holding all the above-mentioned structures in place around hair shafts is typical of the group of fungi in which *P. hortae* is classified, namely, the Loculoascomycetes (31,32,51).

Cutaneous Mycoses

There are many species of dermatophytic fungi possessing the ability to cause disease in humans and animals. All have in common the ability to invade the skin, hair, or nails. In each case, these fungi are keratinophilic and keratinolytic, being able to break down the keratin surfaces of these structures. In the case of skin infections, the dermatophytes invade only the upper, outermost layer of the epidermis, the stratum corneum. Likewise, with hair and nails, being part of the skin, only the keratinized layers are invaded. Often the growth of dermatophytes in the skin per se is not the major reason for skin symptoms but rather the host immune response to the presence of the fungal elements. The outcome of acute versus chronic dermatophytoses is mainly the result of the host immune response rather than the invasiveness of the fungi.

The various forms of dermatophytosis are often referred to as tinea, e.g., tinea capitis (scalp infection), tinea corporis (infection of the trunk), tinea cruris (infection of the groin), tinea pedis (infection of the feet). A major clinical manifestation of dermatophytoses is infection of the nails (onychomycosis or tinea unguium). Three genera of dermatophytes are known (31,32,52): *Microsporum*, *Trichophyton*, and *Epidermophyton*. Without question, the most common pathogenic species observed by clinicians is *T. rubrum*. This species incites the bulk of athlete's foot

and nail infection seen by dermatologists and other clinicians. Other *Trichophyton* species frequently causing human diseases are *T. tonsurans* (infection of the hair and scalp) and *T. mentagrophytes*. Of the *Microsporum* species, *M. canis* and *M. gypseum* are most often seen in the United States. Infection owing to *M. canis* most often follows handling of cats and dogs, where the fungus resides naturally. Infections caused by *E. floccosum* most often affect the feet and the groin.

Dermatophytosis can manifest clinically in diverse ways, but most diagnoses are made based on clinical appearance and patient history. Cultures are always desirable and can be obtained from scraping the affected areas and placing the skin, hair, or nail tissue on any number of laboratory media such as Sabouraud agar, with and without antibiotics, or dermatophyte test medium. An alternative is employment of potato flakes agar, containing the color indicator bromthymol blue (53). With this medium, the dermatophytic fungi change the medium from yellow to blue but do not obscure the pigments produced by various dermatophytes that are often useful as adjuncts in identification. Also, this medium promotes sporulation of dermatophytes to enhance microscopic identification. As with dermatophyte test medium, many nondermatophytic fungi also grow on potato flakes agar and may turn the medium blue. The final identification lies with examination of the fungus under the microscope.

Microscopically, the genus *Microsporum* is identified by observation of its macroconidia, whereas with the genus *Trichophyton*, microconidia are the characteristic structures (Table 52.3). *E. floccosum* produces no microconidia, but its smooth-walled macroconidia borne in clusters of two or three are most distinctive. *M. canis* produces characteristic large, thick- and rough-walled macroconidia (Fig. 52.2). Its macroconidia are multicellular (five to eight cells per conidium). *Trichophyton rubrum* produces microconidia that are teardrop- or peg-shaped borne rough along the sides of hyphae, whereas *T. mentagrophytes* produces both single cigar-shaped macroconidia and grapelike clusters of spherical microconidia. *T. tonsurans* produces variably sized and shaped microconidia, with relatively large spherical conidia often being right alongside of small, parallel-walled conidia and other microconidia of various sizes and shapes (Fig. 52.3).

TABLE 52.3. CHARACTERISTIC FEATURES OF MACROCONIDIA AND MICROCONIDIA OF DERMATOPHYTES

Genus	Macroconidia	Microconidia
Epidermophyton	Smooth-walled, borne in cluster of two or three	Absent
Microsporum	Numerous, large, thick- and rough-walled[a]	Rare
Trichophyton	Rare, smooth, thin-walled	Numerous, spherical, teardrop- or peg-shaped[b]

[a] Except *M. audouinnii*.

[b] Except *T. schoenleinii*.

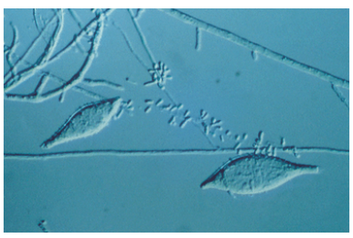


FIGURE 52.2. Macroconidia and microconidia of *Microsporum canis* (Nomarski optics, original magnification ×625).

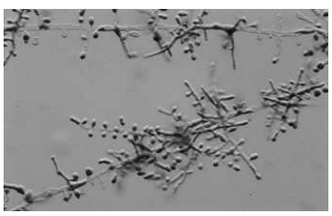


FIGURE 52.3. Microconidia of *Trichophyton tonsurans* (original magnification ×625).

Skin scraping, nail scraping, and observation of hairs directly by microscopy are also most helpful in diagnosing dermatophytic infections. Employing a drop of potassium hydroxide (15% KOH) on a glass slide over the surface of the specimen will often reveal filamentous, clear hyphal elements characteristic of dermatophytes (Fig. 52.4). Calcofluor white has been employed in examining specimens for fungal elements and has demonstrated excellent results.

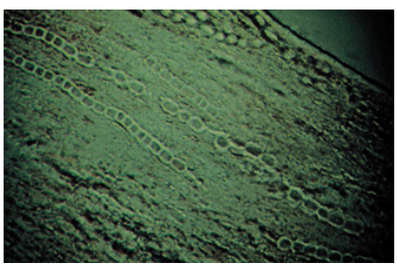


FIGURE 52.4. Potassium hydroxide preparation demonstrates hyphal elements characteristic of dermatophytes. *Trichophyton tonsurans* in hair shaft (original magnification ×600).

Identification of dermatophytes to genus is relatively straightforward, but species-level identification may be considerably more difficult.

Subcutaneous Mycoses

Many fungal pathogens can produce subcutaneous manifestations as part of their disease process; however, some fungi are commonly introduced traumatically through the skin and have a

propensity to involve the deeper layers of the dermis, subcutaneous tissue, and bone. Although they may ultimately be expressed clinically as lesions on the skin surface, they rarely spread to distant organs. In general, the clinical course is chronic and insidious, and, when established, the infections are refractory to most antifungal therapy. The most common subcutaneous mycoses are lymphocutaneous sporotrichosis, chromoblastomycosis, and eumycotic mycetoma (Table 52.4). Although lymphocutaneous sporotrichosis is caused by a single fungal pathogen, *S. schenckii*, the other subcutaneous mycoses are clinical syndromes caused by multiple fungal etiologies (31,54,55 and 56). The causative agents of subcutaneous mycoses are generally considered to have low pathogenic potential and are commonly isolated from soil or decaying vegetation. Exposure is largely occupational or related to hobbies (e.g., gardening, wood gathering). Infected patients generally have no underlying immune defect.

TABLE 52.4. COMMON AGENTS OF SUBCUTANEOUS MYCOSES

		Growth		
Disease	**Organism**	**Tissue**	**Culture (25°C)**	**Culture (37°C)**
Sporotrichosis	*Sporothrix schenckii*	Yeast (rare, cigar shaped)	Mold	Yeast
Chromoblastomycosis	*Cladosporium*	Hyphae[a]	Mold	Mold
	Fonsecaea	Hyphae[a]	Mold	Mold
	Phialophora	Hyphae[a]	Mold	Mold
Mycetoma	*Pseudallescheria*	Hyphae[b]	Mold	Mold
	Madurella	Hyphae[b]	Mold	Mold

[a] Dematiaceous, often blunted to form oval sclerotic bodies.

[b] May be in the form of microcolonies or granules.

Lymphocutaneous Sporotrichosis

Lymphocutaneous sporotrichosis is caused by *S. schenckii*, a dimorphic fungus that is ubiquitous in soil and decaying vegetation (31,32,54). Infection with this organism is chronic and is characterized by nodular and ulcerative lesions that develop along lymphatics that drain the primary site of inoculation. Dissemination to other sites such as bones, eyes, lungs, and central nervous system is extremely rare, occurring in less than 1% of all cases (31,32). Histopathologically, the organism stimulates both acute pyogenic and granulomatous inflammatory reactions. In tissues, the organism appears as a 3- to 5-µm cigar-shaped yeast (Fig. 52.5) but is rarely observed in human lesions. Thus, direct microscopic examination of infected material is usually unrewarding. Definitive diagnosis usually requires culture of infected pus or tissue. *S. schenckii* grows within 2 to 5 days on a variety of mycologic media and appears as a budding yeast at 37°C and as a mold at 25°C. The colonies at 25°C are initially white and moist, turning brown to black with prolonged incubation. Microscopically, the mold appears as delicate branching hyphae with numerous conidia developing in a rosette pattern at the ends of conidiophores (Fig. 52.6). Laboratory confirmation is established by converting the mycelial growth to the yeast form by subculture at 37°C. Alternatively, the organism may be identified immunologically by using the exoantigen test.

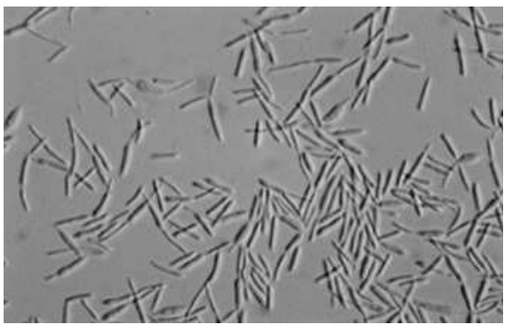


FIGURE 52.5. Microscopic appearance of *Sporothrix schenckii* yeast (original magnification ×1,250).

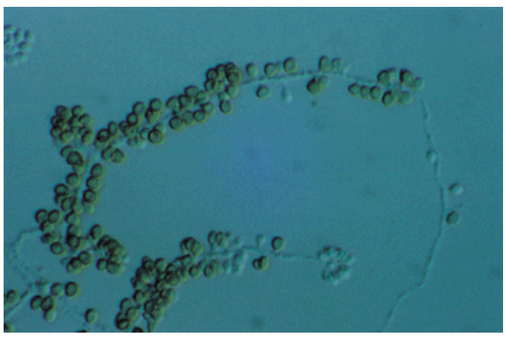


FIGURE 52.6. Microscopic appearance of the mold phase of *Sporothrix schenckii* demonstrates the rosette pattern of conidia at the end of a conidiophore, and dark sessile conidia along the hyphae (Nomarski optics; original magnification ×1,250).

Chromoblastomycosis

Chromoblastomycosis is most commonly seen in the tropics where the warm, moist environment coupled with the lack of protective footwear and clothing predisposes individuals to direct inoculation with infected soil or other organic matter. The organisms most often associated with chromoblastomycosis are dematiaceous (pigmented) fungi of the genera *Fonsecaea*, *Cladosporium*, and *Phialophora* (31,32,51). Multiple species of these genera have been implicated in this disease process. These organisms

are identified according to the pattern and type of sporulation, and in many cases more than one pattern may be observed for a given isolate. For this reason, there remains considerable confusion and controversy regarding the taxonomic placement of many of the dematiaceous fungi that cause chromoblastomycosis. Clinically, the disease is characterized by warty, cauliflowerlike lesions that develop slowly and extend via satellite lesions (31,32,51). The lymphatics are usually not involved and the lesions are painless unless they are involved with secondary bacterial infection. The diagnosis of chromoblastomycosis is generally made by histopathologic examination of infected tissue. Typical lesions demonstrate pseudoepitheliomatous hyperplasia and characteristic brown or copper-colored spherical cells or hyphae known as sclerotic or Medlar bodies (Fig. 52.7). These are the tissue forms of the fungus. The cultures grow as dematiaceous molds but may take weeks to appear and longer to develop the characteristic conidia. Treatment consists of surgery and antifungal therapy. Unfortunately, because most individuals present with advanced disease, therapy is frequently unsuccessful.

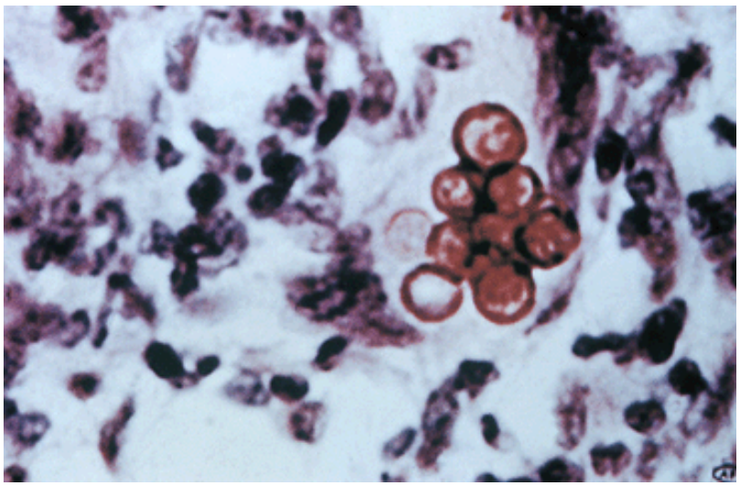


FIGURE 52.7. Histopathologic section of a lesion of chromoblastomycosis demonstrates the presence of sclerotic or Medlar bodies (original magnification ×1,200).

Eumycotic Mycetoma

As with chromoblastomycosis, most mycetomas are seen in the tropics. A mycetoma is defined clinically as a localized chronic granulomatous infectious process involving cutaneous and subcutaneous tissues (55). The process may be quite extensive and deforming with destruction of local bony structures. Most commonly mycetomas are localized to the feet and hands but may also involve other parts of the body such as the back, shoulders, and buttocks. The lesions consist of multiple granulomas and abscesses that suppurate and drain through sinus tracts. The etiologic agents of eumycotic mycetomas encompass a wide range of fungi, including *Pseudallescheria boydii* and *Madurella grisea* (31,32,51,55). Examination of the drainage from sinus tracts may reveal small granules or microcolonies of mycelial filaments. These granules may range from microscopic to 2 mm in diameter and vary according to the infecting species. Histopathologic examination of tissue may reveal hyphal elements and granules (Fig. 52.8). The precise microbiological features depend on the organism involved. Identification of the fungi causing eumycotic mycetomas is by morphology of the asexual conidia formed in culture. It is important to establish the etiology of the infection by culturing specimens because the clinical management depends on the causative organism. An identical picture to eumycotic mycetoma may be caused by various actinomycetes belonging to the genera *Actinomyces*, *Nocardia*, *Streptomyces*, and *Actinomadura* (31,55). These infections, known as actinomycotic mycetoma, may be treated with antibacterial agents, whereas eumycotic mycetomas respond poorly to antifungal therapy and frequently require excision or amputation.

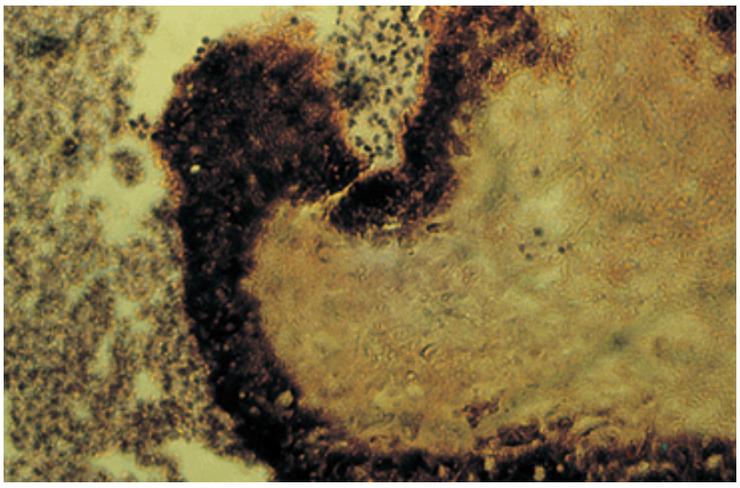


FIGURE 52.8. Tissue section taken from an eumycotic mycetoma illustrates a granule or microcolony (original magnification ×300).

Systemic Mycoses Owing to Dimorphic Fungal Pathogens

The dimorphic fungal pathogens are organisms that exist in a mold form in the natural environment or in the laboratory at 25° to 30°C and in the yeast or spherule form in tissues or when grown on enriched medium in the laboratory at 37°C (31,32,54). The organisms in this group are considered primary or systemic pathogens and include *H. capsulatum* var *capsulatum*, *H. capsulatum* var *duboisii*, *B. dermatitidis*, *C. immitis*, *Penicillium*

marneffei, and *Paracoccidioides brasiliensis* (Table 52.5). These organisms have also been termed endemic pathogens because susceptibility to infection with each organism is generally acquired by living in a geographic area constituting the natural habitat of the particular fungus. In the environment or at 25° to 30°C, these organisms all produce a mycelial form with hyaline, branching, septate hyphae. The tissue form of *H. capsulatum*, *B. dermatitidis*, and *P. brasiliensis* is a budding yeast that may be reproduced on enriched media at 37°C. *C. immitis* forms spherules in tissue or *in vitro* under appropriate conditions. Infection owing to these agents is usually acquired by inhalation of conidia from an environmental source. In recent years, infections with *H. capsulatum* and *C. immitis* have emerged as major opportunistic pathogens in individuals with AIDS (31,54,57). These infections are often acquired primarily but also may be due to reactivation of previous infection in these highly immuno-suppressed individuals.

TABLE 52.5. DIMORPHIC ENDEMIC FUNGAL PATHOGENS

				Growth	
Disease	**Organism**	**Geographic Distribution**	**Tissue**	**Culture (25°C)**	**Culture (37°C)**
Histoplasmosis	*Histoplasma capsulatum* var. *capsulatum*	Ohio and Mississippi River valleys	Intracellular yeast	Mold	Yeast
	Histoplasma capsulatum var. *duboisii*	Africa	Intracellular yeast	Mold	Yeast
Blastomycosis	*Blastomyces dermatitidis*	Ohio and Mississippi River valleys	Large, broad-based budding yeast	Mold	Yeast
Coccidioidomycosis	*Coccidioides immitis*	Southwestern U.S., Mexico, Central and South America	Spherule	Mold	Spherules
Paracoccidioidomycosis	*Paracoccidioides brasiliensis*	Central and South America	Multipolar budding yeast	Mold	Yeast
Penicillosis	*Penicillium marneffei*	Southeast Asia	Intracellular yeast that divide by fission	Mold	Yeast

Histoplasmosis

In 1905, Darling described the causative agent of histoplasmosis as a parasite, then later placed it in the fungal family (44,58). Histoplasmosis is distributed worldwide but is most noted from its endemic regions of the Mississippi-Ohio River valley. Human disease can be linked to either *H. capsulatum* var *capsulatum* (United States) or *H. capsulatum* var *duboisii* (Africa). Disease ranges from slight pulmonary involvement to disseminated disease and is known to involve cells of the reticuloendothelial system.

H. capsulatum var *capsulatum* is most often recovered from respiratory secretions or from blood or bone marrow. Appropriate specimens should be collected and transported to the laboratory as rapidly as possible. It is recommended that the specimen be placed directly on brain-heart infusion (BHI) medium with chloramphenicol and gentamicin to eliminate overgrowth by normal bacterial flora. When an organism is isolated, it should be transferred to potato flakes, Sabhi, or Gorman's media, which is specially formulated to enhance the growth of systemic pathogens and may provide optimum conditions for growth. Cultures incubated at 25°C grow very slowly, resulting in cottony white colonies in 7 to 28 days. The mold phase of *H. capsulatum* is characterized by thin, branching, septate hyphae that produce tuberculate macroconidia and microconidia (Fig. 52.9).

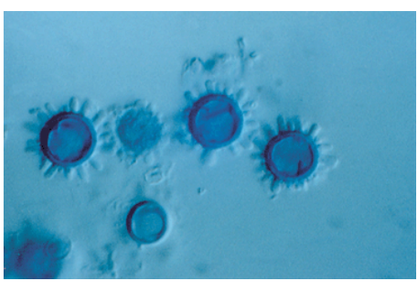


FIGURE 52.9. Microscopic appearance of the mold phase of *Histoplasma capsulatum* var. *capsulatum* demonstrates tuberculate macroconidia (Nomarski optics, original magnification ×1,250).

When *H. capsulatum* is suspected, further laboratory testing is required to differentiate it from saprobic fungi. *H. capsulatum* grows well on cycloheximide-containing medium. Conversion to the yeast form is achieved by incubating the culture at 37°C on BHI infusion agar. This process may take several weeks before conversion is observed. Saprobic fungi will not convert to the yeast form. Other more rapid means of identifying *H. capsulatum* from culture include the exoantigen test and the use of DNA probes (54,59,60 and 61). In the exoantigen test, antigens are extracted from culture material and reacted against anti-*Histoplasma* antibody in an immunodiffusion (ID) test. In the DNA probe test, nucleic acids are extracted from culture material and reacted against a probe specific for *H. capsulatum* ribosomal RNA. With these methods, definitive identification may be made in 2 hours (DNA probe) to 3 days (exoantigen) (54,59,60 and 61).

Additional means of diagnosing histoplasmosis includes histopathologic examination of infected tissue and serology. The tissue phase of *H. capsulatum* is a small budding yeast cell 2 to 5 mm in diameter found almost exclusively within macrophages (Fig. 52.10). It may be visualized with Giemsa or silver stains.

Serologic diagnosis employs both ID and complement fixation (CF) tests and has been quite useful (54,62). An assay to detect antigen in urine has also been developed that has excellent sensitivity and specificity for diagnosis of disseminated infection (54,63,64,65,66 and 67). Unfortunately, its availability is limited to only one laboratory.

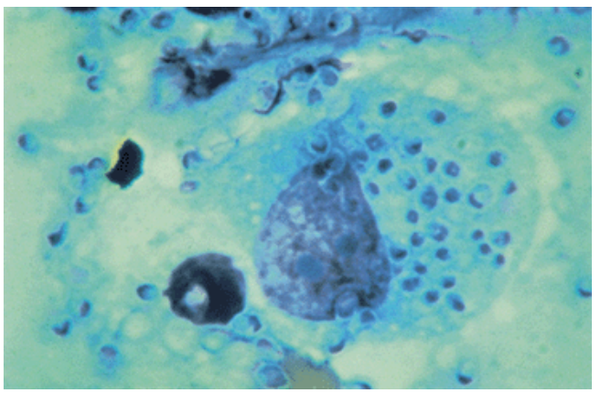


FIGURE 52.10. Microscopic appearance of the tissue phase of *Histoplasma capsulatum* var. *capsulatum* demonstrates intracellular yeast forms (original magnification ×1,000).

Blastomycosis

B. dermatitidis is another of the systemic fungi that become infective on inhalation of conidia. Like other systemic pathogens, *B. dermatitidis* is found in a specific region of the United States, most notably the Ohio-Mississippi River valley region. Disease begins with granulomatous lung involvement but may disseminate to other body sites. The most common form of dissemination is the appearance of cutaneous pustular lesions. Yeast cells can be recovered from biopsies and aspirates of the lesions. Less common involvement may be seen in bone and subcutaneous tissue.

When blastomycosis is suspected, clinical material (e.g., sputum, pus) should be placed on inhibitory media, without cycloheximide, such as BHI agar with chloramphenicol and gentamicin and incubated at 25°C for 7 to 28 days (31,32,54). Growth is very slow and ranges in appearance from a white to a tan fluffy mold. Microscopic examination at this point is not diagnostic. Probable *Blastomyces* isolates must be converted to the yeast phase by subculturing the isolate to a BHI blood agar slant and incubating at 37°C in CO_2 (a candle jar is sufficient). Yeast forms should be visible within a week. Microscopic examination at this point reveals characteristic broad-based budding yeasts.

Although microscopic examination of the yeast gives valuable information, a definitive identification may be made from mycelial phase cultures using the exoantigen test as described for *H. capsulatum*. A commercially available DNA probe assay also allows the rapid identification (2h) of this organism from culture (61). Additional means of diagnosing blastomycosis includes histopathologic examination of infected tissue. The tissue phase of *B. dermatitidis* is a large yeast with broad-based buds (Fig. 52.11). Serologic tests for diagnosing blastomycosis have not been particularly useful. Antigens for CF tests are available, but complement fixing antibodies are absent in as many as 50% of cases.

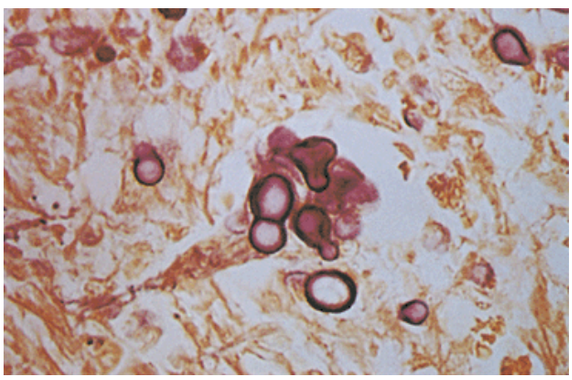


FIGURE 52.11. Microscopic appearance of the tissue phase of *Blastomyces dermatitidis* demonstrates a large, broad-based budding yeast (original magnification ×600).

Coccidioidomycosis

Coccidioidomycosis (Posadas disease, coccidioidal granuloma, valley fever, desert rheumatism, valley bumps, California disease) is commonly a self-limiting, mild to sometimes moderately severe respiratory disease, resulting from the inhalation of arthroconidia produced by the etiologic agent *C. immitis* (31,32,54,57). The organism resides in a narrow ecological niche known as the Lower Sonoran Life Zone, characterized by low rainfall and semiarid conditions. Highly endemic areas include the San Joaquin Valley in California, Maricopa and Pima counties in Arizona, and southwestern Texas. Outside the United States, areas of high endemicity are found in northern Mexico, Guatemala, Honduras, Venezuela, Paraguay, Argentina, and Colombia.

Because *C. immitis* is probably the most virulent of all human mycotic agents, the inhalation of only a few arthroconidia produces primary coccidioidomycosis, which may include asymptomatic pulmonary disease, allergic manifestations (toxic erythema, erythema nodosum or “desert bumps,” and erythema multiforme or valley fever and arthritis (“desert rheumatism”). Primary disease usually resolves without therapy and confers a strong, specific immunity to reinfection, which is detected by the coccidioidin skin test. In patients symptomatic for 6 weeks or longer, the disease progresses to secondary coccidioidomycosis, which may include nodules, cavitary disease, or progressive pulmonary disease; single or multisystem dissemination follows in approximately 1% of this population. Filipinos and blacks run the highest risk of dissemination, with meningeal involvement a common sequela. The gender distribution for clinically significant disease is approximately 9:1 (male:female). The exception is pregnant women among whom the dissemination rate equals or exceeds that for men.

After inhalation, barrel-shaped arthroconidia (2.5 to 4 × 3 to 6 μm) become more rounded as they convert to spherules (Fig. 52.12). At maturity, the spherules (30 to 60 μm in diameter)

produce endospores by a process known as progressive cleavage. Rupture of the spherule walls releases the endospores, which in turn form new spherules. Caution must be exercised with making a diagnosis by histopathology only, as small, empty spherules may resemble the yeast cells of *B. dermatitidis*, and endospores (2 to 5 μm in diameter) can be confused with cells of *C. neoformans*, *H. capsulatum* var *capsulatum*, and *Paracoccidioides brasiliensis*.

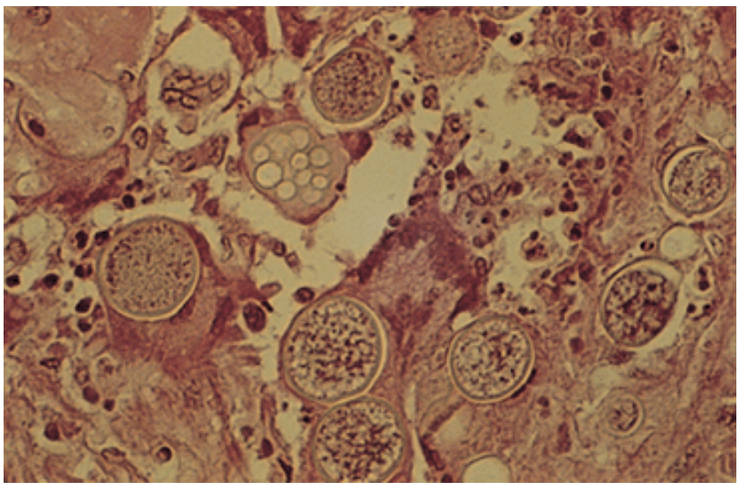


FIGURE 52.12. Tissue section taken from an individual with coccidioidomycosis demonstrates a spherule containing multiple endospores (original magnification ×320).

C. immitis (Rixford et Gilchrist, 1896) is a dimorphic fungus with a variety of mold morphologies at 25°C. Initial growth is white to gray, moist, and glabrous and occurs within 3 to 4 days. It rapidly develops abundant aerial mycelia, and the colony appears to enlarge in a circular "bloom." Mature colonies usually become tan to brown to lavender.

Microscopically, fertile hyphae arise at right angles to the vegetative hyphae and produce alternating (separated by a disjunctor cell) hyaline arthroconidia (Fig. 52.13). When released, conidia have an annular frill at both ends. As the culture ages, the vegetative hyphae also fragment into arthroconidia.

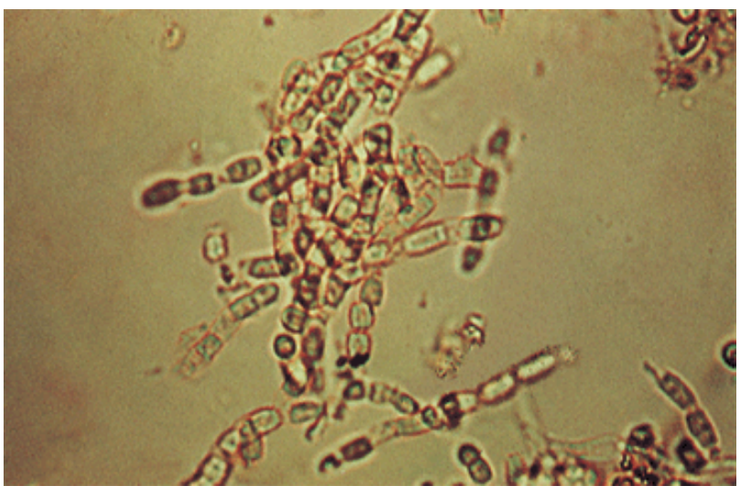


FIGURE 52.13. Microscopic appearance of the mold phase of *Coccidioides immitis* demonstrates arthroconidia (original magnification ×600).

Identification of *C. immitis* from culture may be accomplished by using the exoantigen or DNA probe tests (54,59,61). Additional means of diagnosing coccidioidomycosis include histopathologic examination of infected tissue for the presence of spherules and serologic testing. Several serologic procedures exist for initial screening, confirmation, and prognostic evaluation (31,32,54,62). For initial diagnosis, the combined use of the ID test and the latex particle agglutination test detects approximately 93% of cases. The CF and tube precipitin tests may also be employed for diagnosis as well as for confirmation. Prognostic studies frequently employ serial CF titers.

Paracoccidioidomycosis

Paracoccidioidomycosis (South American blastomycosis, Brazilian blastomycosis, Lutz-Splendore-Almeida disease, paracoccidioidal granuloma) is a chronic, progressive fungal disease endemic to Central and South America (31,32,54,68). Geographic areas of highest incidence (Holdridge plant classification zone) are typically humid, high rainfall areas with acidic soil conditions. As with other systemic mycoses, the gender distribution for clinically significant disease is approximately 9:1 (male:female).

Although the primary route of infection is pulmonary (usually inapparent and asymptomatic), subsequent dissemination leads to the formation of ulcerative granulomatous lesions of the buccal, nasal, and, occasionally, gastrointestinal mucosa; concomitant striking lymph node involvement is also evident. Although the causative agent *P. brasiliensis* has a rather narrow range of temperature tolerance [as evidenced by its predilection for growth in cooler areas of the body (nasal and oropharyngeal)], dissemination to other organs (particularly the adrenals) occurs with diminished host defenses.

Cutaneous and mucosal lesions typically show a pseudoepitheliomatous hyperplasia accompanied by a marked pyogenic and granulomatous response. As a similar tissue reaction is also seen in coccidioidomycosis, and particularly blastomycosis, a definitive diagnosis depends on the demonstration of characteristic yeast cells.

The typical budding yeast cells measure 12 to 40 μm in diameter, with multipolar budding around the periphery, resembling a "mariner's wheel" (Fig. 52.14). These "daughter" cells (2 to 5 μm in diameter) are connected by a narrow base, as contrasted

with the broad base attachment in blastomycosis. Many buds of various sizes may occur, or there may be only a few large buds, giving the appearance of a “Mickey Mouse cap” to the yeast cell. Caution must be exercised when making a diagnosis by histopathology only because the cells in their various sizes and configurations may mimic young spherules of *C. immitis*, the budding cells of *B. dermatitidis*, or the small yeast cells of *H. capsulatum* var *capsulatum*.

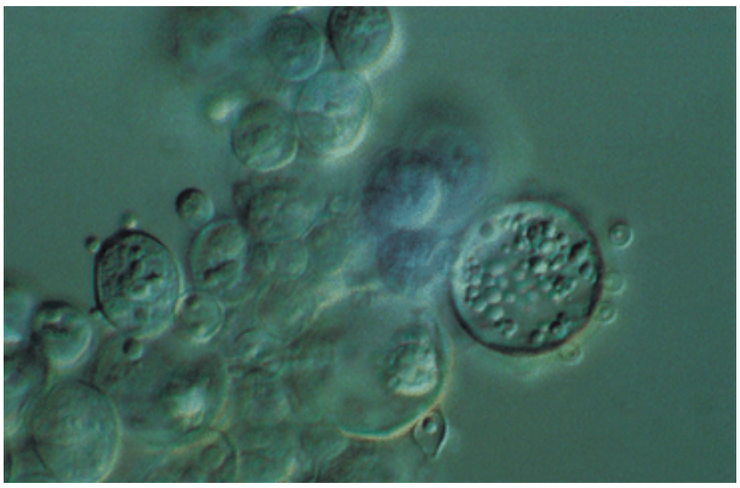


FIGURE 52.14. Microscopic appearance of the tissue phase of *Paracoccidioides brasiliensis* demonstrates multipolar budding yeast forms (Nomarski optics, original magnification ×1,250).

P. brasiliensis is a dimorphic fungus that produces a variety of mold morphologies when grown at 25°C (31,32,54,68). Flat colonies are glabrous to leathery, wrinkled to folded, floccose to velvety, pink to beige to brown, with a yellowish-brown reverse, and resemble those of *B. dermatitidis*. On BHI blood agar at 37°C, the mycelial phase rapidly converts to the yeast phase.

Microscopically, the mold form produces small (2 to 10 µm in diameter), one-celled conidia, generally indistinguishable from those observed with the mold phase of *B. dermatitidis* or the microconidia of *H. capsulatum* var *capsulatum*. When mycelial fragments are transferred to an enriched medium at 37°C, conversion to the characteristic yeast phase occurs.

Diagnosis may be made by culture, histopathologic examination of infected tissues, and serology (31,54,62,68). Both CF and agar gel ID procedures are available for serodiagnosis. The CF test, using yeast-derived antigen, is positive in 80% of active disease. Titers of 1:64 or higher are generally considered diagnostic. Although the CF titer usually decreases with therapy, it may persist in the range of 1:8 to 1:32. The ID test is positive in approximately 95% of active cases, has low cross-reactivity, and diminishes with successful therapy.

Penicillosis

Penicillium marneffei is a dimorphic fungus that appears to be endemic in Southeast Asia. Although infections have been reported in both normal and immunocompromised individuals, most of the cases occur in HIV-infected hosts (9,33,34,35,36 and 37,54). The environmental reservoir for the organism appears to be two species of rat and their burrows (33). Although the fungus does not appear to exist in other parts of the world, increasing international travel makes it likely that infections will be detected for beyond the endemic range of species (33,34 and 35).

Infection with *P. marneffei* is characterized by a variety of signs and symptoms, similar to those of other infections in patients with late-stage AIDS. These include fever, weight loss, generalized lymphadenopathy, anemia, productive cough, pulmonary infiltrates, subcutaneous abscesses and papule-like ulcers (34,35,37,38). The skin lesions may mimic those seen in other infectious processes such as tuberculosis, molluscum contagiosum, and histoplasmosis (36,38).

Detection of *P. marneffei* by histopathologic examination or by culture of blood, bone marrow, skin, and respiratory specimens is the most productive means of making a laboratory diagnosis of infection with this organism (9). In tissue, the presence of 4- to 8-µm diameter cells that divide by fission is characteristic of *P. marneffei* (9). When observed in histocytes, *P. marneffei* can resemble *H. capsulatum*; however, when observed outside of histiocytes, the cells of *P. marneffei* are considerably larger than those of *H. capsulatum* (9). In culture, *P. marneffei* produces a diffuse red pigment and grows as a mold at 25° to 30°C. When incubated at 37°C, it undergoes a phase transition and grows as a single celled yeast form. The red pigment is not produced by the yeast phase of the organism (9).

Opportunistic Mycoses

The opportunistic mycoses occur primarily in immunocompromised individuals, particularly those with malignancies and AIDS, and after major surgery, severe burn injury, and bone marrow and solid organ transplantation. The number and type of fungal pathogens included in this category are increasing rapidly (4,13,14,16,17,27,46,48,49,69,70,71,72 and 73) (Table 52.6). The most well-known causes of opportunistic mycoses include *Candida*, *C. neoformans*, and *Aspergillus* species. The estimated annual incidence of invasive mycoses owing to these pathogens is 73 infections per million population for *Candida*, 66 per million for *C. neoformans*, and 12 per million for *Aspergillus* species (25). In addition to these agents, of increasing importance is the steadily growing list of other opportunistic fungi (Table 52.6). These opportunistic fungi include yeasts other than *Candida* species, nondematiaceous or hyaline molds, agents of zygomycoses, and the pigmented or dematiaceous fungi (13,27,51,56,72,73). Infections caused by these organisms range from catheter-related fungemia and peritonitis to more localized infections involving lung, skin, and paranasal sinuses (13,27,51,56,72,73). Many of these fungi were previously thought to be nonpathogenic and now are recognized causes of invasive

mycoses in compromised patients. Estimates of the annual incidence of these less common mycoses have been virtually nonexistent, but recent data from a population-based survey conducted by the Centers for Disease Control indicate that zygomycosis occurs at a rate of 1.7 infections per million per year, hyalohyphomycosis at 1.2 per million per year, and phaeohyphomycosis at 1.0 per million per year (25).

TABLE 52.6. AGENTS OF OPPORTUNISTIC MYCOSES[a]

I.	*Candida* and other opportunistic yeasts
	Candida albicans
	Candida glabrata
	Candida tropicalis
	Candida parapsilosis
	Candida krusei
	Candida lusitaniae
	Candida dubliniensis
	Cryptococus neoformans
	Malassezia spp
	Rhodotorula spp
	Saccharomyces cerevisiae
	Trichosporon spp
II.	Agents of hyalohyphomycosis
	Aspergillus spp
	Fusarium spp
	Pseudallescheria boydii
	Scedosporium prolificans
	Scopulariopsis spp
III.	Agents of zygomycosis
	Absidia spp
	Mucor spp
	Rhizomucor spp
	Rhizopus spp
IV.	Agents of phaeohyphomycosis
	Alternaria spp
	Bipolaris spp
	Curvularia spp
	Exserohilum spp
V.	*Pneumocystis carinii*

[a] List is not all-inclusive.

The morbidity and mortality owing to infections caused by these opportunistic pathogens are extremely high, and the prevention, diagnosis, and therapy of these infections remain quite difficult (4,13,14,16,17,27,48,49,51,65,69,70,71,72 and 73). Increased recognition of the importance of the opportunistic mycoses has spurred efforts to develop new diagnostic and therapeutic approaches as well as to expand our knowledge of the epidemiology and pathogenesis of fungal infections (11,13,27,51,65,69,73,74 and 75).

Candida and Other Opportunistic Yeasts

The increased incidence of yeast and yeastlike fungi isolated from clinical specimens has magnified the importance of yeast identification to the species level. Gone are the days when *C. albicans*, *Candida* species (not *albicans*), and *Cryptococcus* species are acceptable yeast identifications (15,48,76). *Candida* species has become the number four cause of blood-borne infection in the United States (10,11,18,19,20,21 and 22,24). In addition to *C. albicans*, other *Candida* species such as *C. glabrata*, *C. parapsilosis*, *C. tropicalis*, *C. krusei*, and *C. lusitaniae* are also being recovered more often than in the past and are responsible for more aggressive and refractory infections (1,5,6 and 7,15,20,39,48,71,76,77,78,79,80,81,82,83,84,85 and 86). Emergence of species other than *C. albicans* as causes of infection in some institutions has involved organisms such as *C. glabrata* and *C. krusei* (1,15,20,21 and 22,28,82,84,85 and 86), species that are either innately or relatively resistant to the triazole class of antifungal agents. Although species-specific differences in susceptibility to triazole agents such as fluconazole and itraconazole clearly exist (19,20,21 and 22,24,82,85,86), there are insufficient data to document precisely the role of antifungal drug pressure in the emergence of non-*albicans* species of *Candida* (6,15,24). Nevertheless, given the potential for selection of less responsive organisms by antifungal drug pressure (1,39,85,86), ongoing surveillance may be prudent (15,19,20,21 and 22,24). Clinical labor-atories may need to expand their yeast identification capabilities to facilitate these surveillance efforts (6,11,15,17,19,20,21 and 22,39,48,70).

The *Candida* species are not only frequently present as normal flora of the mucosa, skin, and digestive tract but are also the most notorious agents of yeast infection. These infections range from superficial skin infections to disseminated disease. *C. albicans* certainly reigns as the number one cause of yeast infection in the world today (15,24,28). This organism may be recovered from a variety of sites, including skin, oral mucosa and vagina, as normal flora, but when conditions are altered, it is capable of causing disease in virtually any site. *C. glabrata* is the second most common *Candida* species to incite disease (15,19,20,21 and 22,24,71). This organism is incapable of forming hyphae or pseudohyphae under usual laboratory conditions and assimilates only glucose and trehalose and therefore is easily differentiated from *C. albicans* and other species of *Candida* (50,71). Fungemia with *C. glabrata* is associated with a high rate of serious complications (e.g., endocarditis, endophthalmitis, liver abscess) and high mortality (39,71). Resistance of *C. glabrata* to fluconazole and other azoles has been documented and may contribute to increased colonization and infection with this organism in centers where azoles are used for antifungal prophylaxis (1,6,20,21,24,39,71,85,86). Other notable species of *Candida* include *C. parapsilosis*, *C. tropicalis*, *C. krusei*, and *C. lusitaniae* (15). These isolates are distinguished by differences in their carbohydrate assimilation patterns and by secondary testing (48,50). *C. parapsilosis* has become an important cause of hospital outbreaks of nosocomial infections (11,15,22,77,80,81,83). This organism has a propensity to form biofilms (77,81,83) and is a frequent colonizer of skin (22,87), characteristics that may explain its strong association with catheter-related bloodstream infections (1,77,80,83).

Yeasts are easily distinguished from molds. Macroscopically, they range in color from white to cream or tan, with a few species in the pink to salmon range. Yeasts that are darkly pigmented owing to melanin are associated with polymorphic dematiaceous fungi and are discussed elsewhere in the chapter. Textures vary per strain but can range from mucoid to butterlike to wrinkled to velvety.

Several tests are available to aid in yeast identification (Table 52.7) (48,50,79,88,89,90 and 91). The most basic and easiest test to perform is the germ-tube test. *C. albicans* is capable of forming germ tubes in 3 hours given the proper testing environment. This test requires the use of serum or plasma. Bovine serum can be purchased, but HIV and hepatitis B–negative fresh frozen plasma that has expired from the blood bank can be kept indefinitely at 4°C and used for this testing. A 0.5-mL aliquot of serum or plasma is placed in a disposable test tube. Using a sterile Pasteur pipette, a small portion of an isolated yeast colony is placed into the serum, gently shaken, and placed into a 35°C incubator for 2.5 to 3 hours but no longer. Other yeast isolates are capable of germ-tube formation with extended incubation periods. Care must be taken not to make too heavy a suspension, as this will inhibit germ-tube formation.

True germ tubes are not constricted at the point where they emerge from the blastoconidium (Fig. 52.15). One must take care when reading germ tubes because other isolates are capable of forming false or constricted germ tubes. *C. tropicalis* often germinates under these testing conditions, but germination structures are constricted at the point where they are formed on the yeast cell. After the incubation period, an accurate identification of *C. albicans* can be made when germ tubes are present. *C. dubliniensis* and *C. stellatoidea* are also capable of germ-tube production and may be difficult or impossible to differentiate from *C. albicans* without resorting to additional physiologic, immunologic, or nucleic acid–based testing (31,50,70,88,92). A known germ tube–positive *C. albicans* should be set up in conjunction with all patient testing as a positive control. *Cryptococcus* species should also be set up as the negative control. When test results are negative, the technologist must turn to other procedures for identification. Note that rare *C. albicans* isolates are not capable of germ-tube production.

TABLE 52.7. CHARACTERISTICS OF SELECTED MEDICALLY IMPORTANT YEASTS AND YEAST-LIKE FUNGI

Species	Blastoconidia	Arthroconidia	Chlamydoconidia	Germ Tubes	True Hyphae	Pseudohyphae
Candida albicans	+	–	+	+	+	+
Candida krusei	+	–	–	–	+	+
Candida parapsilosis	+	–	–	–	+	+
Candida tropicalis	+	–	–	–	+	+
Candida glabrata	+	–	–	–	–	–
Cryptococcus spp	+	–	–	–	–	–
Trichosporon spp	+	+	–	–	+	+

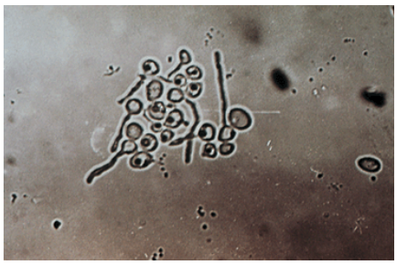


FIGURE 52.15. Germ tube of *Candida albicans* (original magnification ×600).

In addition to the germ-tube test, a rapid colorimetric test based on the detection of *C. albicans*–specific enzymes (L-proline aminopeptidase and ß-galactosaminidase) or the use of agar medium containing chromogenic substrates (CHROMagar Candida, Hardy Diagnostics, Santa Maria, CA) may be useful aids in rapid presumptive identification of yeasts (48,50,79,89,90). CHROMagar combined with observation of microscopic morphology on cornmeal–Tween 80 agar (see below) provides an excellent means of presumptive identification of *Candida* species, including *C. albicans*, *C. tropicalis*, and *C. krusei* (79,89,90). CHROMagar has proven particularly useful in detecting mixed infections with more than one species of *Candida* and formulations incorporating fluconazole into the agar have allowed investigators to simultaneously screen for infections owing to *Candida* and to detect potential fluconazole resistance (75,89,90). Although a single presumptive identification test cannot be used alone for identifying yeasts, a positive germ-tube or colorimetric test or characteristic appearance on CHROMagar medium is generally considered diagnostic for *C. albicans* and further identification is not indicated (48,79,90).

Although sugar fermentations are very valuable, they are very time and labor intensive, thus making them impractical for the routine microbiology laboratory. Sugar assimilations, however, are a series of tests that give significant identification information and can easily be incorporated into any laboratory (31,32,50). These tests identify which sugars a yeast is capable of using as a sole source of carbon. Assimilation results may be determined from means as sophisticated as the automated identification systems to simple manual procedures. Individual laboratories should assess which method is most practical for them.

Although many kits are available for yeast identification, the API 20C yeast identification strip (bioMérieux, Hazelwood, MO) remains a widely used commercial method for assimilation testing. A series of sugars is lyophilized and placed into wells on a plastic strip. Yeasts to be identified are suspended in an agar basal medium and pipetted into the wells. The strips are incubated at 30°C in a humidifier chamber and read at 24, 48, and 72 hours. As sugars are assimilated, the well becomes turbid with fungal growth. Wells in which the sugar is not assimilated remain clear. A code is derived from the assimilation pattern. This code is matched against a computerized database for identification. Identifications are accompanied by a percentage that indicates the probability that the identification is correct. Although this approach is very reliable, other auxiliary testing should accompany assimilation results before a final identification is made (31,32,50). Automated systems are available for yeast identification. Many of these systems use enzyme as well as assimilation reactions to aid in yeast identification.

One test that gives valuable information for separating the clinically significant yeasts is the KNO_3 assimilation (91). A positive KNO_3 assimilation test will turn the medium blue, whereas KNO_3-negative isolates will turn the medium yellow. *Cryptococcus albidus* works well as a positive control, whereas *C. albicans* is a suitable negative control.

Yeast isolates producing the enzyme urease can easily be detected by a simple urea agar. This is another fairly rapid, easily read test that aids in yeast identification. *Cryptococcus* and *Rhodotorula* species are positive, as are a few *Candida* and *Trichosporon* species. Media for detecting urea hydrolysis may be obtained commercially. Positive test results turn the media bright fuchsia pink, whereas negative results will cause little, if any, change in the medium. *C. albidus* works well as a positive control, whereas *C. albicans* functions well as a negative control.

Yeast morphology on cornmeal agar is extremely important when determining whether an isolate produces one or more of the following structures: hyphae, true pseudohyphae, arthroconidia or blastoconidia (31,32,50,88). The morphology is a very important key to yeast identification. As noted previously, when combined with appearance on CHROMagar, microscopic morphology on cornmeal agar allows the identification of a number of commonly encountered yeast isolates (79). A cornmeal plate is inoculated with two parallel 1-cm long streaks. The streaks are then covered with a coverslip and allowed to incubate at room temperature for 48 hours. The Petri dish can be placed on the microscope stage and viewed with a 20× objective. One or more of five morphology types should be observed (Table 52.7 and Fig. 52.16). Blastoconidia are the characteristic budding yeast forms most often seen on direct mounts. *C. albicans* also forms terminal thick-walled chlamydoconidia under some conditions. Pseudohyphae occur when the blastoconidia germinate

to form a filamentous mat. The cross walls are not true separations but rather constrictions, whereas true hyphae remain parallel at cross walls with no indentation. The fifth morphologic type seen with yeasts is arthroconidia. This type begins as true hyphae but begins to break apart at the cross walls as the structures mature. Rectangular fragments of hyphae should be accompanied by blastoconidia to consider an isolate a yeast.

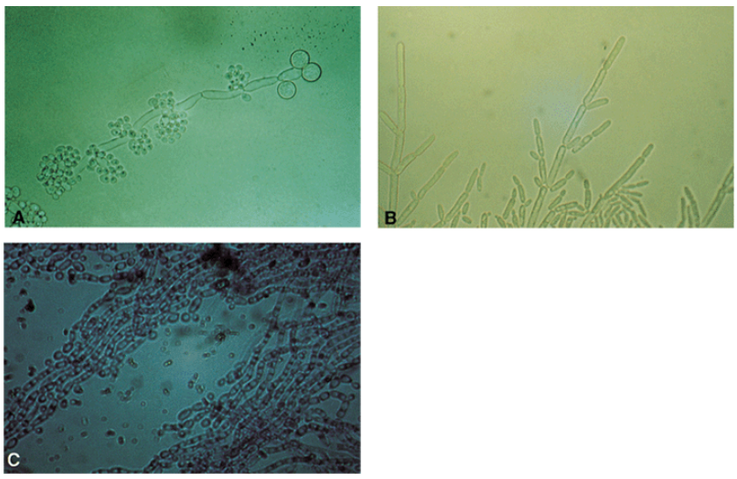


FIGURE 52.16. Yeast morphologies observed on cornmeal agar (original magnification ×600). **A:** True hyphae, blastoconidia, and chlamydoconidia. **B:** Pseudohyphae. **C:** Arthroconidia.

Temperature studies also offer additional information for yeast identification. *Cryptococcus* species have weak growth at 35°C and no growth at 42°C. Several *Candida* species have the ability to grow well at temperatures as high as 45°C. The inability to grow at temperatures of 42° to 46°C is one of the characteristics used to distinguish *C. dubliniensis* from *C. albicans* (92).

Any one of these tests alone, with the exception of the germ-tube test, is not sufficient for proper identification, but when tested in concert, proper identification is often easily accomplished. Armed with the above procedures, the technologist should be able to identify the most commonly encountered yeast isolates.

Among the opportunistic yeasts that merit special mention are the *Cryptococcus* species, which are important causes of meningitis and pulmonary disease (16,31,32,48,50,93,94). These organisms are surrounded by a capsule that causes the characteristic mucoid appearance common with this genus. The capsule can be detected surrounding the budding yeast in spinal fluid with the aid of India ink. The ink stains the cerebrospinal fluid (CSF) while leaving clear halos around each individual yeast cell (Fig. 52.17). *C. neoformans* is the most noted pathogen in this group. This has become one of the major causes of opportunistic infection in AIDS patients. This organism is commonly found in soil contaminated with pigeon droppings and is most likely inhaled before infection. *Cryptococcus* species are noted for not producing true hyphae or pseudohyphae on cornmeal agar (Table 52.7). Some are nitrate positive, whereas all are urease positive. Sugar assimilations vary with each species. The diagnosis of infection owing to *C. neoformans* may be made by culture of blood, CSF, or other clinical material; however, it is most commonly made by detecting capsular antigen (cryptococcal

antigen) in CSF or serum. Detection of cryptococcal antigen is accomplished by using one of several commercially available latex agglutination or enzyme immunoassay (EIA) kits. These assays have been shown to be highly sensitive and specific for the diagnosis of cryptococcal disease (31,32,50,94).

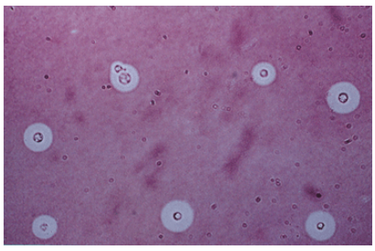


FIGURE 52.17. India ink preparation demonstrates the capsule of *Cryptococcus neoformans* (original magnification ×600).

Rhodotorula species are closely related to the cryptococci in that they bear a capsule and are urease positive. Some species are also nitrate positive. This group of yeasts are remarkable for their bright salmon pink color. They are less common agents of disease but have been known to cause opportunistic infection.

T. asahi (beigelii), a common cause of white piedra, is an emerging agent in disseminated infection (95). This organism has been the culprit in several outbreaks of infection in cancer patients. Its treatment is complicated by the fact that this organism tends to be resistant to amphotericin B, the one antifungal drug known to kill most fungi (95). Amphotericin B is the number one choice for treatment of life-threatening fungal infection but often is not able to clear these infections. *Trichosporon* species are noted for their production of arthroconidia on cornmeal agar (Table 52.7). This, along with the other testing, will ease the identification of this organism. *T. asahi (beigelii)* shares antigenic determinants with *C. neoformans*, and thus serum and CSF from patients with disseminated trichosporonosis may give a positive reaction with tests for cryptococcal antigen (31,32).

There is one last yeast that bears mentioning. It will be negative for all tests. *M. furfur* is a lipophilic organism that requires long-chain fatty acids to live. This organism is a common skin colonizer and when present may cause a condition known as tinea versicolor (52). This condition occurs as hyperpigmented or hypopigmented areas on the skin. It usually causes no discomfort but is undesirable cosmetically. *M. furfur* is also known to cause more serious infections such as fungemia and catheter-related infections (49). Outbreaks of infection were noted in neonatal intensive care units where premature infants were receiving intravenous lipid supplementation. The growth of the organism is favored by the lipid-rich formula, and the organism gains access to the bloodstream via the catheter, thereby causing fungemia (48,49). Treatment usually is not necessary because infection subsides when the lipid infusion is stopped and the intravascular lines are removed. Since the first reports in neonates, this organism has also caused similar syndromes in anorexic patients also receiving lipid infusions through vascular catheters. This organism should be considered when yeasts are seen microscopically in blood culture bottles or clinical material but no organisms are recovered when cultured on routine agar medium. To isolate this organism on agar medium, the plates must be inoculated and then overlaid with sterile olive oil (48,49 and 50).

Aspergillosis

Aspergillosis encompasses a broad spectrum of diseases caused by members of the genus *Aspergillus*. The range of pathologic processes include (a) toxicity owing to ingestion or inhalation of mycotoxins produced by aspergilli, (b) allergies (allergic bronchopulmonary aspergillosis, extrinsic allergic alveolitis), (c) colonization in preformed cavities in the lung (fungus ball), (d) invasive, inflammatory, granulomatous, necrotizing disease of the lungs and other organs, and (e) in the host with abrogated immunity and systemic and fatal disseminated disease (4,11,13,14,16,25,27,31,32,40,41,42 and 43,96,97).

Clinical manifestations, apparently related to the immune status of the host and fungal exposure, are ear infections (incited primarily by *A. niger*), mycotic keratitis (keratomycosis), sinusitis, cutaneous aspergillosis (often sequelae of disseminated disease), pulmonary aspergillosis, central nervous system involvement, bone disease, endocarditis, and fatal disseminated aspergillosis (in organ transplant and leukemic patients) (13,14,16,31,32,40,41,97).

As with other ubiquitous fungi, the diagnosis of aspergillosis necessitates caution when evaluating the isolation of an *Aspergillus* species from clinical specimens. Recovery from surgically removed tissue or sterile sites, accompanied by positive histopathology (moniliaceous, septate, dichotomously branching hyphae) should always be considered significant; isolation from normally contaminated sites requires closer scrutiny.

Although approximately 19 *Aspergillus* species have been documented as agents of human disease, the major clinical entities are caused by *A. fumigatus*, *A. flavus*, *A. niger*, and *A. terreus* (Table 52.8). The increased, clinically significant isolation of *A. nidulans* merits its recognition as a possible emerging pathogen.

TABLE 52.8. DIFFERENTIAL CHARACTERISTICS OF SOME CLINICALLY SIGNIFICANT *ASPERGILLUS* SPECIES

Species	Colonial Morphology	Microscopic Morphology
A. fumigatus	White to gray-green; slate gray with age; reverse yellow to variable; growth at 45–50°C	Smooth-walled conidiophore terminates in dome-shaped vesicle bearing uniseriate phialides from the upper portion only; conidial heads strongly columnar; conidia echinulate
A. flavus	Yellow to lime-green; reverse pinkish; brown to black sclerotia may be present	Rough; thick-walled conidiophores bear globose vesicles; uniseriate or biseriate phialides cover the majority of the vesicle; conidial heads radiate; conidia echinulate
A. nidulans	Dark green to yellow-green; reverse pinkish	Short, sinuous, pale brown conidiophores bearing hemispherical vesicles; biseriate phialides cover upper half of vesicle; globose, rough-walled conidia from short columns; cleistothecia and Hulle cells present
A. niger	White to carbonaceous black; reverse pale yellow	Large, smooth, thick-walled conidiophores bearing globose vesicles; biseriate phialides cover entire vesicle; conidial heads are radiate; conidia thick-walled, echinulate
A. terreus	Cinnamon-buff to brown; reverse dull yellow to brown	Delicate fruiting head compared with above species; short, flexous conidiophores terminate in small, domelike vesicles; biseriate phialides cover most of the vesicle; small conidia to elliptical; single aleurioconidia formed on submerged hyphae; conidial heads columnar

Culture Methods and Identification

Most etiologic agents of aspergillosis grow readily on routine mycologic media lacking cycloheximide, at both 25° and 35°C; *A. fumigatus*, a thermotolerant species, has a maximum growth temperature of 45° to 50°C. Species-level identification of the major human pathogens can usually be made by observing cultural and microscopic characteristics from growth on potato dextrose agar, whereas the less commonly seen isolates require standardized media (Czapek-Dox and 2% malt agar) to utilize identification keys (31,32,41). Microscopic morphology [conidiophores, vesicles, metulae, conidiogenous cells (phialides), conidia] is best observed with the aid of a slide culture (Table 52.8 and Fig. 52.18).

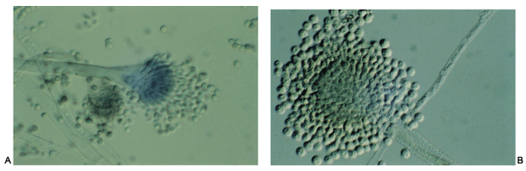


FIGURE 52.18. Microscopic appearance of *Aspergillus* species. **A:** *Aspergillus fumigatus* (Nomarski optics, original magnification ×625). **B:** *Aspergillus flavus* (original magnification ×1,250).

Serologic Diagnosis

Both immunodiffusion and countercurrent immunoelectro-phoresis tests are available for the detection of antibodies to *Aspergillus* antigens, as these antibodies are generally uncommon in sera from healthy individuals (62,96,97). They have limitations, however, in the setting of the immunocompromised host, who is frequently unable to mount an antibody response. Recent efforts for serodiagnosis have been directed toward demonstration of the *Aspergillus* antigen galactomannan employing countercurrent immunoelectrophoresis, radioimmunoassay (RIA), latex agglutination, and enzyme immunoassay methodologies (42,43,47,96,97). Although antigen detection appears promising, further evaluation is needed to establish its utility in diagnosing invasive disease early enough to permit efficacious therapy.

Zygomycosis

Zygomycetes causing human disease encompass several genera in the division of Zygomycota (Table 52.9). The principal human pathogens in the class Zygomycetes are encompassed by two orders: the Mucorales and the Entomophthorales. In the order Mucorales, pathogenic genera include *Rhizopus*, *Mucor*, *Absidia*, *Rhizomucor*, *Saksenaea*, *Cunninghamella*, *Syncephalastrum*, and *Apophysomyces* (31,32,56). As with other saprobic mold-fungi, immunosuppression and trauma are frequent predisposing factors in zygomycosis. Metabolic acidosis, induced by several conditions, appears to be the foremost predisposing factor in classic rhinocerebral zygomycosis caused primarily by *Rhizopus arrhizus*. The predilection of these fungi for the vascular system with resulting infarction and necrosis, combined with their rapid growth rate, makes them some of the most fulminant agents of mycotic disease. The order Entomophthorales contains two pathogenic genera, *Conidiobolus* and *Basidiobolus*. These agents generally incite a more chronic, granulomatous infection of subcutaneous tissues.

TABLE 52.9. PREDOMINANT AGENTS OF ZYGOMYCOSIS

Superkingdom	Eukaryota
Kingdom	Fungi
Division	Zygomycota
Class	Zygomycetes
Order	Mucorales
Family	Mucoraceae
Genera/species	*Rhizopus arrhizus*
	Rhizopus rhizopodiformis
	Mucor circinelloides
	Absidia corymbifera
	Rhizomucor miehei
	Apophysomyces elegans
Family	Syncephalastraceae
	Syncephalastrum racemosum
Family	Mortierellaceae
	Mortierella molfii
Family	Cunninghamellaceae
	Cunninghamella bertholletiae
Family	Saksenaeceae
	Saksenaea vasiformis
Family	Thamnidiaceae
	Cokeromyces recurvatus
Order	Entomophthorales
Family	Ancylistaceae
	Conidiobolus coronatus
Family	Basidiobolaceae
	Basidiobolus ranarum

Histopathology

Histopathologically, fungi in the order Mucorales are seen as ribbonlike, aseptate or sparsely septate moniliaceous hyphae (Fig. 52.19).

In contrast to Hyphomycetes, their diameter often exceeds 10 µm. Because the Zygomycetes are an extremely ubiquitous group of fungi, demonstration of characteristic fungal elements in tissue merits considerably more importance than the mere isolation of a mucoraceous fungus. The tissue presentation seen by members of the Entomophthorales is that of more regularly septate hyphae, often, but not consistently, accompanied by an eosinophilic halo termed the Splendore-Hoeppli phenomenon. The isolation of these agents always merits attention because they have not been found to be laboratory contaminants.

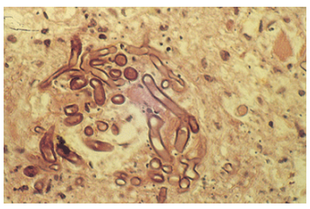


FIGURE 52.19. Tissue section from an individual with zygomycosis demonstrates ribbonlike, aseptate hyphae (original magnification ×600).

Culture Methods

Specimens for culture may be obtained from various sites, but biopsies at the edges of necrotic sites are the most rewarding. As hyphal elements are fragile, tissue should be minced, rather than ground, and processed immediately. If a delay in transport is unavoidable, store at room temperature because Zygomycetes rapidly lose viability when refrigerated.

Growth of most genera of pathogenic Mucorales is readily visible (within 12 to 18 hours) on standard mycologic media lacking cycloheximide and occurs at either 25° or 37°C. It is easily recognized by the rapid growth rate (filling the tube or Petri dish in a matter of days) and gray to brown woolly colonies. Further identification to the genus and species level is based primarily on microscopic morphology (Table 52.10 and Fig. 52.20) because as no practical biochemical or serological methods exist (31,32,56).

TABLE 52.10. DIFFERENTIAL CHARACTERISTICS OF SOME CLINICALLY SIGNIFICANT ZYGOMYCETES GENERA

Genus	Colonial Morphology	Microscopic Morphology
Rhizopus	Rapid growth, cottony, white; becoming gray to brown	Long, unbranched sporangiophores, solitary or in clusters arising from rhizoids; sporangia dark, globose; containing mostly angulated, striated sporangiospores; columella hemispheric
Mucor	Rapid growth, cottony, white; becoming gray to brown	Rhizoids absent; branched or simple sporangiophores arise from hyphae; sporangia globose; sporangiospores globose to cylindrical; columella variable
Absidia	Rapid growth, cottony, gray to olive	Finely branched sporangiophores arise from stolons between rhizoids; pear-shaped sporangia contain globose sporangiospores; conspicuous apophysis merges with round columella; septum usually present in sporangiophore below apophysis
Rhizomucor	Rapid growth, white, becoming gray-brown, low aerial mycelium	Sporangiophores arise from branched aerial hyphae or stolons; rhizoids poorly developed; globose sporangia contain small, round sporangiospores; thermotolerant, growth at 50–55°C
Cunninghamella	Rapid growth, white, becoming dark gray	Erect sporangiophores terminate in globose vesicles; additional smaller, whorled branches of sporangiophores occur beneath primary vesicle; vesicles bear one-spored sporangiola on short stalks (denticles)
Syncephalastrum	Rapid growth, white, becoming gray to black	Branched sporangiophores terminate in globose vesicles; tubular sporangia (merosporangia) contain single row of globose sporangiospores; rhizoids usually present
Saksenaea	Rapid growth, white, woolly	Induce sporulation on nutritionally deficient media; flask-shaped sporangia on short sporangiophores; rhizoids at base of sporangiophore
Apophysomyces	Rapid growth, white, cottony	Induce sporulation on nutritionally deficient media; sporangiophores arise from hyphae, with supporting ‘foot-cells’; pear-shaped sporangia, pronounced dark apophysis; sporangiospores oblong
Mortierella	Rapid growth, white, cottony, garlic-like odor; colonies appear as overlapping “rosettes”	Short tapering sporangiophores arise from rhizoids; small, multispored sporangia rapidly deliquesce; sporangiospores kidney-shaped with double wall
Cokeromyces	Moderate growth, tan, flat, tenacious, becoming brown-gray	Dimorphic; yeast at 37°C; sporangiole stalks arise from vesicle and recurve backward, terminating in multispored sporangiola; zygospores abundant

From Goodman NL, Rinaldi MG. Agents of zygomycosis. In: Balows A, Hausler WJ Jr, Hermann K, et al., eds. *Manual of clinical microbiology*, 5th ed. Washington, DC: American Society for Microbiology, 1991:674–692, with permission.

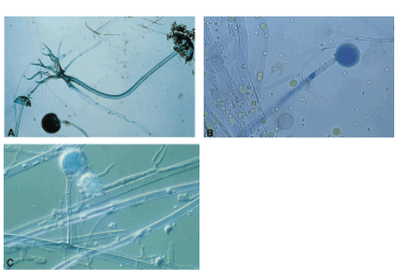


FIGURE 52.20. Microscopic appearance of selected Zygomycetes. **A:** *Rhizopus* species (original magnification ×300). **B:** *Mucor* species (original magnification ×600). **C:** *Absidia* species (Nomarski optics, original magnification ×625).

The Entomophthorales generally require 3 to 4 days before growth is apparent and are characterized by thin, flat, gray to pale yellow waxy colonies adhering to the agar. As they mature, conidia are forcibly expelled onto the Petri dish lid, giving it a ground-glass appearance.

Hyalohyphomycosis

The term hyalohyphomycosis includes all infections owing to nondematiaceous molds that appear as colorless, septate, branched or unbranched hyphal elements in tissue (Fig. 52.21) (27,31,32,41). Numerous genera of moniliaceous (hyaline) Hyphomycetes

are documented agents of hyalohyphomycosis, and the list continues to grow (Table 52.11).

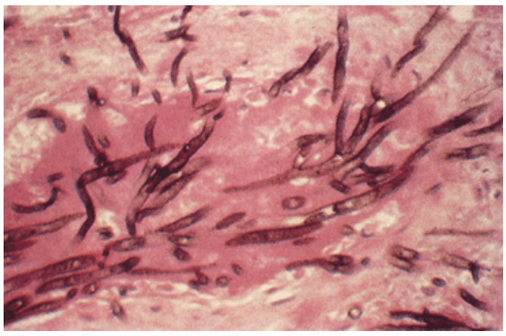


FIGURE 52.21. Tissue section from an individual with hyalohyphomycosis owing to *Fusarium* species (original magnification ×600).

TABLE 52.11. DOCUMENTED AGENTS OF HYALOHYPHOMYCOSIS[a]

Acremonium	*Graphium spp*
falciforme	*Lecythophora*
kiliense	*hoffmannii*
recifei	*mutabilis*
restrictum	*Paecilomyces*
Beauveria bassiana	*variotii*
Chrysosporium spp	*Penicillium*
Cylindrocarpon spp	*chrysogenum*
Fusarium	*citrinum*
chlamydosporum	*commune*
dimerum	*expansum*
moniliforme	*lilacinum*
napiforme	*marneffei*
oxysporum	*spinulosum*
proliferatum	*Scedosproium*
semitectum	*apiospermum*
solani	*inflatum*
sporotrichoides	*Scopulariopsis*
Geotrichum	*brevicaulis*
candidum	*candida*
penicillatum	

[a] This list is not all-inclusive.

Although infections caused by most of these fungi are relatively rare, they appear to be increasing in incidence (12,13 and 14,16,17,27,44,45 and 46,72,73). Most disseminated infections are thought to be acquired by the inhalation of spores or by the progression of previously localized cutaneous lesions. In this chapter, the discussion of specific genera is limited to selected clinically important hyaline mold-fungi, including *Fusarium*, *Pseudallescheria*, *Scedosporium*, and *Scopulariopsis*.

Histopathology

Most agents of hyalohyphomycosis have the same tissue morphology and elicit a similar pathologic response (Fig. 52.21). Consequently, it is impossible to make a correct diagnosis of the etiologic agent based on histopathology alone, and all suspected

material should be submitted for fungal culture and identification (27,31,32,41,47).

Culture Methods and Characteristics of Selected Medically Significant Hyphomycetes

The common etiologic agents of hyalohyphomycosis grow readily on routine mycologic media at both 25° and 35°C. Growth on cycloheximide-containing media shows considerable variability, and this criterion should not be used when determining the potential pathogenicity of an isolate. Although the isolation of moniliaceous agents presents few problems, their identification is occasionally hampered by lack of the production of characteristic structures, e.g., asexual propagules. Growth on potato dextrose agar (PDA) or potato flakes agar generally promotes profuse conidiation and produces the characteristic macroscopic morphologies necessary for separation of genus- or species-level identification. Microscopic morphology is best observed with the aid of a slide culture.

Fusarium species have been recognized with increased frequency as causes of disseminated infection in immunocompromised patients (12,41,44,45 and 46). The most common species isolated from clinical specimens include *F. moniliforme*, *F. oxysporum*, and *F. solani* (41,44,46). The hallmark of disseminated fusariosis is the appearance of multiple purpuric cutaneous nodules with central necrosis. Biopsy of these nodules generally reveals branching, hyaline, septate hyphae invading dermal blood vessels (Fig. 52.21). Cultures of the biopsy material and of blood are useful in establishing the diagnosis of *Fusarium* infection. Although blood cultures are virtually always negative in invasive infections owing to *Aspergillus* species, approximately 75% of patients with fusariosis will have positive blood cultures. In culture, colonies of *Fusarium* species are rapidly growing, cottony to woolly, flat, and spreading. Colors may include blue-green, beige, salmon, lavender, red, violet, and purple. Macroscopic identification should only be made from colonies grown on PDA, as color may be quite variable and is highly medium dependent (31,41,45,46). Microscopically, *Fusarium* species are characterized by the production of both macroconidia and microconidia (Fig. 52.22) (8,27,31,41,45,46). Microconidia are single- or double-celled, ovoid to cylindrical, and generally borne as mucous balls or short chains. Macroconidia are fusiform or sickle shaped and many celled (Fig. 52.22). *Fusarium* species may be misidentified as *Acremonium* when macroconidia are not observed.

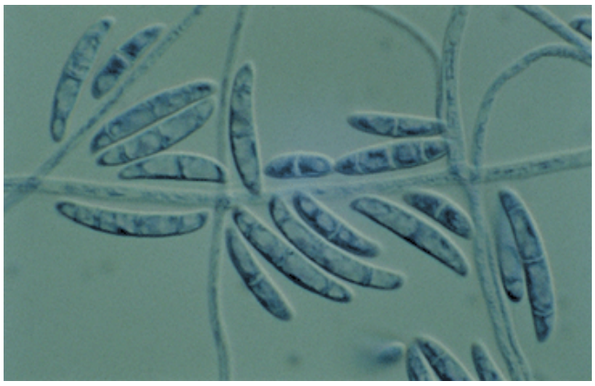


FIGURE 52.22. Microscopic appearance of two-celled microconidia and multicelled, sickle-shaped macroconidia of *Fusarium solani* (original magnification ×625).

Pseudallescheria boydii may be readily isolated from soil and is an occasional cause of mycetoma worldwide; however, it is also the cause of serious disseminated and localized infection in immunocompromised patients (31,32,51,55). In addition to widespread disseminated disease, *P. boydii* has been reported to cause corneal ulcers, endophthalmitis, sinusitis, pneumonia, endocarditis, meningitis, arthritis, and osteomyelitis. *P. boydii* is indistinguishable from *Aspergillus* species and other agents of hyalohyphomycosis on histopathologic examination. Such distinction is important clinically, as *P. boydii* is resistant to amphotericin B and susceptible to miconazole, intraconazole, and voriconazole (31,32,98). Thus, appropriate cultural and mycologic identification is necessary for optimal therapy. The taxonomy and nomenclature of this fungus have caused a tremendous amount of confusion. *P. boydii* (earlier names are *Allescheria boydii* and *Petriellidium boydii*) is a homothallic fungus, i.e., it does not require two mating strains to produce the teleomorph or sexual form. Hence, some strains may produce cleistothecia containing ascospores in culture. More commonly, the anamorph (asexual form) is seen (i.e., *Scedosporium apiospermum*). Macroscopically, colonies are moderately rapidly growing, woolly to cottony, and initially white, becoming smoky brown to green. Microscopically, conidia are one celled, elongate, and pale brown and are borne singly or in balls on either short or long conidiophores (31,32,51).

Scedosporium prolificans (formerly *Scedosporium inflatum*) is a potentially virulent and highly aggressive emerging agent of hyalohyphomycosis (31,73). Although far less important than *Fusarium* or *P. boydii*, infections owing to *S. prolificans* are associated with soft-tissue trauma and are characterized by widespread local invasion, tissue necrosis, and osteomyelitis (73). *S. prolificans* resembles *P. boydii* (*S. apiospermum*) in macroscopic and microscopic morphology. The formation by *S. prolificans* of annelloconidia in wet clumps at the apices of annellides with swollen bases is the most useful characteristic in differentiating this organism from *P. boydii* (*S. apiospermum*) (Fig. 52.23).

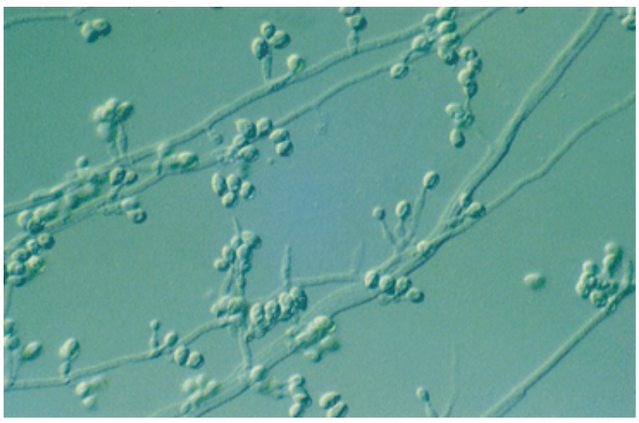


FIGURE 52.23. Microscopic appearance of *Scedosporium prolificans* demonstrates characteristic annellides with swollen bases (Nomarski optics, original magnification ×625).

Scopulariopsis species are ubiquitous soil saprobes that have been rarely implicated in invasive human disease. *S. brevicaulis* is

the most frequently isolated species. Infection is usually confined to the nails; however, serious deep infection has been noted recently in neutropenic leukemia patients and after bone marrow transplantation (31,32,41). Both local and disseminated infections have been described with involvement of the nasal septum, skin and soft tissues, blood, lungs, and brain. Diagnosis is made by culture and histopathology. *Scopulariopsis* species grow moderately to rapidly at room temperature on virtually all laboratory media. Colonies are initially glabrous (smooth), becoming granular to powdery. Young colonies are white and become very pale buff to buff brown depending on the species. Conidiophores are simple or branched; the conidiogenous cells are annellides that form solitarily or in clusters, or may form a scopula, e.g., a "broomlike" structure similar to that seen with *Penicillium* species. Annelloconidia are smooth initially, become rough at maturity, are shaped like light bulbs, and form in basipetal chains (Fig. 52.24).

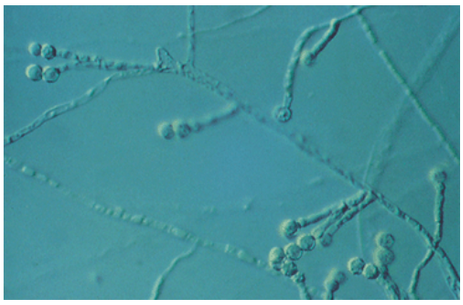


FIGURE 52.24. Microscopic appearance of *Scopulariopsis* species shows anneloconidia in basipetal chains (Nomarski optics, original magnification ×625).

Phaeohyphomycosis

Phaeohyphomycosis is defined as tissue infection caused by dematiaceous (pigmented) hyphae and/or yeasts (31,32,51). Infections owing to dematiaceous fungi constitute a significant and increasingly prevalent group of opportunistic fungal diseases and may take the form of disseminated disease or become localized to the lung, paranasal sinuses, or central nervous system (27,31,32,51). Primary inoculation, resulting in a localized subcutaneous infection, occurs commonly in underdeveloped countries. The dematiaceous fungi that have been documented to cause human and animal infection encompass a large number of different species; however, most human infections have been caused by *Alternaria*, *Bipolaris*, *Cladosporium*, *Curvularia*, and *Exserohilum* species (31,32,51).

Histopathology

In tissue, hyphae with or without yeast forms are present. Because the pigmentation (owing to the presence of melanin in hyphal walls) may not be readily apparent, these fungi may be confused with *Aspergillus* species on histopathologic examination. Staining with the Fontana-Masson technique (a melanin-specific stain) may help visualize the dematiaceous elements (Fig. 52.25). Alternatively, the pigmented hyphae/yeasts may be detected by examining an unstained tissue section by bright-field microscopy (51).

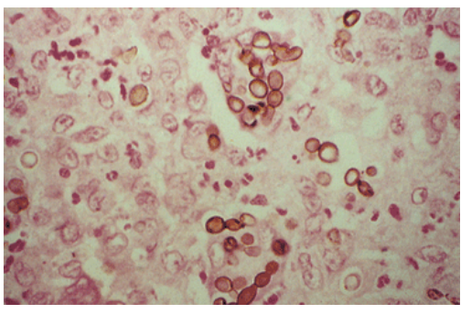


FIGURE 52.25. Tissue section from an individual with phaeohyphomycosis caused by *Bipolaris spicifera*. Stains for melanin, such as Fontana-Masson, may be used to accentuate and confirm the presence of melanin in the fungal cell walls (Fontana-Masson stains, original magnification ×260).

Clinical and Mycologic Characteristics of Selected Medically Important Phaeohyphomycetes

The dematiaceous fungi differ considerably in the clinical spectrum of infection and response to therapy (31,51). Furthermore, the different genera are not readily distinguished on histopathologic examination. Thus, an accurate microbiologic diagnosis based on culture of the infected tissue is important for optimal clinical management of infections attributable to these fungi.

Alternaria species are important causes of paranasal sinusitis in both healthy and immunocompromised patients. Other sites of infection include skin and soft tissue, cornea, lower respiratory tract, and peritoneum. *A. alternata* is the best documented human pathogen in this genus. In culture, *Alternaria* colonies are rapidly growing, cottony, and gray to black. The conidiophores are usually solitary and simple or branched. The conidia develop in branching chains and are dematiaceous, muriform, and smooth or rough and taper toward the distal end with a short beak at their apices (Fig. 52.26) (31,51).

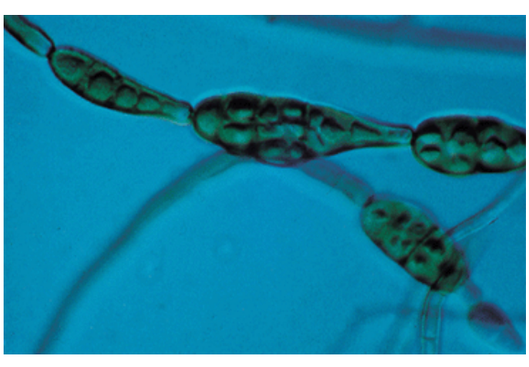


FIGURE 52.26. Microscopic appearance of *Alternaria alternata* (Nomarski optics, original magnification ×1,250).

Cladosporium species are occasionally isolated from cutaneous, eye, and nail infections. These fungi are rapidly growing with a velvety, olive gray to black colony. The conidiophores arise from the hyphae and are dematiaceous, tall, and branching. The conidia may be smooth or rough and single to several celled and form in branching chains at the apex of the conidiophore (31,51). The conidia at the branch points of the chains may have a shieldlike shape and are referred to as shield cells.

Curvularia species are ubiquitous inhabitants of the soil and have been implicated in both disseminated and local infection. Sites of infection include endocarditis, local catheter infection, nasal septum and paranasal sinuses, lower respiratory tract, skin and subcutaneous tissue, bones, and cornea (31,32,51). In clinical specimens, the hyphae may appear nonpigmented (hyaline). Common species causing disease in humans are *C. geniculata*, *C. lunata*, *C. pallescens*, *C. senegalensis*, and *C. verruculosa*. In culture, *Curvularia* colonies are rapidly growing, woolly, and gray to grayish black or brown. Microscopically, the conidia are dematiaceous, two to several celled, and curved (Fig. 52.27). The conidiophores are dematiaceous, solitary or in groups, septate, simple or branched, sympodial, and geniculate (31,32,51).

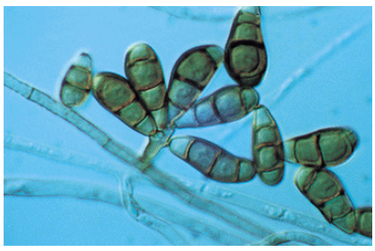


FIGURE 52.27. Microscopic appearance of *Curvularia* species demonstrates conidia and geniculate conidiophore (Nomarski optics, original magnification ×1,250).

Infections caused by the genera *Bipolaris* and *Exserohilum* present similarly to those owing to *Aspergillus* species, except that disease progresses more slowly. Clinical presentations include dissemination with vascular invasion and tissue necrosis, involvement of the central nervous system and paranasal sinuses, and association with allergic bronchopulmonary disease. Most commonly, these organisms cause sinusitis in “normal” (atopic/asthmatic) hosts and more invasive disease in immunocompromised hosts. Marijuana smoking may be a risk factor for infection with these organisms. In culture, *Bipolaris* and *Exserohilum* both form rapidly growing, woolly, gray to black colonies. Microscopically, the conidiophores are sympodial and geniculate. The conidia are dematiaceous, oblong to cylindrical, and multicelled.

Pneumocystosis

Pneumocystis carinii is an organism that causes infection almost exclusively in debilitated and immunosuppressed patients, especially those with HIV infection. It is the most common opportunistic infection among the AIDS population; however, the incidence has decreased in recent years because of the prophylactic use of trimethoprim/sulfamethoxazole and other antimicrobial agents including Highly Active Anti-Retronal Therapy (HAART). Although previously considered to be a protozoan parasite, recent molecular and genetic evidence suggests that *P. carinii* is closer phylogenetically to the fungi (99,100,101,102,103 and 104). Ultrastructurally, both fungi and *P. carinii* have similar cell-wall and mitochondrial structure as well as the presence of intracytoplasmic bodies resembling ascospores (102,105,106). There is homology of the conserved domains of the 16S rRNA with that of the Ascomycetes (99,104,105), the 5S rRNA with the primitive Zygomycetes (103), and the ß-tubulin gene with that of the filamentous fungi (58,107). In addition, the protein synthesis elongation factor EF-3 and thymidylate synthetase are similar to those of the Ascomycetes (99,100). Conversely, there are several factors favoring a protozoan origin including the different morphologic forms found in human tissue and the presence of cholesterol and other unique sterols, instead of ergosterol, in the cell membrane (103,108).

The life cycle of *P. carinii* may include both sexual and asexual components (102,105). During the course of human infection, *P. carinii* may exist as free trophic forms (1.5 to 5 µm in diameter), as a uninucleate sporocyst (4.0 to 5.0 µm), or as a cyst (5 µm) containing as many as eight ovoid to fusiform intracystic bodies (Fig. 52.28) (108,109). After rupture of the cyst, the cyst wall may be seen as an empty collapsed structure (Fig. 52.29).

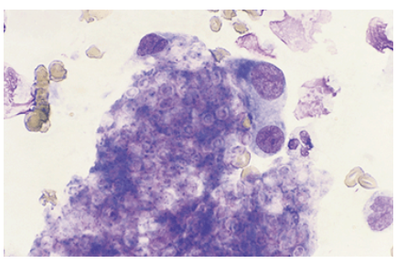


FIGURE 52.28. Microscopic appearance of *Pneumocystis carinii* stained with Giemsa stain (original magnification ×1,000 oil).

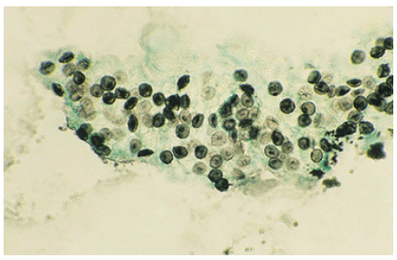


FIGURE 52.29. Microscopic appearance of *Pneumocystis carinii* stained with Gomori methenamine silver stain (original magnification ×1,000 oil).

The reservoir for *P. carinii* in nature is unknown. Although airborne transmission has been documented experimentally among rodents (110,111), it is unlikely that rodents serve as a reservoir for human disease given the fact that rodent strains of *P. carinii* are genetically distinct from those of humans (112,113).

The respiratory tract appears to be the main portal of entry for *P. carinii*. Pneumonia is clearly the most common clinical presentation of *P. carinii* infection; however, among AIDS patients, extrapulmonary sites of infection including lymph nodes, spleen, bone marrow, liver, small intestine, genitourinary tract, heart, eyes, ears, skin, bone, and thyroid have been reported (108,114,115). The high rate of seropositivity among young

children has led to the assumption that most infections were secondary to reactivation of latent infection; however, recent evidence suggests that both reactivation and primary infection occur (116,117,118,119,120,121 and 122). Malnourished, debilitated, and immuno-suppressed patients, especially AIDS patients with low CD4 counts (less than 200/µL), are at high risk.

The hallmark of *P. carinii* infection is an interstitial pneumonitis with a mononuclear infiltrate composed predominantly of plasma cells. The onset of disease is insidious with signs and symptoms including dyspnea, cyanosis, tachypnea, nonproductive cough, and fever. The radiographic appearance is typically one of diffuse interstitial infiltrates with a ground-glass appearance extending from the hilar region but may appear normal or show nodules or cavitation. The mortality rate is high among untreated patients and death is owing to asphyxia.

Histologically, a foamy exudate is seen within the alveolar spaces with an intense interstitial infiltrate composed predominantly of plasma cells. Other patterns including diffuse alveolar damage, noncaseating granulomatous inflammation, and infarctlike coagulative necrosis may also be seen (109).

The diagnosis of *P. carinii* infection is almost entirely based on microscopic examination of clinical material including bronchoalveolar lavage fluid, bronchial brushings, induced sputum, and transbronchial or open lung biopsy specimens. Examination of bronchoalveolar lavage fluid has been shown to have a sensitivity of 90% to 100% and in most instances alleviates the need for transbronchial or open lung biopsy (123,124). Although examination of induced sputum may be useful in some AIDS patients with a high organism load, it has a 20% to 25% false negative rate (123,124). Several stains have been used to detect *P. carinii* including Giemsa, Giemsa-like, toluidine blue, Gomori's methenamine silver, periodic acid–Schiff, Gram-Weigert, calcofluor, and immunofluorescence (108). Giemsa and Giemsa-like (e.g., Wright's) stains demonstrate the trophozoites and do not stain the cell wall (Fig. 52.28). Immunofluorescent techniques stain both trophic forms and the cyst walls (108). *P. carinii* cysts are often stained with Gomori's methenamine silver and demonstrate intact cysts, helmet-shaped forms and cuplike structures as well as the presence of two focal thickenings resembling commas in the cyst wall that are apparently unique to *Pneumocystis* (Fig. 52.29). Generally, a combination of stains is used in most laboratories.

Other diagnostic modalities, including serologic testing, have not proven useful in the diagnosis of *P. carinii* infection (108,125). Although polymerase chain reaction (PCR)-based methods are more sensitive than microscopy in detection of *P. carinii* (126,127,128 and 129), they are generally less specific and more labor intensive (129,130).

ANTIGENIC, BIOCHEMICAL, AND MOLECULAR MARKERS FOR DIAGNOSIS OF INVASIVE FUNGAL INFECTION

Part of "52 - Fungi and Fungal Infections"

The need for more rapid, sensitive and specific nonculture methods for diagnosis of fungal infections is acute. Tests for detection of antibodies, fungal antigens, metabolites, and fungus-specific nucleic acid sequences all have the potential to yield rapid diagnostic information that can guide the early and appropriate use of antifungal therapy (65,76,97,131,132,133 and 134). Some progress has been made in these areas (Table 52.12); however, with few exceptions, these tests have yet to make a significant impact on the diagnosis of fungal infections.

TABLE 52.12. ANTIGENIC, BIOCHEMICAL, AND MOLECULAR MARKERS FOR DIRECT DETECTION OF INVASIVE FUNGAL INFECTIONS

Organism	Cell Wall or Capsule Components	Cytoplasmic Antigens	Metabolites	Genomic DNA Sequences[a]
Candida spp	1. Mannans a. RIA b. ELISA 2. 1,3-B-glucans a. Amebocyte lysate assay 3. Chitin synthase a. Spectrophotometric	1. Enolase a. ELISA b. Immunoblot 2. Antienolase antibody a. ELISA 3. 47-kD breakdown product of HSP-90 a. Enzyme-linked dot immunobinding assay	1. D-Arabinitol a. Rapid enzymatic b. GLC/FID c. Mass spectroscopy/GLC	1. C-14-lanosterol demethylase 2. Ribosomal RNA genes 3. Chitin synthase
Cryptococcus neoformans	1. Capsule-glucuronoxylomannan a. Latex agglutination b. GLC		1. D-Mannitol a. Mass spectroscopy b. ELISA	1. Ribosomal RNA genes
Trichosporon asahi (beigelli)	1. Cell wall a. Cross reactivity in assays for cryptococcal glucuron-oxylomannan			1. Ribosomal RNA genes
Blastomyces dermatitidis	1. Cell wall a. RIA for 120-kD cell wall adhesion protein			1. Ribosomal RNA genes
Histoplasma capsulatum	1. Cell wall a. RIA and ELISA for polysaccharide antigen			1. Ribosomal RNA genes
Aspergillus spp.	1. Galactomannan a. Latex agglutination b. ELISA c. RIA 2. 1,3-B-glucans a. Amebocyte lysate assay 3. Chitin a. Spectrophotometry		1. D-Mannitol a. GLC/FID b. Mass spectroscopy/GLC	1. C-14 lanosterol demethylase 2. Ribosomal RNA genes 3. Alkaline protease

[a] All sequences detected by polymerase chain reaction.
RIA, radioimmunoassay; ELISA, enzyme-linked immunosorbent assay; GLC, gas-liquid chromatography; FID, flame ionization detector.
Adapted from Walsh TJ, Chanock SJ. Diagnosis of invasive fungal infections: advances in nonculture systems. *Curr Clin Topics Infect Dis* 1998;18:101–153.

Determination of antibody and/or antigen titers may be useful in diagnosing fungal infections and when performed in a serial fashion also provide a means of monitoring the progression of disease and the patient's response to therapy (62,66,67,76,94,96,97,135,136). With the exception of antibody tests for histoplasmosis and coccidioidomycosis, however, most tests designed to detect an antibody response as a means of diagnosing invasive fungal infections lack sensitivity and specificity, are poorly standardized and are not widely available (62). Antibody tests for *Candida* and *Aspergillus* are often unable to distinguish active from past infection and colonization from transient fungemia. In addition, a negative test for fungus-specific antibodies does not rule out infection because immunocompromised patients, and some individuals with disseminated infection may not mount an antibody response to the infecting organism (62).

Detection of fungal antigens or metabolites in serum or other body fluids represents the most direct means of providing a serologic diagnosis of invasive fungal infection (42,43,63,66,67,94,96,97). Although significant advances have been made in recent years (Table 52.12), most methods for rapid detection of fungal antigens are available only in research laboratories (97). Notable exceptions are the tests for detection of the polysaccharide antigens of *C. neoformans* and *H. capsulatum* (63,66,67,94,97). The commercially available latex agglutination and EIA tests for cryptococcal antigen are well standardized and widely available and detect more than 95% of cryptococcal meningitis and approximately 67% of disseminated cryptococcal infections (94,97). Both the latex agglutination and EIA tests have supplanted the India ink test (sensitivity of less than 40%) for the diagnosis of cryptococcal meningitis (94). The test for *Histoplasma* antigen is available from a reference laboratory (Histoplasmosis Reference Laboratory, Indianapolis, IN) and is widely used for the diagnosis of disseminated histoplasmosis (63,66,67,97). The test employs either a RIA or EIA format and detects *H. capsulatum* polysaccharide antigen in urine and serum (63). Both the

RIA and the EIA tests have been shown to be rapid (less than 24 hours), sensitive (55% to 99%), specific (more than 98%), and reproducible (63,66,67). In addition to urine and serum, the antigen may be detected in the spinal fluid of 42% to 67% of patients with *Histoplasma* meningitis and in bronchoalveolar lavage fluid of 70% of patients with AIDS and severe pulmonary histoplasmosis (64,66,67,97).

In contrast to the success achieved with the antigen test for cryptococcosis and histoplasmosis, immunoassays for antigens of *Candida* and *Aspergillus* have been disappointing (42,43,76,96,97,133). Efforts to develop tests to detect *Candida* mannan and enolase antigens have been promising but have failed to produce a commercially available product because of problems in achieving optimal sensitivity and specificity (76,97,132,135,136). Likewise, both EIA and latex agglutination assays for detecting *Aspergillus* galactomannan in serum have shown promise but are not widely available (42,43,96,97). Further research is needed in this area to determine both the diagnostically relevant antigens and the testing format necessary to establish a clinically useful assay.

The detection of fungal metabolites is a potential method for the rapid diagnosis of both candidiasis and aspergillosis (41,65,97). The detection of arabitol in serum appears to be an indicator of hematogenously disseminated candidiasis (65,76,97). The diagnostic specificity of arabitol detection may be improved by correcting for renal function (arabitol:creatinine ratio) or by detection of specific isomers (D-arabitol) (76,97). Similarly, the detection of elevated levels of D-mannitol in bronchoalveolar lavage fluid may be useful in the diagnosis of pulmonary aspergillosis (97). The diagnostic utility of metabolite detection remains uncertain owing to significant method-dependent variation in the sensitivity and specificity (97,133).

The use of the PCR to detect fungal nucleic acids directly in clinical material offers great promise for the rapid diagnosis of fungal infections (43,65,97,131,134,137,138,139 and 140). At present, most of the research has concentrated on the diagnosis of invasive

candidiasis (65,76,97); however, PCR has also been applied to the diagnosis of aspergillosis and other fungal infections (Table 52.12) (43,131,134,139).

The amplification and detection of *Candida*-specific DNA from blood and other body fluids are now well established in the research setting (76,137,138) but have yet to gain wide application in the clinical laboratory (65). Amplification targets include the lanosterol demethylase gene, mitochondrial DNA, ribosomal RNA sequences, the actin gene, and a chitin synthetase gene to name a few (65,76,131,133). Detection of as few as two to 10 cells per milliliter has been reported, although most assays do not approach this level of sensitivity in clinical samples. The true sensitivity of PCR for diagnosis of invasive candidiasis is unknown, but sensitivities of 76% to 96% have been reported (65,76,131,133). PCR for detection of *Aspergillus* and other fungal pathogens also shows great promise and the number of studies documenting its utility in this regard is growing steadily (43,97,131,134,139). Current PCR-based approaches to the diagnosis of fungal infections are cumbersome, yet promising prototypes. The distinction between PCR results involving normal flora, colonization, contamination, and truly invasive isolates must be resolved.

In addition to detection of fungal pathogens in clinical material, immunologic and molecular approaches have been applied successfully to the identification of fungi in culture (47,54,59,60 and 61,97,140). Nucleic acid probe–based methods offer great promise for the identification of both common and uncommon clinical isolates (140). Exoantigen and nucleic acid probe–based methods are now well established in the clinical laboratory for the identification of the dimorphic pathogens *H. capsulatum*, *B. dermatitidis*, and *C. immitis* (54,59,60 and 61). In the commercially available nucleic acid probe tests (GenProbe, San Diego, CA), nucleic acids are extracted from culture material and reacted with a chemiluminescent-labeled probe specific for *H. capsulatum*, *B. dermatitidis*, *C. immitis*, or *C. neoformans* rRNA (54,60,61). These tests may be completed within 2 hours and have an accuracy of 99% to 100%.

TABLE 52.13. ANTIFUNGAL AGENTS WITH SYSTEMIC ACTIVITY

Drug	Route	Mechanism	Comments
Amphotericin B	IV	Binds to ergosterol causing direct membrane damage	Established agent; broad spectrum; toxic
Lipid formulations (amphotericin B lipid complex or colloidal dispersion, liposomal formulation)	IV	Same as amphotericin B	Recently licensed agents with broad spectrum activity and decreased toxicity; used for refractory severe infections not responding to conventional therapy
Miconazole	IV	Inhibits membrane sterol synthesis	Toxic agent with modest activity against *Candida* and other yeasts; active against *Pseudallescheria boydii*
Ketoconazole	Oral, topical	Same as miconazole	Established agent with modest broad-spectrum activity
Fluconazole	Oral, IV, topical	Same as miconazole	Triazole with broad-spectrum activity; good central nervous system penetration; good *in vivo* activity against *Candida* and *Cryptococcus neoformans*
Itraconazole	Oral, IV, topical	Same as miconazole	Triazole with broad-spectrum activity including *Aspergillus;* excellent *in vivo* activity against endemic mycoses
Voriconazole, Posaconazole, Ravuconazole	Oral, IV, topical	Same as miconazole	Investigational agents with broad-spectrum activity against species resistant to established agents; used for treatment of aspergillosis
Caspofungin, Anidulafungin	IV	Inhibition of fungal cell wall glucan synthase	Investigational agent with broad-spectrum *in vitro* activity, except against *C. neoformans* and *T. asahi (beigelii)*; used for treatment of aspergillosis
5-Fluorocytosine	Oral	Inhibition of DNA and RNA synthesis	Toxicity and resistance are problems; used in combination with amphotericin B

IV, intravenous.

THERAPY

Part of "52 - Fungi and Fungal Infections"

The emergence of fungal pathogens as important agents of human infection has resulted in an increase in the use of systemic antifungal agents worldwide and the introduction of a number of new antifungal agents with systemic activity (69,74,98,141,142,143,144,145,146,147,148,149,150 and 151). Several antifungal agents are now available for the treatment of fungal infections. Fourteen of these compounds are discussed briefly in this chapter (Table 52.13). They include representatives of the polyene macrolide class (amphotericin B, amphotericin B lipid complex, liposomal amphotericin B, and amphotericin B cholesterol dispersion), the azoles (miconazole, ketoconazole, fluconazole, itraconazole, voriconazole, posaconazole, and ravuconazole), glucan synthesis inhibitors (caspofungin and anidulafungin), and an inhibitor of pyrimidine synthesis (5-fluorocytosine). In addition, the rationale for and limitations of *in vitro* susceptibility testing are discussed with a description

of the current efforts to develop a standardized *in vitro* testing method and to establish clinical correlations for *in vitro* test results.

Antifungal Agents

Amphotericin B and the liposomal and lipid complex formulations of amphotericin B are polyene macrolide antibiotics used primarily in the treatment of systemic and life-threatening fungal infections. These compounds act by binding to ergosterol in the fungal cell membrane, causing loss of membrane integrity and osmotic instability (74). Amphotericin B is the drug of choice for the treatment of aspergillosis, disseminated candidiasis, cryptococcosis, and zygomycosis (69,74,98,142). It is also effective in the treatment of coccidioidomycosis, histoplasmosis, and blastomycosis (Table 52.14). Unfortunately, the therapeutic efficacy of amphotericin B is limited by significant toxicity, specifically nephrotoxicity. The liposomal and lipid complex formulations of amphotericin B were designed to increase the therapeutic index (increase efficacy and decrease toxicity) for amphotericin B therapy and to maximize the delivery of amphotericin B to patients with deep-seated fungal infections such as hepatosplenic candidiasis and invasive pulmonary aspergillosis (74,98). Three different formulations have been evaluated in clinical trials and have been approved for treatment of invasive fungal infection in patients for whom conventional therapy has failed: amphotericin B lipid complex, a 1:1 molar complex of phospholipids and drug; ambisome, a phospholipid/cholesterol liposome formation; and amphocil, a 1:1 molar complex of cholesterol sulfate and amphotericin B. The resultant liposomal and lipid complex preparations have selective toxicity for fungal cells but not for mammalian cells and theoretically promote delivery of the drug to the site of infection, while avoiding the toxicity of supramaximal doses of amphotericin B. The available clinical data indicate that these formulations allow the administration of higher doses of amphotericin B with no increase in acute toxicity and no chronic toxicity (74,98). Importantly, there is no evidence of dose-limiting nephrotoxicity. Unfortunately, the therapeutic efficacy of these formulations does not appear to be significantly better than that of conventional amphotericin B.

TABLE 52.14. ANTIFUNGAL THERAPY FOR DEEP MYCOSES

	Agents	
Infection	**First Choice**	**Alternatives[a]**
Aspergillosis	Amphotericin B (± 5-FC)	Itraconazole
	Lipid formulations of amphotericin B	Caspofungin
Blastomycosis	Amphotericin B	Ketoconazole[b]
	itraconazole[b]	
Candidiasis	Amphotericin B (± 5-FC) fluconazole	Caspofungin
Chromoblastomycosis	Itraconazole	5-FC
Cocciodioidomycosis	Amphotericin B	Ketoconazole[b]
	Fluconazole	Itraconazole[b]
		Intrathecal amphotericin B[c]
Cryptococcosis	Amphotericin B (± 5-FC)	Fluconazole
Fusariosis	Amphotericin B	Lipid formulation of amphotericin B
Histoplasmosis	Amphotericin B	Ketoconazole[d]
	Itraconazole[d]	Fluconazole[d]
Phaeohyphomycosis	Amphotericin B	Itraconazole
Pseudallescherosis	Itraconazole	Miconazole, ketoconazole, or voriconazole
Sporotrichosis	Oral iodides[e]	
	Itraconazole[e]	
Trichosporosis	Amphotericin B[f]	Fluconazole
Zygomycosis	Amphotercin B	None established

[a] Limited data exist for most alternative agents.

[b] Indolent, nonmeningeal.

[c] Meningeal.

[d] Chronic pulmonary or indolent, nonmeningeal disseminated infection.

[e] Lymphocutaneous.

[f] Extracutaneous.

5-FC, 5-fluorocytosine.

The azole class of antifungal compounds consists of a large number of agents with systemic antifungal activity, including the imidazoles (miconazole and ketoconazole) and the newer triazoles (itraconazole, fluconazole, voriconazole, posaconazole, and ravuconazole) (Table 52.13) (74,98,142,145,147,148,150). The azole antifungals interact with cytochrome P-450–dependent enzymes with resulting impairment of ergosterol synthesis and depletion of ergosterol in the fungal cell membrane (150). This results in inhibition of fungal cell growth. The imidazoles miconazole and ketoconazole are established broad-spectrum agents active against a variety of fungal pathogens, including yeasts, dimorphic organisms, dermatophytes, and opportunistic pathogens (Table 52.13 and Table 52.14). The triazoles all show promise as broad-spectrum, orally active, systemic agents with less potential for toxicity than the currently available imidazoles (74,98,142,145,147,148,150). Both fluconazole and itraconazole currently have important roles as

agents of first choice or as major alternatives in the treatment of deep mycoses (Table 52.14) (74,98,152,150). Ongoing and future clinical trials will more clearly define the specific roles of the triazoles in the treatment of systemic mycoses (74,98,142,145,147,148,152,150).

Another target of antifungal activity has been the inhibition of glucan synthesis in the fungal cell wall. This approach takes advantage of the fact that human cells lack the enzyme glucan synthase and thus increases the selectivity for fungal cells. The echinocandins caspofungin and anidulafungin are inhibitors of 1,3-β-D-glucan synthesis, which leads to the lysis of fungal cells. Caspofungin and anidulafungin both have demonstrated potent *in vitro* and *in vivo* activities against a variety of yeasts (including fluconazole- and itraconazole-resistant strains) and molds (141,143,146,151). Notably, both agents lack activity against *C. neoformans* and *T. beigelii* (144).

5-Fluorocytosine is a water-soluble fluorinated pyrimidine antimetabolite used orally (intravenous formulation is investigational) in the treatment of systemic infections caused by susceptible pathogenic or opportunistic yeasts and fungi. The spectrum of susceptible fungal pathogens is narrow and includes *Candida* species, *C. neoformans*, and some agents of chromoblastomycosis (Table 52.14). 5-Fluorocytosine acts as a competitive antimetabolite for uracil in the synthesis of fungal RNA and also inactivates thymidylate synthetase (98). These activities can be antagonized by a variety of purine and pyrimidine bases and nucleosides.

Antifungal Susceptibility Testing

The field of antifungal susceptibility testing has progressed considerably over the past 10 to 15 years. *In vitro* susceptibility tests with antifungal agents are performed for the same reasons that tests with antibacterial agents are performed (98,153,154,155,156,157,158 and 159). Ideally, *in vitro* susceptibility tests will (a) provide a reliable estimate of the relative activities of two or more antifungal agents, (b) correlate with *in vivo* antifungal activity and predict the likely outcome of therapy, (c) provide a means with which to monitor the development of resistance among a normally susceptible population of organisms, and (d) predict the therapeutic potential of newly discovered investigational agents.

At present, the state of the art for susceptibility testing of yeasts is comparable with that of bacteria (158). Standardized methods for performing antifungal susceptibility testing are reproducible, accurate, and available for use in clinical laboratories. The development of quality control guidelines and interpretive criteria for a limited number of antifungal agents (Table 52.15) provides a basis for application of this testing in the clinical laboratory (Table 52.16) (158,159).

TABLE 52.15. INTERPRETIVE GUIDELINES FOR *IN VITRO* SUSCEPTIBILITY TESTING OF YEAST ISOLATES TO FLUCONAZOLE, ITRACONAZOLE, AND 5-FLUOROCYTOSINE USING NATIONAL COMMITTEE FOR CLINICAL LABORATORY STANDARDS APPROVED METHODS[a]

Antifungal Agent	Interpretive Category	MIC (μg/mL) Interpretive Breakpoint	Clinical Outcome (% Success)[b]
Fluconazole	Susceptible	≤8.0	97
	Susceptible-dose dependent	16–32	82
	Resistant	≥64	60
Itraconazole	Susceptible	≤0.12	90
	Susceptible-dose dependent	0.25–0.5	63
	Resistant	≥1.0	53
5-FC	Susceptible	≤4.0	NA[c]
	Intermediate	8.0–16	NA
	Resistant	≥32	NA

[a] Macrodilution and microdilution broth methods as described by Pfaller et al. (158), Rex et al. (159), and NCCLS (155).

[b] Clinical outcome for treatment of candidemia (fluconazole only) and oropharyngeal candidiasis in patients with acquired immunodeficiency syndrome (both fluconazole and itraconazole) as described by Rex et al. (159)

[c] Clinical data not available.

5-FC, 5-fluorocytosine.

Although establishing a correlation between *in vitro* susceptibility tests and clinical outcome has been difficult, it is now clear that antifungal susceptibility testing can predict outcome in several clinical situations, the most notable of which is fluconazole and itraconazole treatment of oropharyngeal candidiasis in AIDS patients (159). The establishment of interpretive breakpoints for *in vitro* susceptibility of yeasts to these antifungal agents has now been accomplished (Table 52.15). The National Committee for Clinical Laboratory Standards Subcommittee on Antifungal Susceptibility Tests has also developed and published a proposed reference method for broth dilution testing of filamentous fungi that cause invasive infections (98,153,156). Studies are now ongoing to further refine this approach and to evaluate the *in vivo* correlation with the *in vitro* data for molds (98,160).

Despite this progress, it remains to be seen how useful antifungal susceptibility testing will be in guiding therapeutic decision making. Guidelines for the use of susceptibility testing in the laboratory have been developed (Table 52.16) and include routine identification of fungi to species level and only selective application of *in vitro* susceptibility testing (158). Future efforts must be directed toward establishing and validating interpretive breakpoints for currently available antifungals as well as those new antifungals under development. In addition, procedures must be optimized for testing non-*Candida* yeasts (e.g., *C. neoformans*) and molds (93,98).

SUMMARY AND CONCLUSIONS

Our understanding of the epidemiology and pathogenesis of fungal infections has improved considerably over the past decade. By weight of the sheer numbers of cases, we have the best understanding of candidal infections; however, there is little doubt that newer fungi previously considered nonpathogens have now emerged as significant human pathogens, particularly in the immunocompromised host. Recognition of these emerging fungal pathogens has resulted in improved understanding of their clinical presentation and response to available therapeutic measures. Clearly, additional efforts are necessary to improve diagnosis and management and to understand better the pathogenesis of infection owing to this diverse group of fungal pathogens. Future studies should include clinical trials of prophylactic agents to prevent colonization, of empiric agents for early treatment of suspected infections, and of therapeutic agents to treat documented infections. Advances in diagnostic methods may provide earlier diagnosis and thus encourage earlier and potentially more effective therapeutic interventions. Finally, epidemiologic studies documenting the rates of colonization and infection owing to these emerging fungal pathogens, as well as studies of the effectiveness of various components of infection control and antibacterial drug use, and optimal guidelines for insertion and management of vascular catheters are urgently needed.

TABLE 52.16. RECOMMENDATIONS FOR STUDIES OF FUNGAL ISOLATES IN THE CLINICAL LABORATORY

Recommendation	Clinical setting
• Species level identification of all *Candida* isolates from deep sites	Routine
• Genus level identification of molds	
• Routine antifungal susceptibility testing not recommended	
• Periodic batch susceptibility testing of *Candida* spp	Epidemiologic survey
• Establish antibiogram for an institution	
• Relevant drugs: fluconazole, itraconazole, 5-FC	
• Routine antifungal susceptibility testing not required	AIDS and oropharyngeal candidiasis
• Susceptibility testing may be useful for patients unresponsive to azole therapy	
• Relevant drugs: fluconazole, itraconazole	
• Isolates of *Candida* (especially non-*albicans*) from deep sites	Invasive candidiasis
• Susceptibility testing offered on request for selected patients	
• Relevant drugs: fluconazole, itraconazole, 5-FC	
• Testing not recommended	Cryptococcosis
• Interpretive criteria have not been established	
• Testing not recommended	Mold infections
• Interpretive criteria have not been established	

5-FC, 5-fluorocytosine; AIDS, acquired immunodeficiency syndrome.
Adapted from Pfaller MA, Rex JH, Rinaldi MG. Antifungal susceptibility testing: technical advances and potential clinical applications. *Clin Infect Dis* 1997; 24:776–784.

REFERENCES

1. Abi-Said D, Anaissie E, Uzun O, et al. The epidemiology of hematogenous candidiasis caused by different *Candida* species. *Clin Infect Dis* 1997;24:1122–1128.

2. Aucott JN, Fayer J, Grossnicklas H, et al. Invasive infections with *Saccharomyces cerevisiae*: report of three cases and review. *Rev Infect Dis* 1990;12:406–411.

3. Banerjee SN, Emori TG, Culver DH, et al. Secular trends in nosocomial primary blood stream infections in the United States, 1980–1989. *Am J Med* 1991;91(Suppl 3B):86S–89S.

4. Beck-Sague CM, Jarvis WR, The National Nosocomial Infections Surveillance System. Secular trends in the epidemiology of nosocomial fungal infections in the United States, 1980 to 1990. *J Infect Dis* 1991;167:1247–1251.

5. Berrouane YF, Hollis RJ, Pfaller MA. Strain variation among and antifungal susceptibilities of isolates of *Candida krusei*. *J Clin Microbiol* 1996;34:1856–1858.

6. Berrouane YF, Herwaldt LA, Pfaller MA. Trends in antifungal use and epidemiology of nosocomial yeast infections in a University hospital. *J Clin Microbiol* 1999;37:531–537.

7. Borg-Von Zepelin M, Eiffert H, Kann M, et al. Changes in the spectrum of fungal isolates: results from clinical specimens gathered in 1987/88 compared with those in 1991/92 in the University Hospital Gottingen, Germany. *Mycoses* 1993;36:247–253.

8. Brunn B, Westh H, Stenderup J. Fungemia: an increasing problem in a Danish University Hospital 1989 to 1994. *Clin Microbiol Infect* 1995;1:124–126.

9. Cooper CR, McGinnis MR. Pathology of *Penicillium marneffei*: a report of eight cases and differentiation of the causative or-ganism from *Histoplasma capsulatum*. *Am J Clin Pathol* 1985;84:323–327.

10. Emory T, Gaynes R. An overview of nosocomial infections, including the role of the microbiology laboratory. *Clin Microbiol Rev* 1993;6:428–442.

11. Fridkin SK, Jarvis WR. Epidemiology of nosocomial fungal infections. *Clin Microbiol Rev* 1996;9:499–511.

12. Minor RL Jr, Pfaller MA, Ginrich RD, et al. Disseminated *Fusarium* infections in patients following bone marrow transplantation. *Bone Marrow Transplant* 1989;4:653–658.

13. Morrison VA, Haake RJ, Weisdorf DJ. The spectrum of non-*Candida* fungal infections following bone marrow transplantation. *Medicine* 1993;72:78–89.

14. Perfect JR, Schell WA. The new fungal opportunists are coming. *Clin Infect Dis* 1996;22(Suppl 2):S112–S118.

15. Pfaller MA. Nosocomial candidiasis: emerging species, reservoirs, and modes of transmission. *Clin Infect Dis* 1996;22(Suppl 2):S89–S94.

16. Pfaller MA. Opportunistic fungal infections. In: Wallace RB, Doebbeling BN, eds. *Maxcy-Rosenthal-Last public health and preventive medicine*, 14th ed. Stamford, CT: Appleton & Lange, 1998:387–392.

17. Pfaller MA. The epidemiology of invasive mycoses: narrowing the gap. *Clin Infect Dis* 1998;27:1148–1150.

18. Pfaller M, Wenzel R. Impact of the changing epidemiology of fungal infections in the 1990's. *Eur J Clin Microbiol Infect Dis* 1992;11:287–291.

19. Pfaller MA, Jones RN, Messer SA, et al. National surveillance of nosocomial blood stream infection due to *Candida albicans*: frequency of occurrence and antifungal susceptibility in the SCOPE Program. *Diagn Microbiol Infect Dis* 1998;31:327–332.

20. Pfaller MA, Jones RN, Messer SA, et al. National surveillance of nosocomial blood stream infection due to species of *Candida* other than *Candida albicans*: frequency of occurrence and antifungal susceptibility in the SCOPE Program. *Diagn Microbiol Infect Dis* 1998;30:121–129.

21. Pfaller MA, Jones RN, Doern GV, et al. International surveillance of blood stream infections due to *Candida* species: frequency of occurrence and antifungal susceptibilities of isolates collected in 1997 in the United States, Canada, and South America for the SENTRY Program. *J Clin Microbiol* 1998;36:1886–1889.

22. Pfaller MA, Messer SA, Houston A, et al. National Epidemiology of Mycoses Survey: a multicenter study of strain variation and antifungal susceptibility among isolates of *Candida* species. *Diagn Microbiol Infect Dis* 1998;31:289–296.

23. Pfaller MA, Lockhart SR, Pujol C, et al. Hospital specificity, region specificity, and fluconazole resistance of *Candida albicans* blood stream isolates. *J Clin Microbiol* 1998;36:1518–1529.

24. Pfaller MA, Messer SA, Hollis RJ, et al. Trends in species distribution and susceptibility to fluconazole among blood stream isolates of *Candida* species in the United States. *Diagn Microbiol Infect Dis* 1999;33:217–222.

25. Rees JR, Pinner RW, Hajjeh RA, et al. The epidemiological features of invasive mycotic infections in the San Francisco Bay area, 1992–1993: results of a population-based laboratory active surveillance. *Clin Infect Dis* 1998;27:1138–1147.

26. Sandven P, Bevanger L, Digranes A, et al. Constant low rate of fungemia in Norway, 1991 to 1996. *J Clin Microbiol* 1998;36:3455–3459.

27. Schell WA. New aspects of emerging fungal pathogens: a multifaceted challenge. *Clin Lab Med* 1995;15:365–386.

28. Voss A, Kluytmans JJAW, Koeleman JGM, et al. Occurrence of yeast blood stream infections between 1987 and 1995 in five Dutch University Hospitals. *Eur J Clin Microbiol Infect Dis* 1996;15:909–912.

29. Wey SB, Mori M, Pfaller MA, et al. Hospital-acquired candidemia: the attributable mortality and excess length of stay. *Arch Intern Med* 1988;148:2642–2645.

30. Wey SB, Mori M, Pfaller MA, et al. Risk factors for hospital-acquired candidemia. *Arch Intern Med* 1989;149:2349–2353.

31. Kwon-Chung KJ, Bennett JE. *Medical mycology*. Philadelphia: Lea & Febiger, 1992.

32. Rippon JW. *Medical mycology: the pathogenic fungi and the pathogenic Actinomycetes*, 3rd ed. Philadelphia: Saunders, 1988.

33. Chariyalertsak S, Sirisanthana T, Supparatpinyo K, et al. Seasonal variation of disseminated *Penicillium marneffei* infections in Northern Thailand: a clue to the reservoir? *J Infect Dis* 1996;173:1490–1493.

34. Drouhet E. Penicillosis due to *Penicillium marneffei*: a new emerging systemic mycoses in AIDS patients travelling or living in Southeast Asia: review of 44 cases reported in HIV infected patients during the last 5 years compared to 44 cases of non-AIDS patients reported over 20 years. *J Mycol Med* 1993;4:195–224.

35. Duong TA. Infection due to *Penicillium marneffei*, an emerging pathogen: review of 155 reported cases. *Clin Infect Dis* 1996;23:125–130.

36. Li JS, Pan LQ, Wu SX, et al. Disseminated *Penicillium marneffei* in China: report of three cases. *China Med J* 1991;104:247–251.

37. Wortman PD. Infection with *Penicillium marneffei*. *Int J Dermatol* 1996;35:393–399.

38. Yuen WC, Chan YF, Loke SL, et al. Chronic lymphadenopathy caused by *Penicillium marneffei*: a condition mimicking tuberculous lymphadenopathy. *Br J Surg* 1986;73:1007–1008.

39. Nguyen MH, Peacock JE, Morris AJ, et al. The changing face of candidemia: emergence of non-*Candida albicans* species and antifungal resistance. *Am J Med* 1996;100:617–623.

40. Pannuti CS, Gingrich RD, Pfaller MA, et al. Nosocomial pneumonia in adult patients undergoing bone marrow transplantation: a 9-year study. *J Clin Microbiol* 1991;9:1–5.

41. Sigler L, Kennedy MJ. *Aspergillus*, *Fusarium*, and other opportunistic moniliaceous fungi. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1212–1241.

42. Verweij PE, Stynen D, Rijs AJ, et al. Sandwich enzyme-linked immunosorbent assay compared with Pastorex latex agglutination test for diagnosing invasive aspergillosis in immunocompromised patients. *J Clin Microbiol* 1995; 33:1912-1914.

43. Verweij PE, Latge JP, Rijs AJ, et al. Comparison of antigen detec-tion and PCR assay using bronchoalveolar lavage fluid for diag-nosing invasive pulmonary aspergillosis in patients receiving treat-ment for hematological malignancies. *J Clin Microbiol* 1995;33:3150–3153.

44. Gamis A, Gudnason T, Giebink GS, et al. Disseminated infection with *Fusarium* in recipients of bone marrow transplants. *Rev Infect Dis* 1991;13:1077–1088.

45. Nelson PA, Toussoren TA, Marasas WFO. *Fusarium species*. University Park, PA: The Pennsylvania State University Press, 1983.

46. Nelson PE, Dignani MC, Anaissie EJ. Taxonomy, biology, and clinical aspects of *Fusarium* species. *Clin Microbiol Rev* 1994;7:479–504.

47. Kaufman L, Standard PG, Jalbert M, et al. Immunohistologic identification of *Aspergillus* spp. and other hyaline fungi by using polyclonal fluorescent antibodies. *J Clin Microbiol* 1997;35:2206–2209.

48. Hazen K. New and emerging yeast pathogens. *Clin Microbiol Rev* 1995;8:462–475.

49. Marcon MJ, Powell DA. Human infections due to *Malassezia* spp. *Clin Microbiol Rev* 1992;5:101–119.

50. Warren NG, Hazen KC. *Candida*, *Cryptococcus* and other yeasts of medical importance. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1184–1199.

51. Schell WA, Salkin IF, Pasarell L, et al. *Bipolaris*, *Exophiala*, *Scedosporium*, *Sporothrix*, and other demateaceous fungi. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1295–1317.

52. Kane J, Summerbell RC. *Trichophyton*, *Microsporum*, *Epidermophyton*, and agents of superficial mycoses. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1275–1294.

53. Rinaldi MG. Use of potato flakes agar in clinical mycology. *J Clin Microbiol* 1982;15:1159–1160.

54. Larone DH, Mitchell TG, Walsh TJ. *Histoplasma*, *Blastomyces*, *Coccidioides*, and other dimorphic fungi causing systemic mycoses. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1259–1274.

55. Padhye AA, McGinnis MR. Fungi causing eumycostic mycetoma. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1318–1326.

56. Richardson MD, Shankland GS. *Rhizopus*, *Rhizomucor*, *Absidia*, and other agents of systemic and subcutaneous zygomycosis. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:in press.

57. Einstein HE, Johnson RH. Coccidioidomycosis: new aspects of epidemiology and therapy. *Clin Infect Dis* 1993;16:349–354.

58. Dyer M, Volpe F, Delves CJ, et al. Cloning and sequence of a ß-tubulin cDNA from *Pneumocystis carinii*: possible implications for drug therapy. *Mol Microbiol* 1992;6:991–1001.

59. Kaufman L, Standard PG. Specific and rapid identification of medically important fungi by exoantigen detection. *Annu Rev Microbiol* 1987;41:209–225.

60. Padhye AA, Smith G, McLaughlin D, et al. Comparative evaluation of a chemiluminescent DNA probe and an exoantigen test for rapid identification of *Histoplasma capsulatum*. *J Clin Microbiol* 1992;30:3108–3111.

61. Stockman L, Clark KA, Hunt JM, et al. Evaluation of commercially available acridinium ester-labeled chemiluminescent DNA probes for culture identification of *Blastomyces dermatitidis*, *Coccidioides immitis*, *Cryptococcus neoformans*, and *Histoplasma capsulatum*. *J Clin Microbiol* 1993;31:845–850.

62. Kaufman L, Reiss E. Serodiagnosis of fungal diseases. In: Rose NR, DeMarcario EL, Fahey JL, et al., eds. *Manual of clinical laboratory immunology*, 4th ed. Washington DC: American Society for Microbiology, 1992:506–528.

63. Durkin MM, Connolly PA, Wheat LJ. Comparison of radioimmunoassay and enzyme-linked immunoassay methods for detection of *Histoplasma capsulatum* var. *capsulatum* antigen. *J Clin Microbiol* 1997;35:2252–2255.

64. Hamilton AJ. Serodiagnosis of histoplasmosis, paracoccidioidomycosis, and penicilliosis marneffei: current status and future trends. *Med Mycol* 1998;36:351–364.

65. Merz WG, Roberts GD. Algorithms for the direct detection of pathogenic fungi in clinical specimens. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1167–1183.

66. Wheat LJ, Connolly-Stringfield P, Kohler RB, Frame PT, Gupta MR. *Histoplasma capsulatum* polysaccharide antigen in diagnosis and management of disseminated histoplasmosis in patients with acquired immunodeficiency syndrome. *Am J Med* 1989; 87:396-400.

67. Wheat LJ, Connolly-Stringfield P, Blair R, Connolly K, Garringer T, Katz BP. Histoplasmosis relapse in patients with AIDS: detection using *Histoplasma capsulatum* variety *capsulatum* antigen levels. *Ann Int Med* 1991; 115:936-941.

68. Brummer E, Castaneda E, Restrepo A. Paracoccidioidomycosis: an update. *Clin Microbiol Rev* 1993;6:89–117.

69. Armstrong D. Treatment of opportunistic fungal infections. *Clin Infect Dis* 1993;16:1–9.

70. Coleman DC, Sullivan DJ, Bennett DE, et al. Candidiasis: the emergence of a novel species, *Candida dubliniensis*. *AIDS* 1997;11:557–567.

71. Fidel PL Jr, Vazquez JA, Sobel JD. *Candida glabrata*: review of epidemiology, pathogenesis, and clinical disease with comparison to *C. albicans*. *Clin Microbiol Rev* 1999;12:80–96.

72. Richter S, Cormican MG, Pfaller MA, et al. Fatal disseminated *Trichoderma longibracheatum* infection in an adult bone marrow transplant patient: species identification and review of the literature. *J Clin Microbiol* 1999;37:1154–1160.

73. Salkin IF, McGinnis MR, Dykstra MH, et al. *Scedosporium inflatum*, an emerging pathogen. *J Clin Microbiol* 1988;26:498–503.

74. Andriole VT. Current and future therapy of invasive fungal infections. *Curr Clin Topics Infect Dis* 1998;18:19–36.

75. Kirkpatrick WR, Revankar SG, McAtee RK, et al. Detection of *Candida dubliniensis* in oropharyngeal samples from human immunodeficiency virus-infected patients in North America by primary CHROMagar *Candida* screening and susceptibility testing of isolates. *J Clin Microbiol* 1998;36:3007–3012.

76. Walsh TJ, Chanock SJ. Laboratory diagnosis of invasive candidiasis: a rationale for complementary use of culture- and nonculture-based detection systems. *Int J Infect Dis* 1997;1(Suppl 1):11–19.

77. Branchini ML, Pfaller MA, Rhine-Chalberg J, et al. Genotypic variation and slime production among blood and catheter isolates of *Candida parapsilosis*. *J Clin Microbiol* 1994;32:452–456.

78. Diekema DJ, Messer SA, Hollis RJ, et al. An outbreak of *Candida parapsilosis* prosthetic valve endocarditis. *Diagn Microbiol Infect Dis* 1997;29:147–153.

79. Koehler AP, Chu KC, Houang ETS, et al. Simple, reliable, and cost-effective yeast identification scheme for the clinical laboratory. *J Clin Microbiol* 1999;37:422–426.

80. Levin AS, Costa SF, Mussi NS, et al. *Candida parapsilosis* fungemia associated with implantable and semi-implantable central venous catheters and hands of health care workers. *Diagn Microbiol Infect Dis* 1998;30:243–249.

81. Pfaller MA, Messer SA, Hollis RJ. Variations in DNA subtype, antifungal susceptibility, and slime production among clinical isolates of *Candida parapsilosis*. *Diagn Microbiol Infect Dis* 1995;21:9–14.

82. Price MF, LaRocco MT, Gentry LO. Fluconazole susceptibilities of *Candida* species and distribution of species recovered from blood cultures over a 5-year period. *Antimicrob Agents Chemother* 1994;8:1422–1424.

83. Weems JJ Jr. *Candida parapsilosis*: epidemiology, pathogenicity, clinical manifestations, and antimicrobial susceptibility. *Clin Infect Dis* 1992;14:756–766.

84. Wingard JR. Importance of *Candida* species other than *C. albicans* as pathogens in oncology patients. *Clin Infect Dis* 1995;20:115–125.

85. Wingard JR, Merz WG, Rinaldi MG, et al. Increase in *Candida krusei* infection among patients with bone marrow transplantation and neutropenia treated prophylactically with fluconazole. *N Engl J Med* 1991;325:1274–1315.

86. Wingard JR, Merz WG, Rinaldi MG, et al. Association of *Torulopsis glabrata* infections with fluconazole prophylaxis in neutropenic bone marrow transplant patients. *Antimicrob Agents Chemother* 1993;27:1847–1849.

87. Strausbaugh LJ, Sewell DJ, Ward TT, et al. High frequency of yeast carriage on hands of hospital personnel. *J Clin Microbiol* 1994;32:2299–2300.

88. Barnett J, Payne R, Yarrow D. *Yeasts: characteristics and identification*, 2nd ed. Cambridge: Cambridge University Press, 1990.

89. Odds FC, Bernaerts R. CHROMagar *Candida*, a new differential isolation medium for presumptive identification of clinically important *Candida* species. *J Clin Microbiol* 1994;32:1923–1929.

90. Pfaller MA, Houston A, Coffman S. Application of CHROMagar *Candida* for rapid screening of clinical specimens for *Candida albicans*, *Candida tropicalis*, *Candida krusei*, and *Candida* (*Torulopsis*) *glabrata*. *J Clin Microbiol* 1996;34:58–61.

91. Pincus D, Salkin I, Hurd N, et al. Modification of potassium nitrate assimilation test for identification of clinically important yeasts. *J Clin Microbiol* 1988;26:366–368.

92. Pinjon E, Sullivan D, Salkin I, et al. A simple, inexpensive, reliable method for the differentiation of *Candida dubliniensis* from *Candida albicans*. *J Clin Microbiol* 1998;36:2093–2095.

93. Sanati H, Messer SA, Pfaller MA, et al. Multicenter evaluation of broth microdilution for susceptibility testing of *Cryptococcus neoformans* against fluconazole. *J Clin Microbiol* 1996;34:1280–1282.

94. Tanner DC, Weinstein MP, Fedorciw B, et al. Comparison of commercial kits for detection of cryptococcal antigen. *J Clin Microbiol* 1994;32:1680–1684.

95. Walsh TJ, Melcher GP, Rinaldi MG, et al. *Trichospron beigelii*, an emerging pathogen resistant to amphotericin B. *J Clin Microbiol* 1990;28:1616–1622.

96. Kurup VP, Kumar A. Immunodiagnosis of aspergillosis. *Clin Microbiol Rev* 1991;4:439–456.

97. Walsh TJ, Chanock SJ. Diagnosis of invasive fungal infections: advances in nonculture systems. *Curr Clin Topics Infect Dis* 1998;18:101–153.

98. Espinel-Ingroff A, White T, Pfaller MA. Antifungal agents and susceptibility tests. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1640–1652.

99. Edman JC, Kovacs JA, Masur H, et al. Ribosomal RNA sequences shows *Pneumocystis carinii* to be a member of the fungi. *Nature* 1988;334:519–522.

100. Edman JC, Sogin ML. Molecular phylogeny of *Pneumocystis carinii*. In: Walzer PD, ed. Pneumocystis carinii *pneumonia*, 2nd ed. New York: Marcel Dekker, 1994:91–105.

101. Pixley FJ, Wakefield AE, Baerji S, et al. Mitochondrial gene sequences show fungal homology for *Pneumocystis carinii*. *Mol Microbiol* 1991;5:1347–1351.

102. Ruffolo JJ. *Pneumocystis carinii* cell structure. In: Walzer PD, ed. Pneumocystis carinii *pneumonia*, 2nd ed. New York: Marcel Dekker, 1994:25–43.

103. Stringer JR. *Pneumocystis carinii*: what is it, exactly? *Clin Microbiol Rev* 1996; 9:489–498.

104. Stringer SL, Hudson K, Blasé MA, et al. Sequence from ribosomal RNA of *Pneumocystis carinii* compared to those of four fungi suggests an ascomycetous affinity. *J Protozool* 1989;36:14S–16S.

105. Ruffolo JJ, Cushion MT, Walzer PD. Ultrastructural observations on life cycle stages of *Pneumocystis carinii*. *J Protozool* 1989;36:53S–54S.

106. Ul-Haque A, Plattner SB, Cook RT, et al. *Pneumocystis carinii* taxonomy as viewed by electron microscopy. *Am J Clin Pathol* 1987;87:504–510.

107. Edlind TD, Bartlett MS, Weinberg GA, et al. The ß-tubulin gene from rat and human isolates of *Pneumocystis carinii*. *Mol Microbiol* 1992;6:3365–3373.

108. Hadley WK, Ng VL. Pneumocystis. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1200–1211.

109. Watts JC, Chandler FW. Pneumocystosis. In: Connor DH, Chandler FW, Schwartz DA, et al., eds. *Pathology of infectious disease, vol. II*. Stamford, CT: Appleton & Lange, 1997:1241–1251.

110. Hughes WT. Natural mode of acquisition for *de novo* infection with *Pneumocystis carinii*. *J Infect Dis* 1982;145:842–848.

111. Hughes WT, Bartley DL, Smith BS. A natural source of infection due to *Pneumocystis carinii*. *J Infect Dis* 1983;147:595.

112. Sinclair K, Wakefield AE, Benerji S, et al. *Pneumocystis carinii* organisms derived from rat and human hosts are genetically distinct. *Mol Biochem Parasitol* 1991;45:183–184.

113. Stringer JR, Stringer SL, Zhang J, et al. Molecular genetic distinction of *Pneumocystis carinii* from rats and humans. *J Eukaryotic Microbiol* 1993;40:733–741.

114. Northfelt DW, Clement MJ, Safrin S. Extrapulmonary pneumocystosis: clinical features in human immunodeficiency virus infection. *Medicine* 1990;69:392–398.

115. Raviglione MC. Extrapulmonary pneumocystosis: the first 50 cases. *Rev Infect Dis* 1990;12:1127–1138.

116. Chusid MJ, Heyrman KA. An outbreak of *Pneumocystis carinii* pneumonia at a pediatric hospital. *Pediatrics* 1978;62:1031–1035.

117. Fenelon LE, Keane CT, Bakir M, et al. A cluster of *Pneumocystis carinii* infections in children. *BMJ* 1985;291:1683.

118. Jacobs JL, Libby DM, Winters RA, et al. A cluster of *Pneumocystis carinii* pneumonia in adults without predisposing illnesses. *N Engl J Med* 1991;324:246–250.

119. Millard PR, Heryet AR. Observations favoring *Pneumocystis carinii* pneumonia as a primary infection: a monoclonal antibody study on paraffin sections. *J Pathol* 1988;154:365–370.

120. Peters SE, Wakefield AE, Sinclair K, et al. A search for *Pneumocystis carinii* in post-mortem lungs by DNA amplification. *J Pathol* 1992;166:195–198.

121. Singer C, Armstrong D, Rosen PP, et al. *Pneumocystis carinii* pneumonia: a cluster of eleven cases. *Ann Intern Med* 1975;82:772–777.

122. Wakefield AE. DNA sequences identical to *Pneumocystis carinii* f. sp. *carinii* and *Pneumocystis carinii* f. sp. *hominis* in samples of air spora. *J Clin Microbiol* 1996;34:1754–1759.

123. Bigby TD, Margolsku D, Curtis JL, et al. The usefulness of induced sputum in the diagnosis of *Pneumocystis carinii* pneumonia in patients with the acquired immunodeficiency syndrome. *Am Rev Respir Dis* 1986;133:515–518.

124. Broadus C, Dake MD, Stulberg MS, et al. Bronchoalveolar lavage and transbronchial biopsy for the diagnosis of pulmonary infections in the acquired immunodeficiency syndrome. *Ann Intern Med* 1985;102:747–752.

125. Peglow SL, Smulian AG, Linke MJ, et al. Serologic responses to specific *Pneumocystis carinii* antigens in health and disease. *J Infect Dis* 1990;161:296–306.

126. Elvin K. Laboratory diagnosis and occurrence of *Pneumocystis carinii*. *Scand J Infect Dis* 1994;94(Suppl):1–34.

127. Leibovitz E, Pollack H, Moore T, et al. Comparison of PCR and standard cytological staining for detection of *Pneumocystis carinii* from respiratory specimens from patients with or at high risk for infection by human immunodeficiency virus. *J Clin Microbiol* 1995;33:3004–3007.

128. Lu JJ, Chen CH, Bartlett MS, et al. Comparison of six different PCR methods for detection of *Pneumocystis carinii*. *J Clin Microbiol* 1995;33:2785–2788.

129. Ribes JA, Limper AH, Espy MJ, et al. PCR detection of *Pneumocystis carinii* in bronchoalveolar lavage specimens: analysis of sensitivity and specificity. *J Clin Microbiol* 1997;35:830–835.

130. Armbruster C, Pokieser L, Hassl A. Diagnosis of *Pneumocystis carinii* pneumonia by bronchoalveolar lavage in AIDS patients. Comparison of Diff-Quik, fungifluor stain, direct immunofluorescence test and polymerase chain reaction. *Acta Cytol* 1995;39:1089–1093.

131. Einsele H, Hebart H, Roller G, et al. Detection and identification of fungal pathogens in blood by using molecular probes. *J Clin Microbiol* 1997;35:1353–1360.

132. Pfaller MA. Laboratory aids in the diagnosis of invasive candidiasis. *Mycopathologia* 1992;120:65–72.

133. Reiss E, Morrison CJ. Nonculture methods for diagnosis of disseminated candidiasis. *Clin Microbiol Rev* 1993;6:311–323.

134. Van Burik JA, Myerson D, Schreckhise R, et al. Panfungal PCR assay for detection of fungal infection in human blood specimens. *J Clin Microbiol* 1998;36:1169–1175.

135. Matthews RC. Comparative assessment of the detection of candidal antigens as a diagnostic tool. *J Med Vet Mycol* 1996;34:1–10.

136. Mitsutake K, Miyazaki T, Tashiro T, et al. Enolase antigen, mannan antigen, Cand-Tec antigen, and ß-glucan in patients with candidemia. *J Clin Microbiol* 1996;34:1918–1921.

137. Fujita SI, Lasker B, Lott TJ, et al. Microtitration plate enzyme immunoassay to detect PCR-amplified DNA from *Candida* species in blood. *J Clin Microbiol* 1995;33:962–967.

138. Kan VL. Polymerase chain reaction for the diagnosis of candidemia. *J Infect Dis* 1993;168:779–783.

139. Makimura K, Murayama SY, Yamaguchi H. Detection of a wide range of medically important fungi by the polymerase chain reaction. *J Med Microbiol* 1994;40:358–364.

140. Sandhu GS, Kline BC, Stockman L, et al. Molecular probes for diagnosis of fungal infections. *J Clin Microbiol* 1995;33:2913–2919.

141. Abruzzo GK, Flattery AM, Gill CJ, et al. Evaluation of echinocandin antifungal MK-0991 (L-743,872): efficacies in mouse models of disseminated aspergillosis, candidiasis, and cryptococcosis. *Antimicrob Agents Chemother* 1997;41:2333–2338.

142. Graybill JR. Future directions of antifungal chemotherapy. *Clin Infect Dis* 1992;14(Suppl 1):S170–S181.

143. Graybill JR, Najvar LK, Luther MF, et al. Treatment of murine disseminated candidiasis with L-743,872. *Antimicrob Agents Chemother* 1997;41:1775–1777.

144. Krishnarao TV, Galgiani JN. Comparison of in vitro activities of the echinocandin LY303366, the pneumocandin MK-0991, and fluconazole against *Candida* species and *Cryptococcus neoformans*. *Antimicrob Agents Chemother* 1997;41:1957–1960.

145. Marco F, Pfaller MA, Messer S, et al. *In vitro* activities of voriconazole (UK-109,496) and four other antifungal agents against 394 clinical isolates of *Candida* spp. *Antimicrob Agents Chemother* 1998;42:161–163.

146. Marco F, Pfaller MA, Messer SA, et al. Activity of MK-0991 (L-743,872), a new echinocandin, compared with those of LY303366 and four other antifungal agents tested against blood stream isolates of *Candida* spp. *Diagn Microbiol Infect Dis* 1998;31:33–37.

147. Pfaller MA, Messer S, Jones RN. Activity of a new triazole, Sch 56592, compared with those of four other antifungal agents tested against clinical isolates of *Candida* spp. and *Saccharomyces cerevisiae*. *Antimicrob Agents Chemother* 1997;41:233–235.

148. Pfaller MA, Messer SA, Hollis RJ, et al. *In vitro* susceptibilities of *Candida* blood stream isolates to the new triazole antifungal agents BMS-207147, Sch56592, and voriconazole. *Antimicrob Agents Chemother* 1998;42:3242–3244.

149. Powderly WG, Kobayashi GS, Herzig GP, et al. Amphotericin B-resistant yeast infection in severely immunocompromised patients. *Am J Med* 1988;84:826–832.

150. Sheehan DJ, Hitchcock CA, Sibley CM. Current and emerging azole antifungal agents. *Clin Microbiol Rev* 1999;12:40–79.

151. Vazquez JA, Lynch M, Boikov D, et al. *In vitro* activity of a new pneumocandin antifungal, L-743,872, against azole-susceptible and -resistant *Candida* species. *Antimicrob Agents Chemother* 1997;41:1612–1614.

152. Rex JH, Pfaller MA, Barry AL, et al. Antifungal susceptibility testing of isolates from a randomized multicenter trial of fluconazole versus amphotericin B as treatment of non-neutropenic patients with candidemia. *Antimicrob Agents Chemother* 1995;39:40–44.

153. Espinel-Ingroff A, Bartlett M, Bowden R, et al. Multicenter evaluation of proposed standardized procedure for antifungal susceptibility testing of filamentous fungi. *J Clin Microbiol* 1997;35:139–143.

154. Fromtling RA, Galgiani JN, Pfaller MA, et al. Multicenter evaluation of a broth macrodilution antifungal susceptibility test for yeast. *Antimicrob Agents Chemother* 1993;37:39–45.

155. National Committee for Clinical Laboratory Standards. *Reference method for broth dilution antifungal susceptibility testing for yeasts. Approved standard.* Document M27-A. Wayne, PA: National Committee for Clinical Laboratory Standards.

156. National Committee for Clinical Laboratory Standards. *Reference method for broth dilution antifungal susceptibility testing for filamentous fungi; proposed standard.* Document M38-P. Wayne, PA: National Committee for Clinical Laboratory Standards, 1999.

157. Pfaller MA, Rinaldi MG, Galgiani JN, et al. Collaborative investigation of variables in antifungal susceptibility testing of yeasts. *Antimicrob Agents Chemother* 1990;34:1648–1654.

158. Pfaller MA, Rex JH, Rinaldi MG. Antifungal susceptibility testing: technical advances and potential clinical applications. *Clin Infect Dis* 1997;24:776–784.

159. Rex JH, Pfaller MA, Galgiani JN, et al. Development of interpretive breakpoints for antifungal susceptibility testing: conceptual framework and analysis of *in vitro-in vivo* correlation data for fluconazole, itraconazole and *Candida* infections. *Clin Infect Dis* 1997;24:235–247.

160. Odds FC, Van Gerven F, Espinel-Ingroff A, et al. Evaluation of possible correlations between antifungal susceptibilities of filamentous fungi in vitro and antifungal treatment outcomes in animal infection models. *Antimicrob Agents Chemother* 1998;42:282–288.

53

Chlamydia, Mycoplasma, and Rickettsia

Stephen G. Jenkins

Gail L. Woods

J. Stephen Dumler

- CHLAMYDIA
- MYCOPLASMA
- RICKETTSIA

CHLAMYDIA

Part of "53 - Chlamydia, Mycoplasma, and Rickettsia"

Chlamydiae are obligate intracellular bacterial parasites of eukaryotic cells. Once called *Bedsonia* species (for Sir Samuel Bedson, who first isolated and described the causative agent of psittacosis), trachoma-inclusion conjunctivitis (TRIC) agents, and *Miyagawanella* species, these organisms were once considered viruses because of their inability to synthesize ATP or to grow on artificial media. They are dissimilar to the viruses, however, in that they contain both RNA and DNA, are susceptible to antibacterial agents, have cell walls similar to Gram-negative bacteria, synthesize protein utilizing their own ribosomes, and divide by binary fission. First isolated in embryonated eggs in 1957 and in cell culture in 1963, these organisms have been described as "energy parasites" (1).

Because of their unique developmental cycle, the chlamydiae are taxonomically classified in a separate order, Chlamydiales (from the Greek *chlamys*, a cloak draped around the shoulder, in reference to the intracytoplasmic inclusions "draped" around the infected cell's nucleus) and family Chlamydiaceae. The family recently was divided into two genera: *Chlamydia* and *Chlamydophila*, which contain three species pathogenic for humans: *Chlamydia trachomatis* (18 serovars), *Chlamydophila* (formerly *Chlamydia*) *psittaci*, and *Chlamydophila* (formerly *Chlamydia*) *pneumoniae* (2). Several distinct clinical syndromes are caused by the specific chlamydial serotypes (Table 53.1). *C. trachomatis* can be differentiated from *C. psittaci* based on its ability to form glycogen inclusions detectable in infected cell cultures with an iodine staining technique and its susceptibility to sulfonamides. Laboratory properties differentiating the chlamydial species are outlined in Table 53.2. Cellular infection with *C. trachomatis* results in a single inclusion that tends to displace the nucleus, whereas the inclusions of *C. psittaci* typically rupture early, resulting in a perinuclear distribution that fails to displace the nucleus. In addition, species can be distinguished definitively using monoclonal antibody stains.

TABLE 53.1. HUMAN CLINICAL DISEASE CAUSED BY CHLAMYDIAL SPECIES

Species	Serotypes	Disease	Major Symptoms	Infection Site	Mode of Transmission
C. pneumoniae	One	Pneumonia	Diffuse lung disease	Pulmonary	Person to person
C. psittaci	Several	Psittacosis (ornithosis)	Subclinical; mild respiratory illness or overt pneumonia	Pulmonary	Birds (including fecal droppings) to man; rarely human to human
C. trachomatis	A, B, Ba, C	Endemic trachoma	Chronic conjunctivitis; may result in blindness	Ocular	Human to human
	L1, L2, L3	Lymphogranuloma venereum	Transient genital ulcers, inguinal (males), pelvic or retroperitoneal (women) lymphadenopathy	Lymph nodes, rectum	Sexual
	D–K	Adult inclusion conjunctivitis	Acute follicular conjunctivitis	Ocular	Genital to eye
		Genital tract infection	Nongonococcal urethritis, cervicitis, epididymitis, proctitis, salpingitis	Genital tract, rectum	Sexual
		Inclusion conjunctivitis of the newborn	Acute mucopurulent conjunctivitis	Ocular	Genital to eye during passage through birth canal
		Pneumonitis/pneumonia of the newborn	Afebrile, chronic diffuse pulmonary disease; pneumonia	Pulmonary	Aspiration during passage through birth canal

C. trachomatis and *C. psittaci* possess a common group-reactive lipopolysaccharide antigen (3), have a similar small genomic size, and demonstrate approximately 10% DNA homology. *C. pneumoniae*, by comparison, lacks significant DNA homology with either of the other species, has different restriction endonuclease patterns, and appears to generally lack extrachromosomal DNA as is typically demonstrable in strains of *C. trachomatis* and the majority of *C. psittaci* isolates (4, 5). DNA homology generally is high within each species.

All three recognized chlamydial species exhibit two distinct forms during their unique developmental cycle. The 300-nm elementary body is especially adapted for extracellular survival and represents the infectious form in the life cycle. The elementary bodies of *C. pneumoniae* are pleomorphic but typically pear-shaped, although under some conditions they are round, similar to the elementary bodies of *C. trachomatis* and *C. psittaci* (2, 6). Upon contact with and attachment to a susceptible cell, the elementary body is actively taken into a phagosome by endocytosis, where the entire replication cycle takes place. Lysosomal fusion, however, fails to occur.

Over the next 6 to 8 hours, the elementary body undergoes reorganization into the metabolically active but noninfectious reticulate or initial body. For 18 to 24 hours, the reticulate body synthesizes new material and divides continuously by binary fission within a membrane-bound vacuole. Adapted for intracellular functions, reticulate bodies cannot survive in the extracellular milieu and are incapable of infecting other cells. After this period of synthesis and division, the reticulate bodies reorganize and condense into compact elementary bodies still bound within the phagocytic vesicle. Depending on the strain, somewhere between 48 and 72 hours after the cellular infection was initiated, the vacuoles release their contents and the cell lyses, releasing infectious elementary bodies.

Chlamydia trachomatis

Spectrum of Disease

Once designated the lymphogranuloma-venereum-trachoma inclusion conjunctivitis (LGV-TRIC) agent, *C. trachomatis* now is recognized as an important cause of cervicitis and pelvic inflammatory disease in women, nongonoccocal urethritis (NGU) and epididymitis in men, and inclusion conjunctivitis and interstitial pneumonia in neonates. The host range of *C. trachomatis* is restricted to humans. *C. trachomatis* is reportedly the most common

sexually transmitted disease in developed countries (7), causing an estimated 4 million infections per year in the United States alone, with resulting annual estimated direct and indirect costs totaling $1.4 billion (8).

TABLE 53.2. LABORATORY CHARACTERISTICS OF HUMAN CHLAMYDIAL SPECIES

Characteristic	*C. trachomatis*		*C. psittaci*	*C. pneumoniae*
	LGV Biovar	Trachoma Biovar		
Inclusions				
Refractile, vacuolar	+	+	0	0
Iodine staining glycogen	+	+	0	0
Intracerebral injection lethal for mice	0	+	+	0
Sulfadiazine susceptibility	+	+	0	0
Cell culture				
L-cell plaques	0	+	+	0
Enhanced recovery following				
Cell treatment with diethylaminoethyl-dextran	+	0	0	+

Inclusion Conjunctivitis

Ocular infections have been documented in approximately one third of neonates treated prophylactically with silver nitrate whose mothers were infected with *C. trachomatis* at parturition (9). Therefore, neonatal chlamydial conjunctivitis may be seen in up to 2 to 6% of infants born in areas where the prevalance of sexually transmitted infection is high.

Typically appearing 2 to 25 days after delivery (earlier in the setting of premature rupture of membranes), inclusion conjunctivitis is characterized by edematous, inflamed conjunctivae with a copious mucopurulent discharge. In cases of chronic or relapsing infection, conjunctival follicles may develop. The infection is usually self-limited, resolving spontaneously over a period of several months with no significant sequelae even without antimicrobial therapy.

Inclusion conjunctivitis in the adult, albeit relatively rare, presents acutely with a foreign body sensation, mucopurulent discharge, photophobia, and follicular conjunctivitis with or without keratitis. Except in cases of reinfection, symptoms disappear spontaneously over a period of several months to 2 years. Adult ocular infection usually occurs in patients whose sexual partner(s) show signs of genital chlamydial infection (10).

Trachoma

Endemic trachoma is a chronic disease characterized by recurrent corneal and conjunctival infection with *C. trachomatis*. Pannus formation and conjunctival scarring occur as sequelae to the chronic follicular keratoconjunctivitis. Historically, trachoma represented the single greatest cause of blindness in the world and probably remains today the single most common cause of preventable blindness (11). In 1973, approximately 400 million people worldwide were afflicted with trachoma, which resulted in blindness in 2 million (12). Although the number and size of endemic areas have been significantly reduced since 1950, trachoma continues to occur, mainly among the impoverished, in less developed regions of the world, especially of northern Africa, the Middle East, and northern India. It is still occasionally seen

in Native Americans in the southwestern United States. In endemic areas, children are often infected with serotypes A to C of *C. trachomatis* during the first 3 months of life. Transmission is not genital but by close contact with infected persons or from flies carrying infectious ocular material from eye to eye (13). Although relatively unusual, acute disease in adults over 20 years of age is much more common in women than in men, apparently due to closer contact with children.

Typically, trachoma begins acutely with inflammation of the bulbar and palpebral conjunctivae. Although initially a mixed cellular response is seen in the ocular exudate, macrophages and lymphocytes coalesce within a few weeks, forming soft, necrotic subconjunctival follicles. At this stage the disease may resolve spontaneously, unless reinfection occurs. Upon repeated infection, however, corneal vascularization begins and progresses concomitantly with conjunctival scarring. Repeated bacterial superinfections contribute to this process. The synergistic combination of bacterial infection, corneal vascularization, and conjunctival scarring results in blindness.

Lymphogranuloma venereum

C. trachomatis serovars L1, L2, L2a, and L3 cause a venereally transmitted disease, lymphogranuloma venereum (LGV), endemic to regions of South America, Africa, and Asia. Climatic bubo, tropical bubo, and estiomene are all synonyms for this disease referred to on occasion in 18th century literature (14). Although uncommon in the United States, with fewer than 500 cases typically reported yearly to public health departments, it is seen more frequently in the warmer, more humid regions of the country, especially the southeast, in people of lower socioeconomic status, in male homosexuals, and in individuals returning from other endemic regions. The disease is reported three times more frequently in men than in women. Transmission following laboratory accidents during aerosolization of the agent without appropriate biosafety measures and through fomites has reportedly occurred and resulted in pulmonary infection with pleural effusion, pneumonitis, and hilar and/or mediastinal lymphadenopathy. Probable reservoirs in the population include persons ignoring symptoms of cervical, urethral, or anorectal infection and individuals with asymptomatic disease.

LGV is the only chlamydial infection resulting in multisystem disease and causing constitutional symptoms. During the primary phase of the illness, 3 days to 3 weeks following exposure, a small painless herpetiform vesicle or ulcer develops, which heals without scarring. The lesions are only infrequently noticed by infected individuals on the affected sexual organ. In gay men and sometimes in women, the primary lesion may be an anorectal ulcer with associated symptoms of diarrhea, tenesmus, and a bloody, mucopurulent rectal discharge resulting from ulceration of the rectosigmoid colon. In such cases, fever and other constitutional symptoms are frequent, and perirectal or inguinal lymphadenopathy may result. A diffuse inflammatory response with mucosal ulceration, crypt abscesses, and granuloma formation is seen upon histologic examination of a rectal biopsy during this stage of the disease.

The primary phase of illness is followed 2 to 6 weeks later by a secondary stage characterized by suppurative regional lymphadenopathy and pronounced constitutional symptoms, including chills, fever, anorexia, headache, and arthralgias. Characteristic small, stellate abscesses surrounded by histiocytes are noted upon histologic examination of affected nodes. Other laboratory abnormalities commonly seen during this period include a leukocytosis, an elevated erythrocyte sedimentation rate, and abnormal liver function tests. During this phase, sequelae resulting from anorectal infection may include perirectal abscesses and rectovaginal or ischiorectal fistulas. Less frequently observed symptoms include hepatitis, erythema nodosum, “aseptic” meningitis, meningoencephalitis, and arthritis with a culture-negative effusion.

The late phase of the disease results from the sequelae caused by fibrotic changes and interference with normal lymphatic drainage. Chronic and progressive, infiltrative or ulcerative involvement of the scrotum, penis, or urethra with fistula development, ulcers, and posterior urethral strictures is not uncommon in men with untreated genital infection. Rectal strictures may result in “pencil stools.” Rarely, lymphedema may cause genital elephantiasis, with vulvar or penile enlargement and disfigurement. Chronic ulceration of the vulva and pedunculated perianal protrusions (lymphorrhoids) also may occasionally occur.

Nongonococcal Urethritis and Prostatitis

Isolated in 30% to 50% of all cases, *C. trachomatis* serovars D to K are the microorganisms, most commonly associated with male nongonococcal urethritis (NGU). Up to one third of men with urethral chlamydial infection may be without symptoms, and the organism can be isolated from a significant percentage of asymptomatic sexually active males. Although the symptoms of NGU are generally less acute than in cases of gonorrhea, the two cannot be reliably differentiated based on symptomology alone. In NGU, the urethral discharge is typically less purulent and smaller in volume than in gonococcal infection and results in less dysuria. Signs of NGU include erythema and tenderness of the meatus and a urethral exudate often demonstrable only upon urethral stripping in the morning prior to urination. One to two percent of chlamydial urethral infections have been reported to evolve into epididymitis.

C. trachomatis is the most frequently isolated agent in young heterosexual men with this disease in whom structural genitourinary tract abnormality is absent (15). *C. trachomatis* also has been reported to be associated with “venereal” Reiter syndrome (16) and has been isolated in up to 80% of cases of postgonococcal urethritis. Among homosexual men, it may cause proctitis or proctocolitis, depending on the serotype of the infecting strain.

Gynecological Infections

Genital infection with *C. trachomatis* is more prevalent in women than in men, but often escapes clinical detection. The adult female may have asymptomatic endocervical infection (17) or may present with mucopurulent cervicitis with vaginal discharge, urethritis, bleeding cervical erosions, proctitis, pelvic inflammatory disease, and/or salpingitis. Vaginitis does not occur. *C. trachomatis* has been isolated from the cervix of 30% to 60% of women with a history of sexual contact with a partner with NGU or gonorrhea or who diagnosed themselves with gonococcal infection, from 10% to 20% of women presenting to sexually

transmitted disease (STD) clinics without a history of contact with parties with urethritis, and from 5% to 10% of young women visiting family planning centers, prenatal clinics, or gynecologic clinics (18). Detection rates are higher in sexually active teenagers, particularly those who are pregnant and those attending STD clinics. The prevalence of infection with *C. trachomatis* also is higher in those of lower socioeconomic status, of black race, residing in urban areas, of divorced marital status, and with increased numbers of sexual partners. It also appears to be higher in users of oral contraceptives (19). In most studies, *C. trachomatis* is several times more prevalent than *Neisseria gonorrhoeae*.

The consequences of chlamydial infection in young women are significant. Up to 10% of women with genital infection show signs and symptoms of acute salpingitis. Among women with ascending infection, important sequelae include chronic pelvic pain, infertility resulting from occlusion of the fallopian tubes, and ectopic pregnancy. Asymptomatic cervical infection is common, and most women with involuntary infertility resulting from bilateral tubal damage fail to give a history of prior pelvic infection. Accordingly, it is highly probable that the proportion of women with ascending chlamydial infection is much higher than 10%. *C. trachomatis* has been detected in approximately two thirds of women with mucopurulent cervicitis. Over 80% of women with chlamydial genital infection demonstrate this syndrome upon gynecologic examination (20).

Acute endometritis resulting from *C. trachomatis* can progress to salpingitis following insertion of an intrauterine device, dilatation and curettage, hysterosalpingography, and subsequent to abortion or parturition. In a number of reported studies, *C. trachomatis* has been isolated from the urethra or cervix of approximately one third of women with laparoscopically documented acute pelvic inflammatory disease and from a comparable portion of women with clinical evidence of salpingitis, symptoms of which include low abdominal pain, discomfort during sexual intercourse, and abnormal vaginal bleeding. Serologic studies have demonstrated that *C. trachomatis* infection has occurred in up to 60% of women with salpingitis. In addition, the agent can be isolated more frequently than *N. gonorrhoeae* from women with this disease, although symptoms typically appear to be less severe. Serologic detection of IgM or high titers of IgG antibodies to *C. trachomatis*, both suggestive of recent or active infection, have been reported in up to 87% of women with acute perihepatitis or peritonitis (Fitz-Hugh Curtis syndrome) (21). Related intraperitoneal disease reportedly associated with chlamydial infection include perisplenitis and periappendicitis.

The risk for developing involuntary infertility as a result of tubal occlusion increases with the severity and number of episodes of salpingitis. The overall risk of infertility reportedly approximates 10% following a single episode of salpingitis, 30% with two episodes, and over 50% following infection on three or more occasions. Supporting this association, *C. trachomatis* has been isolated from a large porportion of peritubal and tubal specimens cultured from infertile women with tubal obstruction (22).

The dysuria-pyuria (urethral) syndrome also has been ascribed to *C. trachomatis* (23). In young women with frequency and dysuria, *C. trachomatis* often can be isolated. The agent also has been detected in pus expressed from the Bartholin's gland in cases of bartholinitis and occasionally from the rectum in women with chlamydial genital infection. Conflicting evidence exists on the importance of *C. trachomatis* as a cause of chorioamnionitis, premature rupture of membranes, premature delivery, spontaneous abortion, and stillbirth and on the influence of prior infection on the success of *in vitro* fertilization procedures.

Neonatal Pneumonia

Onset of classic chlamydial interstitial pneumonia in infants is gradual, with a range from 2 weeks to 3 months. The majority of cases are detected within 3 to 6 weeks of birth. Although the course of illness is protracted, the majority of infants remain consistently afebrile. In most cases, serum IgG and IgM levels are elevated, and the peripheral blood eosinophil count exceeds 300/mm^3 (24). Precedent conjunctivitis and/or ear abnormalities have been noted in many cases (25). Rhinitis with nasal obstruction, typically without nasal discharge, often represents the initial clinical manifestation.

A sequence of closely spaced staccato coughs each separated by a brief inspiration, tachypnea, hyperinflation, and rales with diffuse interstitial involvement typify the illness. Roentgenographic findings usually include pulmonary infiltrates, peribronichial thickening, and focal consolidation. A specific antibody response to the infecting strain of *C. trachomatis* is often demonstrable. A pronounced eosinophilic exudate is observed upon examination of tracheal secretions. Clinical illness usually lasts for several weeks, but rales and radiologic abnormalities may persist for months.

Epidemiology

The basic epidemiology of *C. trachomatis* infection was defined early in this century. The associaton between neonatal conjunctival infection, the maternal cervix, and infection among siblings, suggesting transmission by mucous membrane contact, was described as early as 1942 (26). Although the etiologic agent was thought to be a virus, the full spectrum of disease including female cervical infection, neonatal inclusion conjunctivitis, urethritis, and ophthalmitis in adults was comprehensively delineated prior to the advent of the antibiotic era. Although the role of *C. trachomatis* in NGU is still unclear, the agent can be isolated from a large proportion of men with this syndrome as well as from a significant number of female contacts with these men. Likewise, the majority of fathers of infants with inclusion conjunctivitis are positive for *C. trachomatis* upon urethral culture. Concomitant chlamydial infection at several body sites also is relatively common, and the majority of individuals with chlamydial conjunctivitis have simultaneous pharyngeal or genital colonization and/or infection.

Several conditions are associated with an increased prevalence of female chlamydial infection. *C. trachomatis* cervical infection appears to be more common in women with a history of other sexually transmitted disease. In addition, there appears to be an increased culture recovery rate from specimens collected during the first and fourth weeks of the female menstrual cycle. Risk factors in adolescent women include a younger age at first sexual intercourse, increased number of years of sexual activity, and use of

oral contraceptives. Although contact with infected mucosal membranes represents the most common route of exposure, ocular trachoma may be acquired via fomite inoculation and by swimming in contaminated water. Trachoma also has been acquired during ocular tonometry from equipment improperly sterilized between each procedure.

C. trachomatis also can cause sexually transmitted disease in children. When vaginal, rectal, or pharyngeal chlamydial infection occurs in children outside of the neonatal period, an investigation should be conducted to determine whether sexual abuse has occurred. Although infection with *C. trachomatis* in children does not prove that they are victims of sexual abuse, it should alert caretakers to the possibility of such a problem. Similarly, children undergoing medical evaluation for sexual abuse should be cultured for chlamydiae from pharyngeal, rectal, and vaginal sites, as appropriate.

Diagnostic Procedures

Specimen Selection Collection

Specimens collected for the detection of chlamydiae should be handled with care. Such specimens may contain other agents of potential risk to the health care worker. Following aerosolization and inhalation of LGV strains of *C. trachomatis* in a laboratory setting, accidental infection resulting in pneumonitis and mediastinal lymphadenitis has been reported (27).

All samples collected for chlamydial isolation should be stored at refrigerator temperature prior to processing and inoculated to appropriate cell lines within 48 hours. If specimen processing is expected to be delayed for more than 48 hours after collection, specimens should be frozen at -70°C or below.

Although LGV strains of *C. trachomatis* have been isolated from clotted whole blood, blood culture is not recommended for routine diagnosis of infection. The specimen most frequently submitted to the laboratory for chlamydial isolation is the genital swab. Scrapings are preferred, but specimens from the conjunctivae and the genital tract are most easily collected with a swab. Cervical specimens also may be collected with a cytologic sampling brush. Materials utilized for some swabs have been shown to be toxic to the cell lines used to support the *in vitro* replication of chlamydiae and to inhibit chlamydial multiplication.

Swabs prepared with dacron, rayon, or cotton fibers should be used for specimen collection. Likewise, swabs with plastic or metal shafts are superior to toxic wooden-shaft varieties. Confusing artifacts may result with some dacron swabs when iodine staining techniques are used to demonstrate *C. trachomatis* inclusions. Similarly, alginate swabs can result in the formation of confusing artifacts when dark-field microscopy is utilized to reveal chlamydial inclusions by Giemsa stain.

Removal of mucus and pus prior to obtaining an endocervical sample reduces the rate of bacterial contamination and the frequency of cytotoxicity in inoculated cell cultures and results in better smears for direct fluorescent antibody staining. Infected cells, required for the detection of chlamydiae by cell culture techniques, are also increased in number in samples collected following removal of excess cervical mucus. Concomitant culture of the cervix and the urethra with subsequent specimen pooling increases the culture positivity rate substantially at minimal added expense.

Specimens submitted for diagnosis of chlamydial infection include scrapings or swabs from the conjunctivae, urethra, cervix, and rectum, bubo pus, sputa, throat washings, and tissue. Because epithelial cells, as opposed to mucus or pus, must be collected, vigorous rubbing of the swab against the involved site is required for optimal specimen acquisition. When collecting eye specimens, the eyelid should be everted and the swab rolled directly over the surface of the conjunctivae. In cases of urethritis in the male, the swab should be inserted 3 to 5 cm into the urethra and withdrawn using a rotating motion. In cases of cervicitis in the female, the endocervical specimen should always be collected from the squamocolumnar junction.

Deep posterior nasopharyngeal specimens are required for the diagnosis of chlamydial respiratory tract infection. Techniques used for the collection of samples for the detection of *Bordetella pertussis* are appropriate for this purpose. In suspected cases of lymphogranuloma venereum, pus from an affected lymph node should be apsirated with a syringe needle following insertion through apparently healthy tissue rather than directly into the bubo. Infected fluctuant nodes usually contain frank pus. If none is aspirated, sterile nonbacteriostatic saline should be injected into the node and drawn back into the same syringe in an attempt to collect material for diagnostic testing. Collection of specimens in cases of proctitis or lower gastrointestinal tract infection is best accomplished using direct anoscopic visualization of the suspected sites, optimally from ulcerative or hypertrophic lesions. Isolation of chlamydiae from the upper genital tract represents the only valid technique for documenting *C. trachomatis* involvement in cases of endometritis and acute salpingitis. Optimal gynecologic techniques for the collection of clinical specimens include laparoscopically directed fallopian tube biopsy or needle aspiration and/or guarded endometrial biopsy or washing (28).

No single transport medium has yet been accepted as optimal for the recovery of *C. trachomatis* from clinical specimens. Although labile at temperatures above 6°C, the organism is relatively stable at refrigeration temperatures. A single freeze-thaw cycle results in at least a 1-log loss in infectivity. If a specimen requires storage for less than 72 hours, it is therefore recommended that the sample be maintained at 4 to 6 rather than be frozen. Some laboratorians, however, recommend that all specimens be frozen at less than or equal to -70°C until processing. Two commonly employed transport media are a sucrose-phosphate-glutamate medium and a cell culture medium supplemented with 10% fetal bovine serum. Most chlamydial infections occur in body sites normally colonized with bacterial flora. As a result, specimens collected for the isolation of *C. trachomatis* usually must be treated with antibiotics prior to cell culture inoculation. Antibiotics without chlamydial activity but active against the organisms commonly encountered at the collection site usually are incorporated into the transport medium itself. Aminoglycosides such as gentamicin, uniformly inactive against chlamydiae, have proven to be particularly useful for this purpose. Formulations commonly include vancomycin (100 g/mL) and gentamicin (10 g/mL) as antibacterial agents and nystatin (25 to 50 g/mL) or amphotericin B (2.5 to 4.0 g/mL) to suppress the growth of

yeasts. Tetracyclines, erythromycin, and penicillins should not be used because they interfere with chlamydial isolation procedures. Specimens such as homogenized tissue, semen, and bubo aspirates should be diluted 1:10 and 1:100 in cell culture medium prior to shell vial inoculation because they are often toxic to cells in culture.

Direct Microscopy

An adequate specimen for cytologic examination should include at least 1,000 epithelial cells spread evenly over the surface of the slide such that the cells can be examined individually. Although historically the Giemsa stain proved useful for the detection of chlamydial inclusions, the technique is no longer recommended for routine diagnosis of *C. trachomatis* infection because direct fluorescent antibody and culture systems have proven to be significantly more sensitive. The Giemsa-stained smear may still represent the diagnostic technique of choice in cases of inclusion conjunctivitis of the neonate when large numbers of infected inclusion-containing cells usually can be detected in a rapid and cost-effective manner (Fig. 53.1).

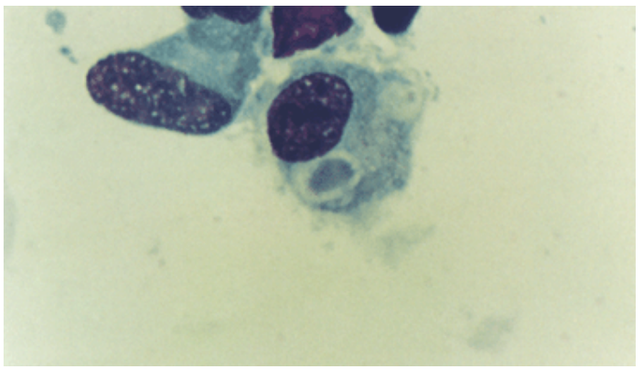


FIGURE 53.1. Giemsa-stained chlamydial inclusion in cell scraping from a case of neonatal conjunctivitis.

Direct visualization of chlamydial elementary bodies using fluorescein-conjugated monoclonal antibodies on urethral, endocervical, conjunctival, and rectal specimens and in respiratory tract or nasopharyngeal aspirates from infants with pneumonitis has proven to be a rapid and sensitive technique for the diagnosis of *C. trachomatis* infection. This technique obviates the need for transport media, and fixed slides are stable for up to 1 week at either room or refrigeration temperature.

Although relatively time-consuming in the laboratory, the method is especially useful in low-volume settings and when specimens must be transported from distant collection sites prior to testing. A distinct advantage of direct fluorescent antibody testing over antigen detection and culture techniques is the capability to evaluate the adequacy of the specimen during slide examination. Columnar epithelial cells attest to the fact that the sample was collected properly, whereas excessive cervical mucus and/or a preponderance of squamous epithelial cells indicate inadequate site preparation or specimen collection. Murine monoclonal antibodies directed against the major outer membrane proteins (MOMP) of all known human serovars of *C. trachomatis* and in both forms of the organism (the infectious elementary body and the replicating, metabolically active reticulate body) are generally felt to be superior to antibody preparations specific for chlamydial lipopolysaccharide (LPS) because the fluorescence is more intense with MOMP preparations. Bacteria other than chlamydiae, however, may sometimes stain with anti-MOMP antibodies either through nonspecific immunoglobulin binding or from cross-reactivity with shared epitopes. This phenomenon is observed frequently with rectal specimens, possibly because of the large numbers of bacteria normally present in such specimens or resulting from cross-reacting antigens in *Peptostreptococcus productus* (29). In specimens from other body sites, cross-reacting bacterial species are less common and can reliably be differentiated from fluorescing elementary bodies on a morphologic basis. Because common epitopes exist between chlamydial and other Gram-negative bacterial lipopolysaccharides, direct fluorescent antibody reagents directed against chlamydial LPS have the potential for cross-reactivity with bacterial species. This would not be expected with antibody preparations directed against the chlamydial MOMP.

A high-quality fluorescence microscope is essential for proper examination of clinical specimens using direct fluorescent antibody (DFA) techniques. Differences in optical alignment, illumination power, and filter quality and wavelength will cause variability in the fluorescent intensity of elementary bodies. To control for these variables, positive and negative controls should be tested routinely with patient samples to substantiate satisfactory reagent and microscope performance. Slides are typically scanned under 400× total magnifiication, and suspicious structures are confirmed using a 1,000× oil immersion objective. The sensitivity of DFA methods when compared with culture is a function of the sensitivity of the comparative cell culture technique, the patient population examined, and the number of elementary bodies required for interpretation of a smear as positive. The sensitivity of DFA in studies of female cervical specimens has ranged from 70% to 100%, whereas the specificity typically exceeds 95% for both men and women. In studies that have employed a third technique to resolve discrepant results between DFA and culture for the diagnosis of chlamydial infection in women, false-negative cultures have been reported for endocervical specimens. DFA appears to be more sensitive for the detection of elementary bodies in female endocervical samples than in male urethral specimens (approximately 70%), probably because of the discomfort associated with specimen collection. Because lower numbers of elementary bodies are typically observed in urethral samples than in endocervical specimens, the number required for the interpretation of a test as positive may have to be lowered for urethral specimens to achieve adequate test sensitivity.

Culture Techniques

McCoy cells, a heteroploid mouse cell of uncertain origin, are most frequently used for the isolation of chlamydiae from clinical specimens. Although a number of other cell lines have proven satisfactory for this purpose, extensive research has been done, and all standard procedures have been developed using this line. Regardless of the cell line used, the most critical step in enhancing chlamydial replication is increased contact between the infectious elementary bodies and the cell monolayer. This can be accomplished mechanically (centrifugation) or by changing the

surface charge of the cell (treating with DEAE-dextran). Cell irradiation or treatment with an antimetabolite (5-iodo-2-deoxyuridine, cytochalasin B, or cycloheximide) further enhances chlamydial replication. Specimen centrifugation onto cycloheximide-treated McCoy cell monolayers, either on coverslips in 1-dram shell vials or in microtiter wells is the method of choice for the culture of chlamydiae at this time.

To process the specimen for cell culture, all cell-containing fluid should be expressed from the swab into the transport medium container, the swab is discarded, and the specimen is vortexed in a centrifuge tube containing glass beads. Coarse debris may be removed from the specimen by centrifuging at a low speed for a brief period of time. Cell monolayers should then be inoculated with the clinical specimen by centrifuging at 2,500 to 3,000 g (centrifugation of microtiter plates is limited to approximately 1,500 g) for 1 hour at 30°C to 35°C.

Controlled temperature centrifuges assist in maintaining a constant acceptable temperature during centrifugation. The inoculum should be removed after 1 to 2-hours (immediately following centrifugation for specimens likely to be toxic to the McCoy cells), and fresh medium containing 1 g/mL of cycloheximide should be added. Cells then should be incubated at 35°C to 37°C in air with 5% CO_2 (CO_2 is unnecessary for stoppered vials) and stained to detect chlamydial inclusions at times appropriate for the staining system employed. Cells should be methanol-fixed prior to staining. The optimal incubation time for both the Giemsa and the iodine (Fig. 53.2) staining techniques is 60 to 72 hours, when glycogen concentrations are at a maximum. Immunofluorescent and immunoperoxidase (Fig. 53.3) procedures are capable of detecting chlamydial inclusions as early as 18 hours after cell culture challenge, but cultures inoculated in duplicate are typically incubated and stained after 24 hours and after 48 hours incubation. Immunoperoxidase and immunofluorescent staining techniques also are more sensitive than the iodine stain, but are more expensive to perform and require more attention to staining.

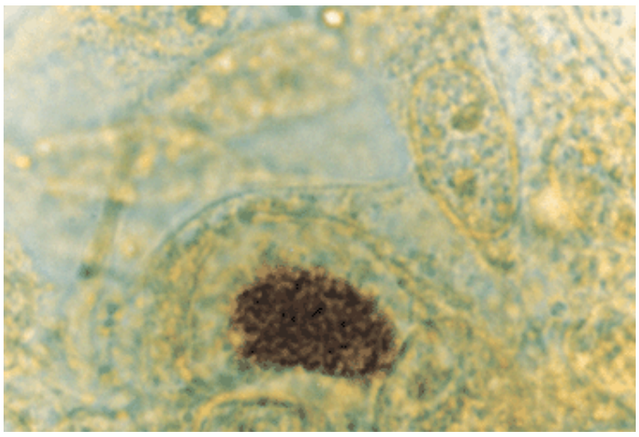


FIGURE 53.2. Iodine-stained *Chlamydia trachomatis* inclusion in infected MyCoy cell culture.

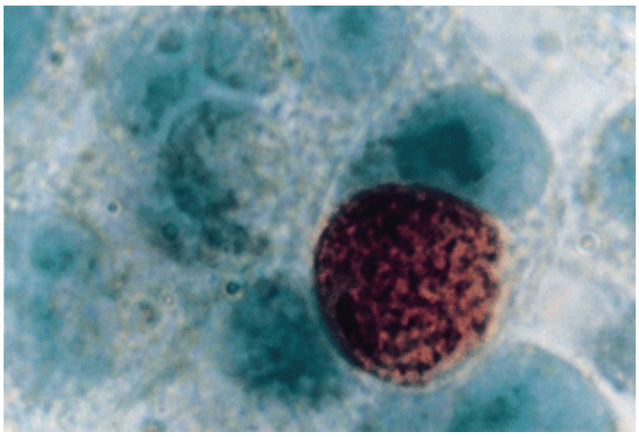


FIGURE 53.3. Immunoperoxidase-stained chlamydial inclusion in infected McCoy cell culture.

The use of 96-well microdilution plates for the isolation of chlamydiae affords considerable cost savings both for media and plastic disposables and expedites the simultaneous processing of multiple specimens. Two possible problems with this technique are the potential for cross-contamination as a result of the proximity of the microtiter wells and the possibility that the decreased cell surface may result in lower isolation rates. Equivalent isolation rates are seen with vials and microtiter plates when specimens are collected from symptomatic individuals. When asymptomatic populations are screened, however, shell vials are clearly superior. Blind passage is required in this setting for maximal recovery rates (approximately 25% of asymptomatic individuals and 10% of symptomatic persons will be positive only when specimens are tested by this technique), and subculture from microtiter wells is considerably more difficult and less efficient than from shell vials. To accomplish the passage, an inoculated cell culture incubated for 72 to 96 hours should be homogenized and the suspended cells used to challenge a second cell monolayer.

Enzyme-Linked Immunoassays

Screening tests based on enzyme-linked immunosorbent assay (EIA) technology were introduced shortly after the commercial availability of DFA tests. In general, EIAs are highly useful for screening symptomatic and asymptomatic women in high- and moderate-risk groups, but they have lower sensitivity for screening women in low-risk populations (30,31,32,33,34,35 and 36). Likewise, these assays are satisfactory for screening symptomatic males for chlamydial infection but demonstrate relatively low positive and negative predictive values when asymptomatic men are tested (37). In a multicenter study, this procedure showed an 87.2% sensitivity and 98.9% specificity compared to urethral swabs tested by cell culture. When screening female urine specimens, however, the current assay appears to lack specificity.

Investigators also have generally found EIA to perform in a satisfactory manner when testing ocular specimens, but limited data suggest that EIA may not perform as well when examining specimens from the respiratory tract (38, 39). Studies also have shown that false-positive results occur when testing clinical specimens for chlamydiae using EIA tests, particularly with samples from the lower gastrointestinal tract. As a result, commercially available EIA kits are not currently approved for the testing of rectal specimens.

Other available techniques include filtration EIA procedures.

In these assays, chlamydiae are trapped in nitrocellulose membranes and detected using monoclonal antibodies directed against the chlamydial LPS using a radiometric or colorimetric antibody detection technique. These tests may offer increased sensitivity over earlier assays and clearly decrease the time required to test an individual specimen. Relatively inexpensive, these procedures also lend themselves to rapid nonbatch testing of multiple specimens as they are received into the laboratory.

Nucleic Acid Probes

An acridinium-ester-labeled DNA probe complementary to *C. trachomatis* ribosomal RNA is commercially available for direct detection of *C. trachomatis* in endocervical, urethral, and conjunctival swab specimens. Instrumentation required for test performance includes a water-bath and luminometer. Test results are available in 2 to 3 hours. The sensitivity of the DNA probe test compared to conventional cell culture as the reference method varies from about 75% to 97% for both endocervical and urethral swab specimens (40,41,42,43,44 and 45). The specificity of the probe test is ≥97% for all specimen types, and it can be improved with a probe competition assay (46).

Nucleic Acid Amplification

Nucleic acid amplification tests for direct detection of *C. trachomatis* in clinical specimens are available from a few manufacturers. Acceptable specimens include endocervical and male urethral swab specimens and male and female urine samples. Total processing time is 4 to 5 hours. Data from several studies indicate that the nucleic acid amplification tests are highly specific and are more sensitive than conventional cell culture (47,48,49,50,51 and 52). However, given that the amplification assays are more expensive than the other commercial nonculture tests, laboratory directors must determine whether the tests will be cost effective in their particular patient population.

Serology

Serologic techniques are not recommended for diagnosis of chlamydial genital tract infection. Although most individuals with documented chlamydial infection have detactable antibodies to *C. trachomatis*, there also appears to be a high prevalence of antibody positivity in uninfected individuals in sexually active populations. This is because chlamydial genital and respiratory tract infection is common, and antibodies to chlamydiae persist following the resolution of infection. In addition, considering the chronic nature of the disease, it is often difficult to demonstrate a diagnostic fourfold rise in antibody titer in persons with active infection.

Skin tests for the diagnosis of LGV, however, have been available since 1938 (53). Originally, skin test antigen was prepared from sterilized bubo or genital exudate material collected from individuals with LGV. The widespread application of the “Frei test” became possible when it was found that large quantities of LGV antigen could be prepared in the yolk sac of developing chick embryos (54). Because the test lacks acceptable sensitivity and specificity (55), serologic techniques have, for the most part, supplanted it. A single titer of 256 or higher by complement fixation generally is considered diagnostic of infection with LGV. This technique lacks sensitivity for diagnosing other infections caused by *C. trachomatis*.

The microimmunofluorescence test (MIF) for the detection of antibodies to chlamydiae was developed in the early 1970s. It is highly reliable in detecting IgG-class antibodies in individuals previously infected with chlamydiae, and, as a result, has proven useful for epidemiologic investigations examining rates of infection with *C. trachomatis* (56, 57). The MIF test also detects IgM antibodies. Its primary use is in the clinical diagnosis of chlamydial pneumonitis in neonates. Although the MIF test primarily detects serotype-specific antibody, chlamydial genus reactivity can be measured if reticulate body antigen is incorporated into the test reagent (58).

Interpretation of Laboratory Results

Most currently available laboratory tests for the diagnosis of *C. trachomatis* infection are relatively expensive and labor-intensive. The U.S. Centers for Disease Control and Prevention (CDC) recommended in 1985 that symptomatic individuals at high risk for sexually transmitted chlamydial infection be treated with appropriate antibiotics empirically (18). This approach is both cost-effective and rational in clinical settings serving high-risk populations for patients willing to acquiesce to treatment for an unproven infection.

Nonculture techniques for the diagnosis of *C. trachomatis* infection generally are highly sensitive but lack specificity, particularly in low-risk populations. As a result, positive screening tests in low-risk patients with nonspecific symptoms should be confirmed with more specific tests, such as culture, before initiation of antimicrobial therapy.

Specific test selection for the diagnosis of infection with *C. trachomatis* usually is based on cost, public health, and laboratory expertise. Culture techniques are generally considered cost-effective for the diagnosis of chlamydial infection in high-risk women presenting to STD clinics, whereas empiric antibiotic therapy appears to be more cost-effective for men in the same setting (59). In populations with chlamydial cervical infection prevalences from 2% to 7%, antigen detection tests have proven cost-effective for screening purposes (60, 61).

Because of the limitations of all of the current testing techniques, the selection of specific test methods for the diagnosis of chlamydial infection by the laboratory should take into account the test cost, the projected number of tests to be performed, personnel requirements, potential problems with specimen transport, and the patient population to be examined. Cell culture techniques are 100% specific, allow for characterization of isolated strains, and are applicable to all types of clinical specimens. They currently represent the methods of choice when unequivocal specificity is required. Demanding somewhat less technical expertise, antigen detection procedures, by comparison, offer the advantages of ease of specimen transport, decreased hazard to the laboratory worker, and generally are less labor-intensive. DFA, ELISA, nucleic acid probe, and culture methodologies appear to be fairly comparable in terms of sensitivity. Nucleic acid amplification tests offer the highest sensitivity.

The lower specificities of direct antigen detection tests in low-risk populations generally preclude their use for screening purposes in such settings. Considerable social stigma is still attached to the presence of sexually transmitted disease. To minimize the emotional cost associated with a report of a positive result, clinicians

should communicate to their patients the predictive value of a positive test with the specific technique employed by the laboratory.

Therapeutic Considerations

Sulfonamides, macrolides, and tetracyclines are all highly active *in vitro* against *C. trachomatis*. In cases of inclusion conjunctivitis, topical therapy with tetracycline, erythromycin, or sulfonamide for 2 to 3 weeks will often suppress signs of infection but will not necessarily eradicate the infecting agent from the nasopharynx or conjunctiva. This is of particular import in infants at risk for the development of chlamydial pneumonia. *C. trachomatis* colonization of the conjunctivae has been shown to persist for up to 40 weeks and nasopharyngeal carriage up to 18 months or more in the face of topical therapy. Therefore, systemic treatment of neonatal inclusion conjunctivitis with erythromycin, or, in older infants, with sulfonamides is indicated to eradicate the ocular infection as well as to prevent pneumonia. Systemic therapy with a tetracycline is recommended for adult disease. In chronic trachoma, therapy is more difficult. Extended treatment with doxycycline appears to be the most effective regimen. Surgical correction of trachoma-associated lid deformities reduces the incidence of blindness.

The best approach to the prevention of neonatal chlamydial infection is prenatal diagnosis and treatment of infected parents. Ocular prophylaxis with silver nitrate does not prevent the establishment of chlamydial infection in the infant, whereas prophylaxis with erythromycin appears to be at least somewhat effective. It does not, however, decrease the overall acquisition of *C. trachomatis* (62) and is ineffective if application to the infant's eyes is delayed.

The recommended antimicrobial agent for the treatment of LGV is tetracycline. Alternative regimens include courses of erythromycin, doxycycline, or sulfamethoxazole. All of these regimens rapidly improve constitutional symptoms, but have little impact in the duration of healing buboes (63). Late complications are not generally affected by antibiotic therapy. Surgical intervention is often required to repair strictures and fistulas and to address elephantiasis.

The most effective agents for the treatment of genital tract infection with *C. trachomatis* are azithromycin and the tetracyclines. Chloramphenicol, rifampin, and oral ampicillins (administered in high doses) and the newer fluoroquinolones such as ofloxacin have all been reported to be effective for treatment of infection with *C. trachomatis* (64,65 and 66), but they are not recommended as first-line agents for chlamydial infection. Optimal treatment during pregnancy for *C. trachomatis* infection has not been established. A regimen of erythromycin (stearate, base or ethylsuccinate; not estolate) or azithromycin is probably the best approach.

With the exception of azithromycin, single-dose therapy with any antibiotic does not reliably eradicate *C. trachomatis* from genital tract sites. Therapy for an extended period of time, at least 10 days, is recommended for complicated infections. All attempts to treat pelvic inflammatory disease on an outpatient basis should include intramuscular injection of a third-generation cephalosporin (such as cefotaxime or ceftriaxone) to provide coverage for *N. gonorrhoeae* as well as tetracycline for chlamydiae. The fluoroquinolones as a group have significant activity against gonococci and some, especially ofloxacin, are highly active against *C. trachomatis*. Further clinical trials will help define the role of these agents for the treatment of chlamydial genital tract infections.

Chlamydophila (Formerly Chlamydia) psittaci

Psittacosis (parrot fever, ornithosis) is a zoonotic disease primarily infecting birds that is accidentally transmitted to man. First described in psittacine birds (parrots and parakeets), *C. psittaci* has been shown to infect most avian species as well as many nonprimate mammals, causing varying degrees of illness. In Switzerland in 1879, Ritter described several human cases of unusual pneumonia associated with tropical bird exposure. Morange named the disease (from the Greek word for parrot, psiittakos) following an outbreak in 1894 in which he attributed the source of the disease to parrots. Meyer chronicled outbreaks in several countries from 1929 to 1930 following large-scale importation of infected South American parrots to the United States and Europe (67). During the investigation of an outbreak at the London Zoo in 1930, Bedson isolated the causative agent from avian and human tissues. Rivers and his associates explained that the probable route of transmission to man was via the upper respiratory tract following inhalation of dried excreta from birds shedding the agent. During ecologic studies on infection in Australian wild birds, Burnet proved that fledglings acquired the infection in the nest from asymptomatic parents. Between 1965 and 1980, 40 to 160 human cases were reported annually in the United States (68) with a 1% to 5% mortality rate.

Pathogenesis and Spectrum of Disease

In man, following inhalation into the respiratory tract, *C. psittaci* migrates to the reticuloendothelial cells of the spleen and liver, where it replicates intracellularly within macrophages. Invasion of the lung and other organs then follows by hematogenous seeding. This two-step process explains the relatively long incubation period (7 to 15 days or longer) observed in this disease. A predominantly lymphocytic inflammatory response occurs in both the alveolar and interstitial spaces of the lungs, particularly in the dependent lobes. This results in alveolar wall thickening, edema, and sometimes necrosis. Small hemorrhages may occur, accounting for the hemoptysis sometimes seen clinically. Mucus plugging of the brochioles contributes to the dyspnea and cyanosis often observed in severe disease. The epithelial lining of the bronchial tree typically remains intact. Liver, spleen, and hilar node enlargement may occur, and focal splenic and hepatic necrosis can result. In fatal cases, meningeal, pericardial, brain and adrenal changes have been noted upon histologic examination of tissues collected at autopsy.

Although the lung is the organ most prominently involved, *C. psittaci* causes a systemic illness in man. The course of clinical disease, however, varies widely. The infection generally manifests itself with sudden onset of chills and high fever (38°C to 40.5°C), but may present insidiously with slowly increasing malaise and fever over a 2- to 4-day period. Severe headache is often the patient's major complaint. Arthralgias and painful myalgias (especially in the back of the neck), anorexia, and malaise are other

common symptoms. A pale macular rash (Horder spots) has been described. Jaundice is noted only in very severe cases.

A persistent dry hacking cough, which can appear at any time during the course of infection, is a prominent feature of the disease. Small amounts of mucoid, occasionally blood-streaked sputum may be produced. Gastrointestinal symptoms are experienced by only a minority of patients and, when present, include diarrhea, nausea, vomiting, and abdominal bloating. Other uncommon complaints include epistaxis and mild sore throat. Changes in mental status, associated with hypoxia, may be observed toward the end of the first week of illness and occasionally progress to lethargy, delirium, and even coma.

Epidemiology

Virtually any species of bird can serve as a host to *C. psittaci*. As a result, psittacosis is truly a cosmopolitan disease. Although psittacine birds remain the major reservoir for human infection, cases associated with exposure to infected ducks, cockatiels, sparrows, canaries, pigeons, and occasionally, various mammalian species have all been reported. Many cases in recent years have occurred in persons processing carcasses of domestic fowl, especially turkeys (69). Usually a sporadic disease, approximately one-half of all cases occur in pet bird owners. The disease is considered an occupational hazard for pet shop employees, zoo workers, veterinarians, and others who capture or raise psittacine birds for sale and profit.

A dramatic drop in the incidence of cases of psittacosis in the United States has been noted over the past few decades. Factors contributing to this decline include the use of tetracycline-containing poultry feed and the federal requirement that all commerically imported psittacine birds be treated for 30 days with antichlamydial drugs prior to transport into the country. Although antibiotics do not totally eradicate the agent from infected birds, they greatly reduce the transmission rate of the organism. Imported and domestic birds often are housed together or in close proximity in commercial settings. Because birds rapidly become infectious under stressful and crowded conditions, the potential for the spread of the organism in such settings always remains a real possibility.

C. psittaci is found in relatively high titer in the excreta, feathers, blood, and tissues (particularly the kidneys, spleen, and liver) of infected birds. Although birds can show signs of overt illness and die from infection, more frequently they demonstrate only minor symptoms such as lethargy, diarrhea, failure to feed, and ruffled feathers. This often results in close contact with the owner investigating the cause for the bird's altered behavior. Asymptomatic birds and those that have recovered from illness often continue to shed the organism for several months.

The agent can survive in dried bird excreta for long periods of time, and aerosolization of such contaminated materials with subsequent inhalation usually results in human infection. Bird bites, mouth-to-beak intimate contact, and manipulation of contaminated tissues and feathers have also led to infection. Prolonged, close contact with infected birds is not required for acquisition of the organism. Well-documented cases in which individuals have had only momentary contact with potentially contaminated environments have been reported. This may explain why approximately 20% of cases have no known route of exposure. No documented cases exist from ingestion or handling of dressed, eviscerated poultry.

Several reported instances of direct person-to-person transmission, primarily involving hospital personnel, have been recorded. Disease in these cases has reportedly been more severe than that following exposure to infected avians. The reasons for these observations are speculative, but include theoretical alterations in strain virulence following human passage and a dose effect.

Diagnostic Procedures

Culture

Although the diagnosis of psittacosis can be confirmed by isolation of the organism, attempts to isolate *C. psittaci* from clinical or animal specimens may be hazardous to laboratory personnel, and such procedures should be offered only by facilities with specialized training in the handling of the etiologic agent under laboratory conditions. Cell culture probably is the most sensitive technique for primary isolation of *C. psittaci*. Several cell lines have been used to cultivate *C. psittaci*. L cells efficiently support the growth of the agent, most strains forming cytopathic plaques that result in additional cycles of chlamydial replication, and eventually infecting the entire monolayer. Specific identification can be made by staining infected cells with commercially available genus-specific immunofluorescent antibody reagents. *C. trachomatis* can be differentiated from *C. psittaci* with species-specific immunofluorescent antibody preparations or based on the failure of *C. psittaci* intracellular inclusions to stain with iodine.

Serology

Because of the high titers of complement-fixing antibody produced in cases of psittacosis, complement fixation tests are used by public health laboratories to confirm the diagnosis of the disease. Although a single titer of 32 or higher is consistent with a presumptive diagnosis of psittacosis, a fourfold rise in titer between acute and convalescent serum specimens collected a minimum of 3 weeks apart is interpreted as definitive evidence of *C. psittaci* infection. Antibody titers generally rise by the second week of illness. Antimicrobial therapy inititated early in the disease can delay or diminish production of complement-fixing antibodies and, as a result, decrease the sensitivity of the test (70).

Technically demanding, the accuracy of the complement fixation test is highly dependent upon the quality of available test antigen. Unfortunately, satisfactory reagents are not always available commercially. Because *C. psittaci* is a common cause of inclusion conjunctivitis in guinea pig colonies, all lots of guinea pig-derived complement must be carefully tested for the presence of interfering antibodies before being placed in service (71).

Interpretation of Laboratory Results

Routine laboratory tests usually are of little value in establishing diagnosis of psittacosis. White blood cell counts typically fall within the normal range, although an unremarkable leukocytosis or leukopenia may be observed. Occasionally, patients become anemic during the course of the disease. During acute stages of the illness, protein may be detectable in the urine. Despite the fact that hepatomegaly is common, results of liver function

tests are usually within normal limits. Analysis of cerebrospinal fluid is unrewarding.

The diagnosis, therefore, usually is dependent upon the results of serologic testing. False-positive rises in antibody are rare, but have been reported in some patients with legionellosis. Although infrequently necessary, clinical specimens may be shipped, upon approval of the state public health laboratory, to the CDC for attempts to isolate the agent.

Therapeutic Considerations

The tetracyclines are extremely effective agents for the treatment of psittacosis. Following initiation of antimicrobial therapy, clinical improvement can usually be observed within 48 to 72 hours. Relapses, however, are common unless treatment continues for 10 days to 2 weeks after defervescence. Case fatality rates ranged from 20% to 40% in the preantibiotic era, but today the prognosis is good with appropriate therapy. All suspected cases should be reported promptly to public health authorities.

Chlamydophila (Formerly Chlamydia) pneumoniae

Spectrum of Disease

C. pneumoniae was originally designated the TWAR agent, a name constructed by combining the laboratory identifiers from the first two isolates, TW-183 and AR-39. The organism generally causes a mild self-limited pneumonia in teenagers and young adults. Early symptoms including pharyngitis, with or without laryngitis, may represent the only clinical evidence of infection. Other individuals may develop bronchitis, which in turn may further progress to a pneumonia lacking any distinguishing clinical or roentgenographic characteristics. The organism is being recognized increasingly as a cause of serious respiratory infection in persons with AIDS (72) and in other immunocompromised individuals. Fever, cough, and sore throat are the most common symptoms. Young individuals with mild pneumonia typically fail to show abnormalities in white blood cell counts or leukocyte differentials. In older adults with other chronic illnesses, disease is often more severe and may result in hospitalization. Such patients are more apt to show a leukocytosis.

The erythrocyte sedimentation rate is typically elevated. Some patients have a biphasic illness where an initial episode of pharyngitis and laryngitis resolves spontaneously, but is followed in 2 to 3 weeks by bronchitis and pneumonia. Although the first isolate of *C. pneumoniae* was recovered from a conjunctival specimen, no further association with ophthalmic infection has been reported. Data from several studies suggest that the organism *C. pneumoniae* might be a factor in the pathogenesis of cardiovascular disease (73, 74).

Epidemiology

Unlike *C. psittaci*, no avian or other animal reservoirs have been identified for *C. pneumoniae*. Human-to-human transmission appears to be the only significant mechanism for spread of the agent. Although the mode of transmission has not been specifically identified, it is presumed to be via a respiratory route. *C. pneumoniae* does not appear to be sexually transmitted (75).

Antibody prevalence is generally low during the childhood years, rising rapidly in teenagers. In adults, the prevalence of *C. pneumoniae* antibodies reportedly ranges from 25% to 50%. Antibody persistance in the elderly suggests repeated infection with resulting periodic boosts in titer. The prevalence of antibody is reportedly higher in males than in females. In several outbreaks among military recruits, boarding school students, and in the general population, retrospective serologic studies have demonstrated that *C. pneumoniae* was the cause of pneumonia. Prospective studies have since incriminated *C. pneumoniae* in similar outbreaks among college students. One recent investigation indicated that *C. pneumoniae* was an important cause of hospital-acquired pneumonia in intubated patients having undergone some type of surgical procedure (76). Reported outbreaks have occurred throughout the world during all seasons of the year.

Diagnostic Procedures

Culture

HeLa-229 cells are more susceptible to infection with *C. pneumoniae* than the McCoy cells utilized in most laboratories for the isolation of *C. trachomatis*. Multiple cell passage may be required to recover the organism. Egg cultivation, however, may be necessary for the isolation of the agent from some specimens. If HeLa-229 cells are used for the isolation of *C. pneumoniae*, pretreatment of the monolayers with diethylaminoethyl-dextran increases the susceptibility of the cells to infection.

Serology

Two patterns of antibody response to *C. pneumoniae* are seen. During primary infection, complement-fixing antibodies appear early, followed in 10 days to a month by IgM antibodies measurable by microimmunofluorescence techniques. Typically, IgG-class antibodies are not detected until 6 to 8 weeks after onset of illness. In the second pattern, attributed to reinfection, complement-fixing antibody is not detectable, and IgM antibody occurs in low titer or cannot be detected at all. A rapid anamnestic IgG antibody response occurs upon reinfection. Whole inclusion immunofluorescent testing and ELISA tests utilizing broadly cross-reactive LGV antigen appear to be of limited value because of cross-reactivity with *C. trachomatis*.

Therapeutic Considerations

Few comparative antibiotic trials have been conducted to determine clinical efficacy in the treatment of *C. pneumoniae* infection. *In vitro* clarithromycin has been reported to be more active on a weight basis than other macrolides, tetracyclines, and ciprofloxacin. Limited clinical experience suggests that erythromycin may be inadequate for the treatment of serious infection, resulting in prolonged symptoms or relapse. Based on clinical experience in other types of chlamydial infections, tetracyclines currently appear to be the drugs of choice.

MYCOPLASMA

Part of "53 - Chlamydia, Mycoplasma, and Rickettsia"

Mycoplasmas, the smallest organisms capable of self-replication, were first isolated by Roux and Nocard from cattle with contagious

bovine pleuropneumonia in 1898. Subsequent isolates from other sources were termed pleuropneumonialike organisms. Members of the class Mollicutes (Latin: *mollis*, soft; *cutis*, skin), they are characterized by the absence of a cell wall and are, therefore, refractory to treatment with ß-lactam class antibiotics. All mycoplasmas require lipids and lipid precursors for plasma membrane synthesis.

Like viruses, they pass through filters that retain bacteria but, except for cell wall, possess all other cellular constituents of bacteria. Unlike cell wall-deficient bacteria (L forms), they fail to revert to bacteria when cultured under appropriate conditions.

Microbiology and Taxonomy

Despite superficial cultural similarities, the mycoplasmas comprise a heterogenous group of microorganisms differing from each other in DNA composition, metabolic reactions, nutritional requirements, antigenic composition, and host species specificity. The differentiating characteristics of the various human mycoplasmas are outlined in Table 53.3. They are divided into three taxonomic families: The Mycoplasmataceae and Spiroplasmataceae (both requiring cholesterol for growth), and the non–cholesterol-dependent Acholeplasmataceae. Included in the family Mycoplasmataceae are the *Mycoplasma* species, which do not hydrolyze urea, and the genus *Ureaplasma*, strains of which produce abundant quantities of urease.

TABLE 53.3. CHARACTERISTICS OF HUMAN MYCOPLASMAS

Species	Frequency of Isolation	Usual Site of Isolation	Nature of Disease	Metabolic Activity			Optimum pH	Growth-Enhanced in Gaseous Environment			Days to Detection of Colonial Growth
				Arginine	Urea	Glucose		Aerobic	Anaerobic	CO_2	
M. orale	Frequent	Oropharynx	Unknown	+	0	0	7.0	+	+ +	LD[a]	3–10
M. salivarium	Frequent	Oropharynx	Unknown	+	0	0	6.0–7.0	+	+ +	+	2–5
M. buccale	Rare	Oropharynx	Unknown	+	0	0	7.0	+	+ +	LD	3–10
M. faucium	Rare	Oropharynx	Unknown	+	0	0	7.0	+	+ +	LD	3–10
M. lipophilium	Rare	Oropharynx	Unknown	+	0	0	7.3	LD	LD	LD	LD
Acholeplasma laidlawii	Extremely rare	Oropharynx	Unknown	0	0	+	6.0–8.0	+ +	+ +	+ +	1–5
M. fermantans	Infrequent	Genitourinary tract	Unknown	+	0	+	7.3	+	+ +	LD	3–20
M. primatum	Rare	Genitourinary tract	Unknown	+	0	0	LD	LD	LD	LD	LD
M. genitalium	Rare	Genitourinary tract	Nongonococcal urethritis	0	0	+	7.3	LD	0	+ +	Slow
M. hominis	Frequent	Genitourinary tract	Septicemia, abscesses, arthritis, endometritis, other reproductive disorders	+	0	0	5.5–8.0	+ +	+ +	+ +	1–4
Ureaplasma urealyticum	Frequent	Genitourinary tract	Nongonococcal urethritis, arthritis, other reproductive disorders	0	+	0	5.5–6.5	+ +	+ +	+ +	1–4
M. pneumoniae	Frequent	Oropharnyx, respiratory tract	Tracheobronchitis, pneumonia	0	0	+	6.5–7.5	+ +	0	+ +	3–20

There are 51 recognized species in the genus *Mycoplasma* and two, *Ureaplasma urealyticum* and *Ureaplasma diversum*, in the genus *Ureaplasma*. Ten Mycoplasma species have been detected in the clinical laboratory. *Mycoplasma pneumoniae* is the cause of primary atypical pneumonia. *Mycoplasma salivarium* and *Mycoplasma orale* are common inhabitants of the oropharynx, particularly in individuals with poor oral hygiene, whereas *Mycoplasma buccale*, *Mycoplasma faucium*, *Mycoplasma primatum*, and *Mycoplasma lipophilum* are less frequent oropharyngeal colonizers. *Mycoplasma fermentans* is a rare isolate from genital tract and oropharyngeal specimens. *Mycoplasma hominis*, generally considered a genital tract mycoplasma, also can cause infection at other body sites. *Mycoplasma genitalium*, which may cause some cases of nonspecific urethritis, and *U. urealyticum* also are isolated from genital tract specimens. An additional, probably pathogenic, mycoplasma has been identified recently and is tentatively termed *Mycoplasma incognitos* (77).

Mycoplasmas are ubiquitous in nature as saprophytes and/or parasites of plants and animals. *Mycoplasma* species also cause contamination of cell cultures used for research and diagnostic purposes, interfering with their metabolic processes and slowing their rate of growth. The spiroplasmas that infect insects and plants have a helical structure, whereas most other mycoplasmas are spherical. The morphology of individual organisms, however, varies based on differences in growth conditions and the stage of the cell cycle. Some mycoplasmas form branching filaments 0.3 to 0.4 µm in diameter and up to 1,500 µm long, which divide to form new spherical bodies. Most mycoplasmas are facultative anaerobes. Strains isolated from cattle and sheep rumen that grow only under strict anaerobiosis, however, are termed anaeroplasmas. Most mycoplasmas divide by binary fission. The genomes of ureaplasmas and mycoplasmas approximate 500 kd, whereas those of the acholeplasmas and the spiroplasmas are about 1000 kd. The morphology of *M. pneumoniae*,

M. genitalium, *M. pulmonis*, and *M. gallisepticum* is characterized by specialized structures at one or both poles of the organisms. *M. genitalium* and *M. pneumoniae* have a tapered, filamentous tip containing a dense, central rodlike core.

Specialized structures also have been observed in strains of *M. pulmonis*, and pear-shaped polar blebs have been reported in *M. gallisepticum*. These structures may assist the organisms in attachment to the respiratory or genital mucosa, as they usually represent the points of intimate contact by the mycoplasmas with the membranes of the host cells. They also may be involved in locomotion because these four mycoplasmas demonstrate a gliding motility in which the organisms move tip first.

Mycoplasmal cells do not Gram-stain, but may be stained, albeit poorly, with a Giemsa technique. Because of their plasticity, mycoplasmas all have a tendency to grow down into solid media. This may result in the production of colonies with a characteristic "fried egg" appearance (an opaque central zone of growth in the agar surrounded by a translucent peripheral surface zone). These "typical" colonies, however, do not always develop. Colonial morphology is a function of the medium constituents and degree of hydration, and atmospheric conditions. Mycoplasma colonies vary in size from 50 to 600 µm and consequently are best visualized through a clear medium with a dissecting microscope. On primary isolation, colonies of *M. pneumoniae* typically appear mulberrylike without translucent peripheral growth, although larger colonies with the characteristic zone of peripheral growth do sometimes develop. Ureaplasmas usually produce small colonies (15 to 30 µm in diameter), but on media buffered to pH 6.0, colonies up to 300 µm in size often develop. Because of the tiny colonies characteristically produced in culture, ureaplasmas were originally designated T strains or T mycoplasmas.

Mycoplasmas are highly fastidious organisms, requiring enriched media, typically containing peptones, yeast extract (supplying preformed nucleotide precursors), and animal sera (10% to 20%). The serum provides a complex of lipoprotein and cholesterol, which is incorporated into the lipid membrane of sterol-requiring mycoplasmas. Because acholeplasmas synthesize saturated long-chained fatty acids and carotenoids, which substitute for cholesterol in the plasma membrane, they have no need for exogenous cholesterol. Multiplication rates for mycoplasmas are generally lower than those for bacteria, ranging from as short as 1 hour to as long as 6 to 9 hours (*M. pneumoniae*). The microbial yield from broth cultures is low (1 to 20 mg of protein/L). Cells are visualized best in broth culture using phase-contrast or dark-field microscopy. *Mycoplasma* species usually utilize either glucose or arginine as a major energy source, although some strains are capable of using both. The carbohydrate-metabolizing species break down glucose by glycolytic pathways, mainly to lactic acid. In most mycoplasmas, the respiratory pathways are flavin-terminated, and heme compounds, catalase, and cytochromes therefore are absent. Species that metabolize arginine use a three-enzyme system to convert it to ammonia via ornithine, the process providing the organism with ATP. By comparison, the ureaplasmas utilize neither glucose nor arginine, but convert urea to ammonia by means of urease, although their requirement for urea as an energy source has not been proven.

Mycoplasma pneumoniae

Spectrum of Disease

M. pneumoniae (once termed the Eaton agent) is a human pathogen that primarily causes acute respiratory disease including pneumonia. Tracheobronchitis is the most common clinical syndrome resulting from *M. pneumoniae* infection. The organism also accounts for up to 10% of radiographically proven cases of pneumonia. Serious disease is seen principally in older children and young adults, whereas milder disease typically occurs in infants and young children. Although it occurs sporadically in all seasons and in epidemics, lasting months to years, some reports indicate that infection is more common in the fall and winter months. Following exposure, the incubation period to illness is approximately 3 weeks, with the majority of cases occurring within 15 to 25 days. Onset of disease is typically insidious, and constitutional symptoms including headache, fever, and malaise are usually noted 2 to 4 days before disease localization. Pneumonia, pharyngitis, and tracheobronchitis all clearly are caused by infection with this organism. A variety of complications associated with these syndromes also have been described. It has been further suggested that even in the absence of respiratory disease, some cases of otitis media, erythema multiforme and other skin diseases, pericarditis and myocarditis, and neurological disease may result from *M. pneumoniae* infection. Clinical manifestations attributed to *M. pneumoniae* infection are outlined in Table 53.4.

TABLE 53.4. CLINICAL ILLNESS ATTRIBUTED TO *MYCOPLASMA PNEUMONIAE* INFECTION

	Primary illness		Complications	
Definite Causation	**Possible Association**	**Reported Association**	**Definite or Probable Association**	**Reported**
Pneumonia	Otitis media	Arthritis	Pleural effusion	Pleuritis
Tracheobronchitis	Myringitis	Pancreatitis	Atelectasis	Pneumatocele
Bronchiolitis	Erythema multiforme (major and minor)	Fever of unknown origin	Respiratory insufficiency	Lung abscess
Pharingitis			Relapse	Bronchiectasis
Rhinitis	Myocarditis		Sinusitis	Bronchiolitis obliterans
	Pericarditis		Otitis media	Bacterial superinfection
	Meningoencephalitis		Myringitis	Cavitary disease
	Neuritis		Maculopapular rash	Nephritis
	Meningitis		Vesicular rash	Hepatitis
			Urticaria	Pancreatitis
			Erythema nodosum	lgA nephropathy
			Erythema multiforme	Pediatric nephrotic syndrome
			Intravascular erythrocytic hemolysis	Thrombocytopenic purpura
				Pelger-Huet anomaly
			Intravascular coagulation	Hemophagocytosis
			Raynaud's syndrome	Transverse myelitis
			Myocarditis	Leukencephalitis
			Pericarditis	Optic neuritis
			Arthritis	Cerebrovascular accident
			Meningoencephalitis	Hearing loss
			Neuritis	Polymyositis
			Meningitis	Pediatric benign myositis
			Cerebellar ataxia	
			Changes in mentation	

Although in cases of tracheobronchitis, tracheitis with frequent bouts of paroxysmal coughing and substernal pain may be conspicuous, basic bronchitis with frequent cough is more commonly noted. Respiratory illness in infants or older children accompanied by wheezing suggests bronchiolitis. Bronchitic patients tend to have fewer and milder systemic symptoms, a shorter duration of cough and fever, and a lower frequency of high-titered serum cold agglutinins than patients with pneumonia.

Pharyngitis often accompanies lower respiratory tract infection with *M. pneumoniae*, but it may be the predominant symptom in the disease (78). In such cases, symptoms usually are indistinguishable from pharyngitis caused by other bacteria or by viruses. The onset is usually insidious with headache and fever, but sore throat appears early and becomes the chief complaint. Coryza and cough also may be present. The posterior pharynx usually is diffusely erythematous, and tonsillar and pharyngeal exudates may be observed. Tender anterior cervical nodes usually are noted.

Isolation of *M. pneumoniae* from respiratory secretions in such cases does not necessarily establish it as the cause of the pharyngitis. Two studies in children have shown that *M. pneumoniae* may be isolated as frequently from asymptomatic patients as from those with pharyngitis and that a rise in antibody titer only rarely correlates with organism isolation and/or clinical pharyngitis (78, 80).

Rhinitis is reported to be the major clinical finding in children less than 1 year of age with *M. pneumoniae* infection. Isolated cases of bullous myringitis, with associated severe ear pain, may complicate *M. pneumoniae* respiratory tract infection, but the previously held theory that most isolated cases of bullous myringitis are caused by this organism is probably invalid.

Lower respiratory illness caused by *M. pneumoniae*, once called “walking pneumonia,” usually is mild. Hospitalization generally is not required; only about 2% of children and 10% of adults with *M. pneumoniae* illness are admitted to hospitals. Prodromal headache is more common in teenagers and adults than in children. Before cough appears, headache, fever, and malaise increase in severity over a 2- to 4-day period. A cough (nonproductive) then becomes the major manifestation of illness, and its absence makes the diagnosis of *M. pneumoniae* infection unlikely. Diffuse or substernal soreness in the chest that increases on inspiration is common, but pleuritic pain is rare. Mild sore throat, laryngitis, coryza, and earache may be seen. Fever is sustained unless altered by antipyretics, and maximum temperature may reach 104°F, but is typically less than 102°F. Shaking chills are uncommon. Tender anterior cervical nodes are often reported. Muscle tenderness is common, and nausea, vomiting, anorexia, and arthralgias may occur. Skin rash develops in approximately 15% of cases a few days after onset. Usually maculopapular in nature, it also may be vesicular or bullous. Clinically inapparent sinusitis is common.

Although the course of pneumonia is variable, recovery is the rule. Death has been reported, but is rare. In untreated cases, fever lasts from 2 days to 2 weeks. Slow, progressive, symptomatic improvement is typical. Cough, malaise, and radiographic abnormalities frequently persist for 2 to 6 weeks. The most common complications of *M. pneumoniae* infection are pulmonary in nature. Extensive lung involvement with lobar consolidation may result in respiratory insufficiency. In mild cases, significant pleural effusion and transient atelectasis may occur. In patients with sickle cell disease, severe illness with significant pleural effusion is common.

Clinical relapse occurs within 2 to 3 weeks of the initial illness in approximately 10% of all cases and may be accompanied by a radiographically confirmed infiltrate in the lung segment initially involved or in another segment of the same or opposite lung. Residual pleural abnormality and subsequent pneumatocele development, lung abscess, cavitation, bronchiolitis obliterans, secondary bacterial infection, and bronchiectasis all have been reported following pneumonia caused by *M. pneumoniae*, although the role of the organism in these sequelae is unclear. Myringitis may be seen, and secondary bacterial ear infections in children are not uncommon.

Illness caused by *M. pneumoniae* may be complicated by urticaria, erythema nodosum, or more commonly by erythema multiforme or Stevens-Johnson syndrome. Erythema multiforme major may be primarily caused by *M. pneumoniae* and, although a history of preceding illness usually exists, it can occur in the absence of associated respiratory disease. The organism has been cultured from vesicular fluid of these patients. Significant intravascular hemolysis can occur in association with high-titer cold agglutinins, but clinically inapparent hemolysis is probably more common. In patients with clinical hemolysis cold agglutinin titers in excess of 1:500 typically are seen and hemagglutinins also are usually demonstrable at 37°C. Intravascular hemolysis usually occurs late in the illness and may coincide with return of the body temperature to normal. Other occasionally reported hematologic complications include Raynaud's phenomenon and disseminated intravascular coagulation.

Neurological complications of *M. pneumoniae* infection have been reported in up to 10% of hospitalized patients (81, 82). The most commonly described syndromes are meningitis, encephalitis, meningoencephalitis, and mono- or polyneuritis. Less commonly reported complications include cerebrovascular accident, “toxic” psychosis, cerebellar ataxia, transverse myelitis, and

several types of mononeuropathy including cranial nerve palsy, phrenic nerve paralysis, and sudden deafness. *M. pneumoniae* has been reported to be the possible cause of up to 6% to 8% of all cases of aseptic meningitis (83) and many cases of meningitis, neuritis, and meningoencephalitis attributed to *M. pneumoniae* are not associated with antecedent respiratory illness. Although the course of such infections has not been well characterized, residual neurological deficits and death have been described, and *M. pneumoniae* has been isolated from the cerebrospinal fluid (CSF). In cases of meningitis and meningoencephalitis, various types of leukocytic CSF responses have been observed, but mononuclear cells usually predominate. The CSF protein level is usually elevated, and the glucose concentration is normal or low.

Reported cardiac complications range from clinically inapparent changes in electrocardiograms to significant peri- and myocarditis with prominent pericardial effusion. Marked arthralgia is common, but overt arthritis is rare. Other reported complications include juvenile benign myositis, polymyositis, hepatitis, thrombocytopenic purpura, hemophagocytosis, Pelger-Huët abnormalities, pancreatitis, nephritis, IgA nephropathy, and nephrotic syndrome in children (84).

Epidemiology

M. pneumoniae is a cosmopolitan cause of respiratory disease, but epidemics have been best documented in temperate climates. In large urban areas, the organism is endemic, and infection and disease occur year-round. During the summer months, it may cause up to 50% of all pneumonias. Some evidence suggests that epidemics occur at 4- to 8-year intervals both in military and civilian populations and, during such periods, disease is most prevalent in the late summer and early fall months. Typical epidemics start in the fall, are slow in onset, and may persist in the community for as long as 2 years. Infection during epidemic periods is three to five times more frequent than in nonepidemic years, with overt disease being most common in older children and young adults. Smaller outbreaks, however, can be observed at any time. The proportion of all pneumonias attributable to *M. pneumoniae* varies with the population studied. Frequencies have ranged from 4% in an indigent civilian population to 44% in a population of military personnel. In a large community study in Seattle, Washington, approximately 15% of all documented pneumonias were attributed to *M. pneumoniae* (85). This is probably representative of communities throughout the developed world. The relatively mild nature of typical *M. pneumoniae* infection is underscored by the fact that less than 5% of resultant pneumonia cases are severe enough to require hospitalization.

M. pneumoniae infection rates are highest in school-age children and young adults. It is the most common etiologic agent of pneumonia in children 5 to 15 years of age. In the Seattle survey, this organism was responsible for 30% to 60% of all pneumonias in the 5- to 20-year-old age group. Although pneumonia attributable to *M. pneumoniae* is less common in adults, it remains a significant pulmonary pathogen throughout life. In addition, *M. pneumoniae* is reported to be the most common cause of tracheobronchitis and lower respiratory disease, with wheezing in children 5 to 15 years of age (86). Pneumonia appears to be somewhat more common and illness more severe in young males than in young females. Higher infection rates are seen, however, in women 30 to 39 years old than in men of comparable age, probably reflecting increased exposure to school-aged children. Although data are somewhat conflicting, both day care center and family studies have shown that symptomatic *M. pneumoniae* infection is also common in children less than 5 years of age, although the disease is typically mild (87,88 and 89). Despite common complaints of coryza and wheezing, most children with *M. pneumoniae* infection are afebrile.

M. pneumoniae transmission between school-aged children is frequently followed by introduction into and slow but extensive spread within families. The limiting factor determining whether an exposed family member will acquire the organism appears to be immunity resulting from previous infection. Transmission in institutionalized and school populations is generally slow, requiring close contact with an infected individual. Spread between playmates appears to be more efficient than among classmates in a schoolroom setting. High transmission rates also have been observed in college populations and in military recruits where close contact in dormitories and barracks appears to play an important role (90,91,92 and 93). Transmission by individuals with asymptomatic infection has not been documented. Because of the evidently low communicability, the requirement for close contact for efficient transmission, and the relatively long incubation period, the spread of *M. pneumoniae* in the community is slow, and infections appear endemic in nature. Epidemics, however, have occurred, and point-source outbreaks lacking characteristics of prolonged close contact have been reported, implying that small-particle aerosol transmission may occur. Risk of clinical disease from *M. pneumoniae* infection is low among children less than 5 years of age, high for those in the 5 to 20-year-old range, and again, lower among adults over 20 years of age. When it occurs, though, illness usually is more severe in middle-aged and elderly individuals.

Diagnostic Procedures

Specimen Selection/Collection

M. pneumoniae infects the entire respiratory tract and, as a result, may be isolated from throat swabs or washings, nasopharyngeal swabs, tracheal aspirates, lung biopsies, and sputum. Since the organisms attach to epithelial cells of the respiratory mucosa, specimens should contain cellular material for optimal yield. *M. pneumoniae* has, on rare occasions, also been isolated from nonrespiratory sites including blood, CSF, and internal abdominal organs. Because the natural course of infection lasts for 6 to 8 weeks, the collection time is not critical, although specimen collection prior to initiation of antimicrobial therapy should be accomplished whenever possible. Because *M. pneumoniae* is relatively fastidious and very susceptible to drying, direct specimen inoculation onto growth media as soon as possible after acquisition of the specimen is strongly encouraged. Vials of liquid SP-4 medium or trypticase soy broth with 0.5% bovine albumin may be inoculated at the patient's bedside with swab specimens and transported expeditiously to the laboratory at room temperature. Antibiotics, such as penicillin, usually are incorporated in transport media to suppress bacterial overgrowth. Sputum, throat washings and other liquid specimens, and tissue samples can be

transported to the microbiology laboratory in sterile leak-resistant containers, preferably on ice, for processing. If delay in processing (more than 24 hours) is anticipated, specimens should be frozen at -70°C or below. Organisms survive indefinitely under these storage conditions. Likewise, if specimens are to be transported over long distances, shipment on dry ice is recommended. Containers placed on dry ice should be tightly sealed to prevent absorption of CO_2 by specimens. Short-term storage of specimens at -20°C is acceptable, but longer-term storage at this temperature (more than a few weeks), results in loss of *M. pneumoniae* viability. Provided the medium contains protein, mycoplasmas are unique in their ability to withstand repeated episodes of freezing and thawing.

Currently, there are no useful staining techniques for the direct microscopic detection of *M. pneumoniae* in clinical specimens. As a result, laboratory diagnosis of mycoplasmal infection usually depends upon recovery of the organism in culture, detection of its presence using genetic probes, or on the results of serologic assays.

Culture Techniques

Media for the isolation of mycoplasmas usually contain a base of soybean or beef protein supplemented with serum (10% to 20%) as a source of mandatory sterols and yeast extract to provide peptides and other growth factors including preformed nucleic acid precursors. Because *M. pneumoniae* replicates more slowly than most bacteria and many fungi commonly encountered in clinical specimens, antimicrobial agents are generally added to make the media selective, preventing overgrowth by more rapidly growing microorganisms. Exploiting the fact that mycoplasmas lack the cell walls characteristic of most bacteria and therefore are intrinsically resistant to cell-wall active agents, ß-lactam antibiotics such as penicillin and polymyxins (typically active against most Gram-negative species) are commonly employed for the inhibition of contaminating bacteria. Thallium acetate also is used sometimes to inhibit the growth of contaminating bacteria. Because it interferes with the growth of *U. urealyticum* and *M. genitalium*, it should be avoided when attempting to isolate these species. Care should be taken when handling this compound because it is also toxic to humans. Likewise, the addition of antifungal agents such as amphotericin B is useful in preventing overgrowth by contaminating yeasts. Methylene blue also is utilized sometimes as a medium supplement to inhibit other species of mycoplasma.

Because the *Mycoplasmatales* represent a heterogeneous group of organisms, no single medium will provide optimal conditions to accomplish the isolation of all species. A pH of 6.0 is required by *U. urealyticum* for optimal replication, and little growth is observed above pH 7.0. By contrast, most other species grow well in media with a 7.0 pH. *M. pneumoniae* is the most pH-tolerant species, growing in media with a pH range from 5.5 to 8.0. For pH to be monitored visually, inclusion in the medium of a pH indicator, such as phenol red, is recommended. The incubation atmosphere is also an important growth factor. *M. genitalium* and *M. pneumoniae*, both obligate aerobes, are stimulated by atmospheric CO_2. *M. salivarium*, *M. orale*, *M. buccale*, and *M. faucium* grow best under anaerobic conditions and grow poorly on H agar aerobically. *M. fermentans*, *U. urealyticum*, and *M. hominis* appear to be minimally affected by incubation atmosphere.

Quality control testing should be performed on all prepared media to demonstrate that optimal growth of mycoplasmas can be achieved. For this purpose, it is highly recommended that laboratories use recent clinical isolates of all species being sought in culture. Stock strains adapted to culture conditions often grow readily on media that fail to support the growth of organisms from clinical specimens. Artifacts also are a major problem in the detection of mycoplasmas in clinical materials because debris and host cells can resemble fried-egg colonies. It is therefore imperative that all laboratories prove that observed colonies represent transferable entities with specific properties before issuing a report indicating that a mycoplasma has been isolated from a clinical specimen. Some lots of serum induce the formation of pseudocolonies, whorls of magnesium and calcium soap crystals appearing on the surface of agar plates. Although transferable, these crystals can be differentiated from true colonies of mycoplasmas because they fail to produce metabolic products.

To isolate *M. pneumoniae*, a 0.1-mL volume of transport medium in which a swab has been extracted should be simultaneously inoculated onto a solid medium, such as H agar or SP-4 medium, and into a biphasic broth medium. Isolation of the agent is best accomplished from sputum or tracheal aspirate, but nasopharyngeal and oropharyngeal swab specimens also often prove fruitful. Body fluids, sputa, and disrupted tissue specimens should be diluted 1:10 and 1:100 before media inoculation to lower the concentration of inhibitory substances commonly found in tissue samples. Agar plates should be incubated aerobically at 35°C to 37°C in a sealed container and examined for the presence of characteristic colonies with a stereoscopic microscope under oblique light at 20× to 40× total magnification after 2, 10, 15, and 21 days' incubation. This provides the best working distance and resolution for observation of colonies on an agar surface through the bottom of an unopened plate. This technique prevents contamination of the media surface with microorganisms, including mycoplasmas, during the multiple-plate examinations. Biphasic cultures should be observed for a pH decrease as indicated by a color change of the phenol red indicator from red to yellow-green from acid production by *M. pneumoniae* resulting from glucose fermentation. They also should be examined microscopically by viewing the broth through the side of the container for the presence of spherules, fluid medium colonies appearing as early as 5 days after medium inoculation. *M. pneumoniae* does not produce turbidity in liquid media, and cloudy culture medium is usually indicative of bacterial or fungal contamination. Bacterial overgrowth also sometimes will result in an acidic pH change in the medium, but typically within the first couple of days of incubation. A color change resulting from *M. pneumoniae* in the biphasic system tends to occur 1 to 2 days after the appearance of characteristic colonies on agar plates. Both systems require a minimum of 5 days of incubation and sometimes as long as 14 to 21 days for growth detection. Because mycoplasmas grow down into agar media, subculture of colonies to a broth medium is best accomplished by excising the agar around a colony and transferring it to broth. For subculture to an agar medium, colonies are first emulsified in broth and then transferred to a plate. Broth from biphasic systems should be subcultured to solid media after 21 days' incubation, and these cultures should be observed for an additional 21 days. The

majority of positive specimens will yield typical small colonies on agar and spherules in broth as well as cause a detectable pH change within 10 to 12 days. Some strains may require up to 30 days for primary isolation.

Recovery of *M. pneumoniae* in culture has little immediate clinical diagnostic value, because of the length of time required for primary isolation. It is useful, though, for epidemiologic purposes and in outbreaks of infection in settings where close physical contact has occurred between possible patients. The recovery of *M. pneumoniae* in culture, however, is always significant clinically because the organism is never part of the “normal” flora.

Identification Methods

A presumptive positive report for *M. pneumoniae* can be issued if an organism has been recovered that grows slowly on agar, forms spherules in broth, and ferments glucose. With the exception of *U. urealyticum*, which can be identified with special strains, definitive differentiation of species cannot be based on colonial morphology. A potential confounding factor, albeit uncommon, is the isolation of cell wall-deficient bacteria, which, in many cases, resemble mycoplasmas.

Ordinarily forming large fried egglike colonies, these organisms transfer poorly and usually revert to typical bacterial morphologies on ß-lactam-free media. Confirmation is accomplished using a test for hemolysis (94). To perform the hemolysis test, suspect colonies are overlaid with a thin layer of 8% guinea pig red blood cells in saline agar, incubated overnight at 37°C, and examined for clear zones of ß-hemolysis surrounding individual or groups of colonies.

M. pneumoniae can be differentiated from the only other two mycoplasmas isolated from respiratory specimens under recommended culture conditions based on colony size and growth rate. *M. hominis* forms larger colonies and grows more rapidly than *M. pneumoniae*. *Acholeplasma laidlawii*, which hemolyzes guinea pig erythrocytes, also grows more rapidly than *M. pneumoniae* and forms large colonies within several days of subculture. Unequivocal identification of an isolate as *M. pneumoniae* can be accomplished using growth inhibition assays with specific antibody. Aliquots of serial 10-fold dilutions of broth cultures are spread uniformly over H agar plates and allowed to adsorb. Paper discs soaked with specific antisera then are placed on the agar plates which are incubated for 4 to 6 days and examined for zones of growth inhibition.

Serology

There is considerable confusion and disagreement in the literature and among physicians regarding the clinical utility of the serum cold agglutinin assay for the diagnosis of *M. pneumoniae* infection. A simple and inexpensive procedure, it appears to be of some value in serious lower respiratory tract infections, when more specific assays are not available. It is of particular value in hospitalized patients because it is positive early enough in the course of disease to assist the clinician in the choice of appropriate antimicrobial therapy and more frequently is positive in patients seriously ill with *M. pneumoniae* infection. The presence of cold agglutinins, however, is not specific for *M. pneumoniae* infection (95). Serum cold agglutinins at a titer of 1:32 or greater have been reported in 50% to 90% of patients with confirmed *M. pneumoniae* lower pulmonary tract infections. In general, the intensity of the cold agglutinin response directly correlates with the severity of pulmonary involvement. Although elevated cold agglutinin titers have been reported in 18% of cases of adenovirus pneumonia in a military population (96), as a rule, the higher the cold agglutinin titer, the more likely a particular respiratory infection is caused by *M. pneumoniae*. Only a fraction of cold hemagglutinin-positive nonbacterial pneumonias are caused by *M. pneumoniae*, but the majority of infections in patients with titers of 1:128 or higher are caused by this organism.

A useful rapid screening “bedside” test for the presence of serum cold agglutinins may be performed by adding blood to a tube containing an equal volume of an anticoagulant such as sodium citrate. The tube then is placed in ice water (0°C to 4°C) for 15 to 30 seconds, removed, and immediately examined for coarse agglutination by tilting the tube on its side. On warming, the agglutination should disappear and should form again upon recooling. A positive screening test correlates with a standard test result of 1:64 or above. In the standard cold agglutinin test, blood should be allowed to clot at 22°C to 37°C before removal of the serum. The test is then performed by making a serial dilution of the patient's serum, mixing each dilution with a suspension of standardized human group O erythrocytes, and incubating the test overnight at 4°C. IgM-class antibodies directed against *M. pneumoniae* react with the I antigen on the red blood cells membrane at 4°C. Hemagglutination disappears on warming to 37°C. Cold agglutinins appear during the first and second week and disappear by the sixth week of illness. A single titer of 1:128 or higher generally is considered significant. Because cold agglutinins appear in only about half of all patients with proven *M. pneumoniae* infection, a negative result is of little value.

Like many other infectious diseases, the specific serologic diagnosis of illness caused by *M. pneumoniae* can be based on a significant change in IgG antibody titer between timed acute and convalescent serum samples. The acute phase specimen should be collected as early as possible in the illness, and the convalescent sample should be collected 2 to 4 weeks later. The two specimens should be tested concomitantly. Complement fixation (CF), ELISA, and immunofluorescent assays are commonly employed for this purpose. A fourfold increase in antibody titer, demonstrable in 50% to 85% of patients tested with immunofluorescent or CF techniques, is diagnostic of acute *M. pneumoniae* infection. CF antibodies, typically develop during the second and third weeks of illness and persist for 6 to 12 months.

Because CF antibodies often are relatively short-lived, a fourfold decrease in titer can also be helpful in establishing etiologic significance in a specific patient. Although CF titers of 1:256 or higher usually are indicative of recent infection, unequivocal causation of infection with *M. pneumoniae* should not be assumed under such conditions. Infection with *M. pneumoniae* is typically associated with a prolonged incubation time. As a result, antibody development has often already begun at the time of clinical presentation. Fourfold changes in titer can thus occur over a very short period of time (in as few as 5 days), and serum samples collected 5 to 7 days apart will often demonstrate a significant rise in antibody titer. False-positive results, however, have been reported in patients with *M. genitalium* infection, patients with documented bacterial meningitis (97), and in some

cases of acute pancreatitis. Reports based on CF tests that implicate *M. pneumoniae* in various neurological syndromes must, therefore, be interpreted with caution as they may be the result of these cross-reacting antibodies.

Both immunofluorescent and ELISA techniques, comparable in sensitivity and specificity to CF assays, offer the advantage of permitting separate detection of IgG and IgM class antibodies. ELISA techniques using solubilized whole organisms and unfractionated extracts lack sensitivity because the active components are represented in such small quantities in the antigenic mixture that they are undetectable. By comparison, the use of purified adhesion P1 protein in ELISA system appears to offer acceptable levels of sensitivity while increasing the specificity of the test significantly over techniques using whole lysed *M. pneumoniae* organisms.

Interpretation of Diagnostic Test Results

Roentgenographic chest examination typically reveals unilateral infiltrates that appear diffusely reticulonodular or interstitial. Bilateral involvement, however, occurs in about 25% of patients. Although small pleural effusions are noted in approximately one fourth of cases, significant pleural effusion is rare. Peripheral white blood cell counts are typically within normal range, but may be elevated. Leukocyte differentials usually demonstrate 60% to 85% polymorphonuclear leukocytes with occasional bands. Erythrocyte sedimentation rates are typically strikingly elevated. Gram-stained sputum specimens usually reveal large numbers of leukocytes, but no predominant bacterium. Usual upper respiratory flora generally is reported on sputum culture. Standard blood cultures are negative.

In the setting of a compatible clinical syndrome, isolation of *M. pneumoniae* in culture should be considered diagnostic. Because carriage of *M. pneumoniae* often can be demonstrated for 6 to 8 weeks after recovery from illness, the potential for confusion does exist in cases of subsequent but unrelated illness. Under optimal conditions, only 85% of patients demonstrating a four-fold rise in antibody titer to *M. pneumoniae* are typically culture-positive. The precise value of a negative probe or culture, therefore, is not known.

Therapeutic Considerations

M. pneumoniae is susceptible *in vitro* to many broad-spectrum antimicrobial agents including tetracyclines, macrolides, chloramphenicol and the aminoglycosides but are resistant to inhibitors of cell wall syntheseis such as the ß-lactams, because they lack typical bacterial cell walls. Although *in vitro* susceptibility testing of *M. pneumoniae* usually is not performed and standardized procedures for such testing have not been developed, erythromycin-resistant strains have been reported. Untreated, illness caused by *M. pneumoniae* typically resolves in 2 to 3 weeks, but symptoms resolve more rapidly with appropriate antimicrobial therapy. Recommended antibiotics for treatment of pneumonia caused by *M. pneumoniae* in adults are tetracycline and its analogues as well as erythromycin. Tetracyclines must not be given to children or pregnant women. Use of other therapeutic modalities including antitussives, antipyretics, assisted ventilation, and oxygen therapy should be based on assessment of individual patient needs. Because illness is clinically indistinguishable from that caused by any number of viruses, routine treatment of individuals with pharyngitis and tracheobronchitis is not recommended unless a specific diagnosis has been made based on laboratory findings. If, however, prolonged illness is observed and *M. pneumoniae* has been established as the etiologic agent, either tetracycline or erythromycin therapy may be indicated. Because antibiotic-treated individuals continue to harbor organisms in the posterior pharynx for up to 2 to 3 months, antimicrobial therapy does not appear to decrease transmission rates to close contacts.

No method has proven uniformly effective in preventing infection and disease with *M. pneumoniae*. As a result, patients probably should be isolated, and attempts should be made in home settings to discourage close contact with infected individuals.

Mycoplasma hominis and Ureaplasma urealyticum

Spectrum of Disease

Mycoplasma hominis

M. hominis colonizes the genitourinary tract (20% to 40% of women and a smaller percentage of adult males) and, occasionally, the oropharynx. It has, however, been reported in infections following facial and transpharyngeal oral surgery (98) and in cases of exudative pharyngitis (99,100 and 101). In addition, it has been isolated sometimes in pure culture, from the upper urinary tract of patients with acute pyelonephritis. In many such infections, antibody to *M. hominis* has been demonstrable in both urine and serum. Data from several studies suggest that the organism is an infrequent cause of acute pyelonephritis and acute exacerbations of chronic pyelonephritis. Likewise, *M. hominis* is frequently recovered, when appropriate culture procedures are utilized, from tubo-ovarian and pelvic abscesses as well as from inflamed fallopian tubes. In both Swedish and British studies (102), the organism was isolated directly from the fallopian tubes of approximately 10% of women with acute salpingitis using laparoscopic techniques to collect specimens. In these studies, antibody was detectable in the serum of approximately 50% of patients with salpingitis compared to 10% of healthy women. Seroconversion also was demonstrable in approximately one half of women from whom *M. hominis* was isolated from the lower genital tract. Other studies, however, have reported that women with gonococcal pelvic inflammatory disease were not more likely to demonstrate detectable serum antibodies to *M. hominis* than those without *N. gonorrhoeae*, suggesting that damage caused by other agents may be important in such antibody responses, thereby obfuscating the true importance of *M. hominis* in this disease.

M. hominis also is commonly seen in association with postpartum fever. It can be isolated from the blood of approximately 10% of febrile women following abortion but not from nonfebrile postabortal women or following normal pregnancies. Rises in antibody titer are demonstrable in approximately 50% of febrile women following abortion, whereas only a small fraction of women without fever show detectable antibody to *M.*

hominis. Bloodstream infection with *M. hominis* typically is a transient and often asymptomatic sequela of uncomplicated vaginal delivery. Immediately postparturition, *M. hominis* can be recovered for a short period of time from the blood, but in less than 10% of women tested. This compares to women with postpartum fever in whom the agent reportedly has been isolated more than 24 hours after delivery in many cases and in whom a rise in antibody titer has been demonstrated. Presumably, *M. hominis* induces postpartum fever by causing endometritis. Typically, such patients have low-grade fever for 24 to 48 hours after delivery, show minimal symptoms, and recover uneventfully without antimicrobial therapy.

Although isolated from up to 30% of patients with nongonococcal urethritis, results of several studies have failed to incriminate *M. hominis* as a cause of this disease. Some hypogammaglobulinemic patients have, however, developed a chronic urethrocystitis, which, failing to isolate other potential etiologic agents, is apparently caused by persistent ureaplasmal or *M. hominis* infection. Less frequently, *M. hominis* causes extragenital infections including meningitis, brain abscess, septicemia, and arthritis. *M. hominis* has been isolated from the synovial fluid of hypogammaglobulinemic patients with nonbacterial septic arthritis and should be considered as a potential etiologic agent of arthritis in these patients. Likewise, *M. hominis* occasionally has been isolated after childbirth from otherwise normal mothers who develop sudden-onset arthritis. Response to tetracyclines and other agents to which mycoplasmas are typically susceptible further supports the belief that *M. hominis* was responsible for this arthritis. Immunosuppression also may predispose an individual to proliferation with *M. hominis* and resultant peritonitis and septicemia have been reported.

Because neonates are also relatively hypogammaglobulinemic, they appear to be at increased risk of infection with *M. hominis*. In one report, *M. hominis* was the sole organism recovered from purulent material collected from a neonate following incision and drainage of several supraclavicular abscesses (103) and in another reported neonatal infection, *M. hominis* was the etiologic agent in a case of submandibular lymphadenitis (104). Likewise, a neonate with severe burns reportedly developed septicemia from *M. hominis*. The organism also has been isolated from amniotic fluid in a case where the infant subsequently died with respiratory distress syndrome (105). An interstitial pneumonia was observed on postmortem examination of lung tissue. *M. hominis* also is clearly the cause of some ocular infections in neonates. Other *M. hominis* infections reported include a scalp abscess complicating placement of an intrapartum monitoring device (106) and a massive pericardial effusion in a newborn (107). In cases of neonatal meningitis and brain abscess when *M. hominis* has been recovered from CSF or abscess material, the neonate probably either became infected *in utero* or became colonized at birth, subsequently developing infection. In cases of *M. hominis* meningitis, CSF examination usually reveals a pleocytosis with a predominance of neutrophils, an increased protein, and a decreased concentration of glucose. Because CSF generally is not cultured for mycoplasmas, *M. hominis* should be considered in the differential diagnosis of cases of culture-negative neonatal CNS infection.

Ureaplasma urealyticum

Recent studies suggest a fairly strong association between *U. urealyticum* and certain reproductive and genitourinary infections, including nongonococcal or nonspecific urethritis, prostatitis, and, in women, acute urethral syndrome. Having been isolated from the epididymis in association with an antibody response, recent studies also indicate that the organism may be a rare cause of acute epididymitis (108). *U. urealyticum* also may play a role in involuntary infertility, habitual spontaneous abortion and stillbirth, low birth weight, and premature delivery associated with silent chorioamnionitis The agent also has been isolated from aspirates of scarred renal tissue in patients with reflux nephropathy (109), but their role in this disease is not known. Current data do not support a role for *U. urealyticum* in cervicitis, vaginitis, postabortal or postpartum fever (although on rare occasions it has been isolated from the blood of women with low-grade postpartum temperature increases), or Bartholin gland abscess.

Nongonococcal urethritis is more common than gonococcal urethritis in most developed nations, and *U. urealyticum* accounts for 20% to 30% of such infections (110). At least two studies have shown that infants colonized with *U. urealyticum* have statistically lower mean birth weights than babies in whom colonization was not detected (111, 112). The association with low birth weight, however, was not related to a shortened gestational period. Although *U. urealyticum* was recovered significantly more often from babies whose placentas demonstrated histologically confirmed chorioamnionitis than from babies whose placentas were read as normal in one study (113), no adverse effects could be attributed to either colonization of the neonates or the placental lesions. The organism has, however, been recovered from fetal tissues following spontaneous abortion. *U. urealyticum* also has been isolated more often from neonates with pneumonia than from asymptomatic controls (21% versus 4%) and from a significant proportion (16%) of infants less than 3 months of age hospitalized with pneumonitis (114, 115). Fatal neonatal pneumonia resulting from an intrauterine infection with *U. urealyticum* also has been reported (116). Like *M. hominis*, in adult patients with hypogammaglobulinemia, *U. urealyticum* has been shown to cause invasive disease (117). *M. hominis* and *U. urealyticum* also have been the sole organisms recovered from the CSF of neonates with meningitis (118, 119) and, in prospective studies, represent the most common microbiological isolates in preterm neonates undergoing lumbar puncture for suspected sepsis or hydrocephalus (120, 121). The significance of these findings, however, is unclear because many such CNS infections resolve without specific antibiotic therapy, and the long-term sequelae following such infections have yet to be determined.

Epidemiology

The major reservoir for both *M. hominis* and *U. urealyticum* is the genital tract of adult women and men. Following puberty, colonization with genital mycoplasmas occurs primarily as a function of sexual contact, and antibodies are first detectable at this time. *U. urealyticum* has been isolated from up to 81% of adult women, while 9% to 50% are culture-positive for *M. hominis* (122).

U. urealyticum is isolated more frequently from the vagina than from the cervix, endometrium, and upper genital tract, whereas *M. hominis* is recovered more often from the cervix than from endometrium or the vagina (123). Although the incidence of colonization is lower in men than in women, from 30% to 50% of sexually active males are colonized with *U. urealyticum* (124). The majority of individuals colonized with these organisms are without associated symptoms. When infection does occur, however, it appears to be caused by a person's endogenous flora. The frequency with which *U. urealyticum* and *M. hominis* are recovered from the cervix and vagina of sexually active women implies that in these body sites they represent part of the usual bacteria flora. These two organisms reportedly have been isolated more frequently from blacks than whites but, because genital mycoplasma colonization also appears to be related to socioeconomic status, the extent to which this reflects differences in susceptibility is not clear. Genital mycoplasmas also tend to be recovered more often from pregnant than nonpregnant women. By comparison, they are isolated significantly less frequently following menopause. These findings suggest that sex hormones may have some influence on colonization with these organisms.

Colonization of infants with genital mycoplasmas generally occurs during passage through the birth canal. As with adults, most neonates have no associated clinical manifestations of infection, although *U. urealyticum* can cause a clinically silent amnionitis which, despite the lack of symptoms, can persist and result in premature delivery (125). Infants delivered vaginally are much more frequently colonized with these agents than babies delivered by cesarean section. *U. urealyticum* has been isolated from at least one body site in 30% to 62% of all infants born to women positive by cervical culture for the agent and, likewise, *M. hominis* can be recovered from 18% to 45% of babies born to women harboring the organism (126, 127). Ureaplasmas have been recovered from genital specimens of approximately one third of infant females. *M. hominis* is isolated somewhat less often from these same specimens. These organisms are recovered significantly less frequently from genital specimens of infant boys. This is probably because of the decreased exposure of their genital mucosal surfaces.

Persistence of neonatal colonization with genital mycoplasmas generally ends by 2 years of age (128), although the organisms occasionally persist to puberty in young girls. In one reported study, up to 20% of prepubertal females remained colonized with *U. urealyticum* and approximately 6% with *M. hominis*. Genital mycoplasmas have only rarely been isolated from prepubertal males outside of the neonatal period. Both organisms are found considerably more often in sexually abused children (129).

Diagnostic Procedures

Specimen Selection/Collection

With appropriate collection and culture techniques, genital mycoplasmas can often be recovered from the urine and from cervical, vaginal, or urethral swabs of asymptomatic individuals. *U. urealyticum* also has been isolated in pure culture from several usually noncolonized sites including (i) the blood of patients with low-grade temperature elevations; (ii) lung tissue of infants with neonatal pneumonia; (iii) prostatic secretions of patients with prostatitis; (iv) the upper urinary tract of patients with renal calculi and upper urinary tract disease; (v) fluid from arthritic joints of hypogammaglobulinemic individuals; and (vi) fetal membranes or tissues. *M. hominis* likewise has been isolated from several normally noncolonized body specimens including the blood of postpartum patients with fever, the CSF of neonates with signs and symptoms of meningitis and sepsis, surgical drainage, mediastinal and abdominal fluids, joint fluid of arthritic persons with hypogammaglobulinemia, bronchoalveolar lavage, throat swabs, abscesses, prostatic secretions, semen, and tissue specimens.

These organisms appear to represent usual flora in the female genital tract. As a result, care must be exercised when collecting samples from internal sites not to contaminate specimens with cervical, vaginal, or urethral secretions. Swabs should be suspended in a suitable transport medium (such as 0.5% albumin in 2 mL of soybean-casein digest broth containing 200 to 400 U.S. units of penicillin per milliliter). Transport media should contain both protein and a peptone to provide maximum protection to the cell wall-deficient mycoplasmas. A transport medium is not required for submission of fetal tissues and membranes, but such specimens should be maintained in a moist environment until culture inoculation can be accomplished. Although swabs with wooden shafts should not be used for specimen collection (plastic and aluminum are acceptable), the tip composition (cotton, rayon, or calcium alginate) does not appear to be a factor in the recovery of these organisms. Ideally, specimens should be transported to the laboratory without delay and inoculated to an appropriate culture system. Specimens may, however, be held up to 24 hours at 4°C without significant loss of organism viability. For longer delays, specimens should be frozen at -70°C.

Urine specimens should be transported to the laboratory on ice (specimen transport devices containing boric acid preservatives are not recommended) and centrifuged at 600× g to sediment epithelial cells and other debris. The pellet from the centrifuged urine specimen should be diluted 1:2 in transport medium prior to culture inoculation. Because urine protein concentrations are typically low, a urine sample submitted for recovery of mycoplasmas should not be frozen unless the specimen pellet is reconstituted with transport media containing a protein supplement. Body fluids, sputa, and tissues should be diluted 1:10 or 1:100 in transport media prior to culture inoculation to remove inhibitory substances. Specimen storage for more than several days at -20°C is not recommended because the organism is known to lose viability at this temperature.

Blood specimens for the isolation of genital mycoplasmas should be collected in a syringe without anticoagulant because *M. hominis* and *U. urealyticum* are inhibited to some degree by these compounds. Blood should be inoculated into an appropriate blood culture medium at the patient's bedside. Because *M. hominis* will replicate in most blood culture systems (but does not alter the appearance of the media), blood for the isolation of this organism can be inoculated directly into routine blood culture bottles, but must be subcultured blindly on a daily basis for 5 days to detect growth of the organism. Alternatively, a medium specifically formulated for isolation of *M. hominis* from blood (28)

may be used. Blood should be diluted at least 1:10 in medium to prevent inhibition by factors normally present in clinical specimens. In automated blood culture systems, the instrument may detect the organism's metabolism and respond with a positive growth signal (130). Although Gram stains of the media will be negative, the organism can be visualized using the acridine orange staining procedure, and the organism can be recovered upon subculture of the broth medium.

Culture Techniques

The four mycoplasmas commonly encountered in genital tract specimens have substantially different growth properties. *M. hominis* and *U. urealyticum*, the two most important clinically, grow rapidly and can be readily isolated. *M. fermentans* and *M. genitalium* are more difficult to detect in culture. Because *M. hominis* tends to overgrow *U. urealyticum* when isolated concomitantly, separate culture systems are required for the recovery of the two organisms. *M. hominis* grows over a pH range from 6 to 8, but *U. urealyticum* fails to grow at a pH greater than 7.0 and multiplies optimally in a pH range from 5.5 to 6.5. As a result, *M. hominis* can replicate in media designed for the isolation of *U. urealyticum*, but the converse is not generally true. *M. hominis* is susceptible to lincomycin, whereas *U. urealyticum* is not. Thus, inclusion of 20 to 50 g/mL of lincomycin in the medium effectively selects for *U. urealyticum*. Erythromycin has likewise been used successfully as a selective agent in media designed to recover *M. hominis*. Two media systems commonly used are U agar and U broth for the isolation of *U. urealyticum* and H agar and H broth for the recovery of *M. hominis*. Because the final pH of both H agar and H broth is greater than 7.0, they effectively select for *M. hominis*. A number of other media have been used successfully for isolation of these organisms (131,132,133 and 134).

Because of the steep death phase resulting from exhaustion of urea and the elevated pH of the medium from urease activity, *U. urealyticum* presents unique problems during isolation attempts. In both broth and solid media it is essential that the culture system be properly buffered and that a sufficient quantity of urea be supplied. After a color change in broth medium, *U. urealyticum* may remain viable for as few as 12 hours, whereas colonies survive for approximately 2 days after detection on solid media. 2-(N-morpholino) ethanesulfonic acid (MES) at a pH of 6.3 is a very effective buffer for this purpose. Molar ratios of buffer to urea for broth cultures are 2:1 to 4:1. In agar, a 10:1 buffer to urea ratio can prevent inhibitory pH increases at the site of colonial growth, preventing resulting loss of organism viability. In broth systems the addition of a reducing agent, such as 1mM sodium sulfite, reduces the lag phase for many strains of *U. urealyticum*. The organism is detected in culture by its production of a readily recognizable pH change resulting from urea hydrolysis.

Of the potentially pathogenic mycoplasmas, *M. hominis* is the least fastidious. In addition to H agar and H broth, nonhemolytic colonies may be observed on many routine blood-containing bacteriologic media. The most productive are those with an enriched base such as Columbia CNA, supplemented anaerobic blood agar, and chocolate agar; growth on standard sheep blood agar is inconsistent. On these media, *M. hominis* grows fastest under conditions of anaerobiosis and somewhat slower under hypercapneic incubation. Colonies are so small that they usually are only detectable when present in large numbers on these media and when plates are carefully examined. The organism may be observed on plates incubated anaerobically after 48 hours; but some isolates may require 72 hours before growth is visible. Growth is typically noted 24 hours later with cultures incubated in 5% CO_2. In either environment, *M. hominis* colonies are approximately 0.05 mm in diameter and may be interpreted as specimen debris or water condensation when plates are examined with the unaided eye. Debris, however, is seldom observed outside of the area of primary inoculation. Strong light assists in plate examination. Colonies may be noted only when media are examined under light at an oblique angle. Plate examination under a dissecting microscope also assists in colony recognition.

Whereas *U. urealyticum* produces no discernible turbidity in broth culture, *M. hominis* can cause a faint, yet distinct haze. For optimal isolation of both organisms, 0.1-mL aliquots of the clinical specimen should be inoculated with a loop or pipette to both agar and broth media. Both U and H agar plates should be streaked for isolation. Plates may be incubated at 35°C to 37°C under aerobic, anaerobic, or hypercapneic conditions. Anaerobic incubation of these media does not appear to increase recovery rates for either *M. hominis* or *U. urealyticum*, but such conditions may discourage the growth of aerobic bacteria and fungi. Strict anaerobic conditions, however, enhance growth of *M. fermentans* and prevent the growth of *M. genitalium*. Care should be exercised to avoid drying of agar plates under all incubation conditions. Aliquots of specimen (0.1 mL) also should be inoculated into H broth and a U diphasic medium system and incubated aerobically. Plates should be examined for growth after 1, 2, 3 and 4 days' incubation with a stereoscopic microscope at 40 to 60× total magnification, with the agar surface viewed through the bottom of the plate to avoid media contamination. Oblique lighting enhances colony visibility. Standard compound microscopes work poorly because of the short working distance of the lenses and poor contrast. The U broth must be examined twice daily to detect a color change indicative of mycoplasmal growth.

Nearly all isolates of both *M. hominis* and *U. urealyticum* will be detected within 5 days of inoculation in broth media. Cultures rapidly decline in viability after peak growth, which occurs prior to any observable pH shift. This complicates choosing an optimal time for subculture. Aliquots from all positive broth cultures should be immediately subcultured to both a fresh U broth and a U agar plate. *U. urealyticum* colonies can be identified using a $CaCl_2$ staining technique (0.1 M $CaCl_2$ and 0.1 M urea) or with the single-reagent test of Shepard (135, 136). Upon addition of a drop of reagent to the plate, colonies of *U. urealyticum* produce a brown halo discernible under transmitted light within 1 to 5 minutes, whereas colonies of *M. hominis* demonstrate no such reaction. Positively staining transferable isolates that hydrolyze urea may be definitively reported as *U. urealyticum*.

M. hominis produces relatively large colonies (200 to 300µm in diameter) on both U and H agars, unless lincomycin is included in the medium as a selective agent. Although most large-colony isolates recovered on H agar are *M. hominis*, bacteria-producing colonies similar to *M. hominis* are sometimes recovered. *M. hominis* can be distinguished from cell-walled bacteria based

on its failure to Gram stain. *M. hominis* colonies recovered on blood-containing media will stain, however, using the traditional Dienes technique. Subculture to a penicillin-containing medium, such as A-7 agar, which inhibits some bacteria, may in some case clarify whether an isolate is indeed a mycoplasma. Large-colony isolates may be presumptively identified as *M. hominis* if they are transferable and utilize arginine (in H broth supplemented with 50 mM arginine). A commercially available arginine broth containing penicillin, erythromycin, polymyxin, and amphotericin (Remel, Lenexa, Kansas) originally formulated for the isolation of *M. hominis* from contaminated urogenital specimens, prevents false-positive reactions that can occur if organisms being tested are arginine-utilizing bacteria other than mycoplasmas. The medium can be inoculated with a few drops of a saline suspension of growth from either the primary or the subculture plates. The inoculum can be prepared by transferring a small block of colony-bearing agar to a tube of saline and vortexing the tube vigorously. Both the inoculated tube and an uninoculated control should be incubated overnight and read for a color change from salmon to red, indicating arginine hydrolysis. A final examination of a Gram-stained smear of positive arginine broths assures that the reaction is not the result of contaminating bacteria.

Colonies may be transferred with an inoculating needle, but an agar block method generally is considered more convenient. With this technique, a small piece of medium bearing the colonies is cut from the agar plate and transferred face down to the medium of choice. The block is then rubbed over the first quadrant of the plate to transfer the colonies, and the remainder of the plate is streaked in a conventional manner using a needle or loop. The inoculated medium is incubated and examined after 24 and 48 hours. Colonies should have the characteristic “fried-egg” appearance commonly associated with mycoplasmas. Upon subculture, colonies are typically larger than on primary isolation media, and, as a result, more easily detectable with the unaided eye. Attempts to isolate *Mycoplasma* and *Ureaplasma* species are, unfortunately, commonly hampered by the presence of confounding artifacts on culture media. Mammalian cellular debris may closely resemble colonies on agar plates. Debris, however, does not transfer.

Serology

Serologic assays for detection of antibodies to either *U. urealyticum* or *M. hominis* have been used in certain research laboratories (137, 138). These tests, however, are not commercially available and, therefore, are rarely useful for clinical diagnosis.

Interpretation of Laboratory Results

Interpretation of the significance of recovery of either *U. urealyticum* or *M. hominis* from a clinical specimen is sometimes difficult. Under most circumstances, isolation of *U. urealyticum* from cervical specimens has little significance because of the high colonization rate in apparently healthy women. Although reported colonization rates with *M. hominis* are somewhat lower, its recovery from cervical specimens must likewise be interpreted with caution. Isolation of either organism from usually sterile cites such as blood, amniotic fluid, fallopian tube, upper urinary tract, and other internal sites has greater significance. Recovery of *U. urealyticum* in large numbers from the urethra of a male with nongonococcal urethritis is probably significant.

Positive culture results should be interpreted only in conjunction with other laboratory results and clinical findings.

Therapeutic Considerations

Mycoplasmas are characteristically susceptible to many broad-spectrum antibiotics including fluoroquinolones, chloramphenicol, tetracycline, and clindamycin. At clinically achievable levels, tetracycline inhibits most strains of *M. hominis* and *U. urealyticum*. By comparison, erythromycin demonstrates only moderate activity against *U. urealyticum* and lacks activity against *M. hominis*. In part because standardized procedures are not currently available, *in vitro* susceptibility testing usually is not performed on clinical isolates of these organisms.

Both *C. trachomatis* and *U. urealyticum* apparently cause NGU. A 100-mg dose of a tetracycline, such as doxycycline, given twice daily for 7 days, represents standard therapy for this syndrome. About 10% of ureaplasmas are resistant to tetracyclines, however, and patients with infections caused by such strains typically fail to respond to treatment with a tetracycline. Most strains of tetracycline-resistant *U. urealyticum* appear to be susceptible to erythromycin.

When cases of prolonged postabortal or postpartum fever are presumed to be caused by *M. hominis*, treatment with a tetracycline should be considered, realizing that tetracycline-resistant strains may be involved. In other more serious illnesses in which *M. hominis* is thought to be responsible, such as arthritis in a hypogammaglobulinemic person and neonatal infections, tetracycline resistance always should be considered.

AIDS-Associated Mycoplasma (Mycoplasma incognitos)

A previously unrecognized pathogenic human mycoplasma tentatively designated *M. incognitos* was first recognized by Lo and coworkers (139). The organism, initially thought to be a large DNA virus, was isolated as the result of an experiment in which purified DNA derived from Kaposi's sarcoma tissues of persons with AIDS was introduced into NIH3T3 cells by transfection. Cells were monitored for transforming foci, and a transformant, designated sb51, was recovered. The agent, which passed through a 0.22 m-pore-size filter, was subsequently shown to infect normal cells. Electron microscopic examination of infected cells demonstrated nearly spherical viruslike particles 140 to 280 nm in diameter. The agent, detected in the cytoplasm, along the cell membrane, and occasionally in the nuclei of infected cells, showed minimal or no cytopathic effect. The small size of the membrane-bound intracellular particles, their capability to result in persistent cell infection, and their ability to transfect susceptible cells resulted in the conclusion that the particles were viral in nature. Southern blot analysis, however, revealed that no cross-hybridization existed with any known virus tested. Ribosomal RNA analysis, using *Escherichia coli* probes, revealed significant homology, suggesting that the VLIA particles were procaryotic in nature. Because the agents were filterable, subsequent testing with DNA from several *Mycoplasma* species was attempted. Results

indicated that the particles were related to mycoplasmas.

Biochemical, antigenic, and DNA analysis demonstrated that *M. incognitos* was distinct from all other recognized human *Mycoplasma* species, but very closely related to *M. fermentans*. The organism can be grown in SP-4, a cell-free medium designed to support the growth of fastidious mycoplasmas. It is also similar to *M. fermentans* in that it utilizes glucose both anaerobically and aerobically, metabolizes arginine, and is susceptible to tetracycline, but resistant to erythromycin (140).

Several differences between the "Lo mycoplasma" and *M. fermentans*, however, do exist and may justify assigning the organism to a new species, *M. incognitos*. The organism is capable of intracellular parasitism, a trait not reported with *M. fermentans*, and individual colonies and particles are smaller than those of *M. fermentans*. It also appears to be somewhat more fastidious than *M. fermentans* and has not been isolated directly from infected human or experimental animal body fluids or tissues. In addition, monoclonal antibodies directed against *M. incognitus* fail to cross-react with *M. fermentans*. Restriction enzyme digests also differ between the two organisms. Finally, the pathogenic properties of *M. incognitos* are quite different from those described for *M. fermentans*.

The Lo mycoplasma was originally detected in peripheral blood mononuclear cells and autopsy tissues of persons with AIDS. The organism since has been visualized in brain, liver, and spleen tissues of infected patients using electron microscopy. Histopathologic tissue examination reveals a variety of results ranging from little change to occasional extensive necrosis, with or without evidence of an inflammatory response. No other known infectious agents were demonstrable in the lesions. Whether *M. incognitos* represents a cofactor, an opportunist, or a primary etiologic agent in the lesions described in these immunocompromised patients remains unclear. The organism is clearly infectious, and is apparently responsible for disease progression in some individuals. In one recent report, six geographically separated, previously healthy non–HIV-positive patients with acute influenza-like symptoms died with fulminant *M. incognitos* infection with multisystem involvement 1 to 7 weeks after onset of illness. The organism was found extracellularly and intracellularly in necrotizing lesions of spleen, lung, liver, and adrenal gland tissues in all six patients, as demonstrated by electron microscopic, immunoserologic, *in situ* hybridization, and polymerase chain reaction (PCR) techniques. No other potential infectious agents were demonstrable.

Monoclonal antibodies directed against epitopes of *M. incognitos* also have been used to immunohistochemically demonstrate the organism in liver, spleen, lymph node, thymus, or brain in 22 patients with AIDS, as well as in two placentas delivered by patients with AIDS. A 33S-labeled DNA probe specific for *M. incognitos* and an *in situ* hybridization technique also have detected *M. incognitos*-specific genetic material in these tissues. Likewise, ultrastructural analysis of the specific areas of tissue highly positive for *M. incognitos* antigens has revealed characteristic organisms both intracellulary and extracellularly. The organism is, therefore, apparently cytopathic and cytocidal. *M. incognitos* also has been shown to cause fatal systemic infection in experimental silver leaf monkeys. Experimentally infected animals showed only a transient immune response and developed wasting syndromes, resulting in death within 7 to 9 months. Progressive weight loss, tissue necrosis, and a minimal inflammatory cellular and antibody response were noted in all infected animals. The organism was found in the nuclei and the cytoplasm of infected tissues in the absence of other possible etiologic agents. The organism apparently has little specific tissue tropism.

Neither the prevalence nor the significance of *M. incognitos* in human disease is known. The organism appears to be invasive in some immunocompromised patients and capable of inducing immunosuppression in previously healthy persons. The functional deficit of various organ systems, in the absence of other etiologic agents, is apparently associated with cryptic infection with *M. incognitos*, therein the proposed species name. The high prevalance of *M. incognitos* infection in persons with AIDS has led to conjecture that it plays a significant disease-promoting role in HIV-infected individuals. Independently, French investigators have reported that tetracycline analogs inhibit the cytopathic effects of HIV in T-lymphoblastoid tumor cell lines without suppression of virus production (141). Consistent with this theory, *M. incognitos* infection has been shown to accentuate the cytocidal effects of HIV in CD4-positive human cell lines (142). Because the Lo mycoplasma is potentially susceptible to antimicrobial therapy, this possibility warrants in-depth investigation. *In vitro*, ciprofloxacin appears to be the most active antibiotic against *M. incognitos*. Carefully controlled and matched clinical trials in appropriate persons with AIDS using antibiotics to which the Lo mycoplasma is highly susceptible *in vitro* should help to clarify this issue.

The suggestion that a procaryotic cell can be generated by transfection of a eucaryotic cell is highly controversial. Other plausible explanations exist for this reported finding, including contamination of the cell line with the Lo mycoplasma during laboratory manipulations or introduction of the organism from an unidentified source during routine passage of the cells in culture or during early transfection cycles. Alternatively, *M. incognitos* may have been present in the Kaposi's sarcoma cells that served as the source of transfecting DNA. This process would involve the transcription and translation of the entire *M. incognitos* genome (approximately 500 megadaltons) by the genetic apparatus of a eucaryotic cell and the final assembly of a functional procaryotic cell. Although highly unlikely, further research is likewise justified to investigate this possibility.

RICKETTSIA

Part of "53 - Chlamydia, Mycoplasma, and Rickettsia"

Agents of rickettsial disease were originally grouped together based on the concept that arthropods are an important part of the organisms' life cycle and in part based on the presumed obligate intracellular associations of these agents. Thorough studies of these organisms were greatly hindered because most were difficult or impossible to cultivate; therefore, classification based strictly on this schema is flawed. The advent of molecular taxonomic classification has confirmed a close genetic and phenotypic relationship between some of the genera (*Rickettsia* and *Ehrlichia*) and proved that convergent evolution has yielded characteristics that led to the inclusion of other more diverse genetic and phenotypic genera and species such as *Coxiella* and *Rochalimaea* (now reclassified as *Bartonella*) (143). The species previously known as *Rickettsia tsutsugamushi* has been determined significantly divergent to be classified into a new genus, *Orientia* (144). In general, four genera are now considered to include important human pathogens in the family Rickettsiaceae: *Rickettsia*, *Orientia*, *Ehrlichia*, and *Coxiella*. *Bartonellaceae* remain a separate family in the order Rickettsiales.

The members of these genera are true bacteria that contain DNA and RNA and divide by binary fission. By ultrastructure, these bacteria have gram negative-type cells walls, and when examined by light microscopy they appear as small gram-negative intracellular bacteria. Aside from *Bartonella* spp., none have been successfully cultivated on artificial, cell-free medium; cultivation in tissue culture has been accomplished for many of these species. Members of the genus *Rickettsia* and *Orientia* live in the cytoplasm of the host cell, those in the genera *Ehrlichia* and *Coxiella* live in membrane-bound vacuoles, whereas *Bartonella* spp. may reside in either extracellular or intracellular niches depending upon the host. Ehrlichiae grow to form a cytoplasmic aggregate of bacteria called a morula. Conversely, *C. burnetii* actively proliferates within the acidified phagolysosome to form aggregates containing in excess of 100 bacteria.

Rickettsia species contain lipopolysaccharide (LPS) as in other related Gram-negative bacteria, and some epitopes in this LPS convey group and genus serologic cross-reactivity. *O. tsutsugamushi* lacks both LPS and peptidoglycan in its cell wall, and *Ehrlichia* species have no detectable LPS, but apparently do contain peptidoglycan in their cell walls (145, 146). *C. burnetii* undergo an LPS "phase variation" similar to that described for smooth and rough strains of the Enterobacteriaceae. Immunoblot analysis and gene cloning has demonstrated the presence of highly conserved proteins as well as species- or genus-specific proteins among genetically related species. Evidence is now accumulating to suggest that some of the major outer membrane proteins of *Rickettsia*, *Ehrlichia*, *Coxiella*, and *Bartonella* species function as cellular ligands to promote host cell attachment and invasion, while others serve as structural components (147,148 and 149). Owing to the inability to achieve gene deletion and supplementation for some of these species, definite virulence factors with proven in vivo significance have yet to be identified.

Rickettsial diseases of humans are separated into groups on the basis of clinical characteristics, etiologic agent, insect vectors, and epidemiology, including geographic distributions (Table 53.5). With the exception of *C. burnetii* and possibly some *Bartonella* spp., rickettsiae usually are transmitted among mammalian reservoirs by arthropod vectors. Humans are not an essential link in their natural cycle but become chance hosts when they are bitten by an infected vector, scratch or rub infectious arthropod feces into the skin, inhale or ingest the organism, or have the infectious agent inoculated into the skin via cat scratch or bite. The presence of indigenous reservoirs and vectors, plus the continuous possibility that new strains of rickettsiae may be introduced from other parts of the world, make the threat of new outbreaks and continued low prevalence, high morbidity infections in the United States a continuing concern.

TABLE 53.5. SUMMARY OF FEATURES IN HUMAN RICKETTSIAL DISEASES

Disease	Agent	Arthropod Vector-Transmission	Hosts	Diagnostic Tests[a]
Spotted fever				
Rocky Mountain spotted fever	*Rickettsia rickettsii*	Tick bite	Dogs, rodents	IFA, DFA, IH
Boutonneuse fever (Mediterranean spotted fever)	*Rickettsia conorii*	Tick bite	Dogs, rodents	IFA, DFA, IH
Rickettsialpox	*Rickettsia akari*	Mite bite	Mice	IFA
Typhus				
Murine typhus	*Rickettsia typhi/Rickettsia felis*	Rat flea or cat flea feces	Rats, opossums	IFA, DFA
Epidemic typhus	*Rickettsia prowazekii*	Louse feces	Humans	IFA
Brill-Zinnser disease (recrudescent typhus)	*R. prowazekii*	Reactivation of latent infection	Humans	IFA
Flying squirrel (sylvatic) typhus	*R. prowazekii*	Louse or flea of flying squirrel	Flying squirrels	IFA
Scrub typhus				
Scrub typhus	*Orientia tsutsugamushi*	Chigger bite	Rodents?	IFA
Ehrlichioses				
Human monocytic ehrlichiosis	*Ehrlichia chaffeensis*	Tick bite	Deer, dogs?	IFA, PCR
Human granulocytic ehrlichiosis	*Ehrlichia phagocytophila group*	Tick bite	rodents	IFA, PCR, tissue culture
Human granulocytic ehrlichiosis – USA midwest	*Ehrlichia ewingii*	Tick bite	dogs	IFA, PCR
Q fever	*Coxiella burnetii*	Ticks?	Cattle, sheep, goats, cats, rabbits	IFA
Bartonelloses				
Bacillary angiomatosis and visceral peliosis	*B. henselae*	Cat scratch or bite	Cats	Culture, PCR,
	B. quintana	Louse feces	Humans	IFA, IH
Cat scratch disease	*B. henselae*	Cat scratch or bite	Cats	Culture, PCR
	B. quintana	Louse feces	Humans	IFA, IH
	B. clarridgeiae	Cat scratch or bite	Cats	
	B. vinsonii	?Rodent ectoparasite	?Rodent	
Febrile bacteremia, endocarditis	*B. henselae*	Cat scratch or bite	Cats	Culture, PCR,
	B. quintana	Louse feces	Humans	IFA
	B. clarridgeiae	Cat scratch or bite	Cats	
	B. elizabethae	?Rodent ectoparasite	?Rodent	
	B. vinsonii	?Rodent ectoparasite	?Rodent	
Carrion's disease and Oroya fever	*B. bacilliformis*	Sandflies	?Humans	Culture, Wright or Giemsa-stained blood smear

DFA or IH test can be used to detect *Rickettsia* in tissues samples. PCR may be performed to detect *Ehrlichia, Coxiella*, or *Bartonella* nucleic acids in acute-phase blood using specific oligonucleotide primers. The preferred confirmatory serologic tests are IFA or ELISA (where available). Cultivation of *Bartonella* may be accomplished in a BHL2 laboratory, whereas cultivation of other rickettsiae should be attempted only in specialized BHL3 health laboratories.

Although grouped together under the name "rickettsial diseases," the pathogenetic mechanisms of each of the various genera in the family Rickettsiaceae are different. Members of the genus *Rickettsia* are vasculotropic and infect endothelial cells, whereas both ehrlichiae and *Coxiella burnetii* reside within phagocytes, which accounts for some of the clinicopathologic differences encountered. Although patients with human monocytic and granulocytic ehrlichiosis often present with an illness similar to the vasculotropic rickettsioses, the pathogenesis is clearly different because endothelial cell infection and vasculitis are rare events. In contrast to these rickettsial infections, humans acquire *C. burnetii* predominantly after inhalation of infectious aerosols.

Rickettsiae, ehrlichiae, and *Coxiella burnetii* have been cultivated in tissue culture cells; a specific diagnosis may be rendered within as few as 48 hr. for RMSF and Q fever. Culture for *E. chaffeensis* may take up to 1 month to achieve, while culture for the HGE agent usually is more rapid. *Bartonella* species are fastidious but can be cultivated by prolonged routine axenic culture. Some consider rickettsial cultivation hazardous; however, given the widespread application of standard precautions, such methods should be reconsidered for early and specific identification of rickettsial infections.

Spotted Fever Group Rickettsioses

Rickettsia rickettsii

Rocky Mountain spotted fever (RMSF) is the most frequent rickettsial disease in the United States. Systemic endothelial cell infection by the obligate intracellular bacterium *Rickettsia rickettsii* leads to disease after proliferation of the rickettsia within the cytoplasm of endothelial cells, vasculitis, microvascular leakage, tissue hypoperfusion, and end-organ damage (150). This potentially rapidly fatal infection is infrequently diagnosed, but often is considered in the differential diagnosis of patients after tick exposure who present with fever, headache, and rash.

Epidemiology

RMSF occurs in almost every state of the continental United States, southwestern Canada, Mexico, Central America, and South America. However, most cases are diagnosed in North Carolina, Georgia, Virginia, Tennessee, Oklahoma, and Maryland. Despite this, the geographic prevalence varies over years and decades potentially influenced by changes in ecology of tick vectors and climate. RMSF incidence varies cyclically over decades peaking last in 1981 with the last nadir in 1996. Since 1996, the incidence of RMSF has risen, suggesting yet another peak of cases and attendant fatalities within the next decade. Despite the name, only a small proportion of all cases occur in the Rocky Mountain regions. Habitats that favor activity of *Dermacentor variabilis* (American dog tick, in the Eastern United States and Canada) or *Dermacentor andersoni* (wood tick in the Western United States), including wooded areas or coastal grassland and salt marshes are where the highest incidence rates are documented. However, well documented foci of infection are observed in both rural and urban settings, including the South Bronx (151). Most cases in the United States occur during April through October with peak tick activity and human exposure. The incidence of disease is highest in children 5 to 9 years of age. Ticks are the natural hosts, reservoirs, and vectors of *R. rickettsii*. Ticks naturally maintain infection by transovarial transmission (passage of the organism from infected ticks to their progeny),

but natural maintenance also requires transient infection of small and medium size animals from which uninfected tick lines derive infection (152). Humans are infected only accidentally and are generally considered dead-end hosts. Rarely, RMSF is acquired by inhalation of infectious aerosols in laboratory workers propagating the infectious agent; only one case of transfusion-transmitted RMSF has been documented (153).

Pathology and Clinical Manifestations

Vasculitis results after infection of endothelial cells and direct rickettsia-mediated cellular injury (150). Perivascular lymphohistiocytic infiltrates and edema without significant endothelial damage are the first findings and coincide with the appearance of macules and maculopapules. The characteristic lymphohistiocytic or leukocytoclastic vasculitis of small venules and capillaries (Fig. 53.4) leads to erythrocyte extravasation and skin petechiae (154), but also the more serious consequences of widespread vascular leakage, hypoperfusion and ischemic injury of organs. Rarely, vessels become completely thrombosed and cause infarction or hemorrhagic necrosis of organs, digits, or extremities. The most significant complications of RMSF result from microvascular leakage leading to noncardiogenic pulmonary edema or cerebral edema after meningoencephalitis (150).

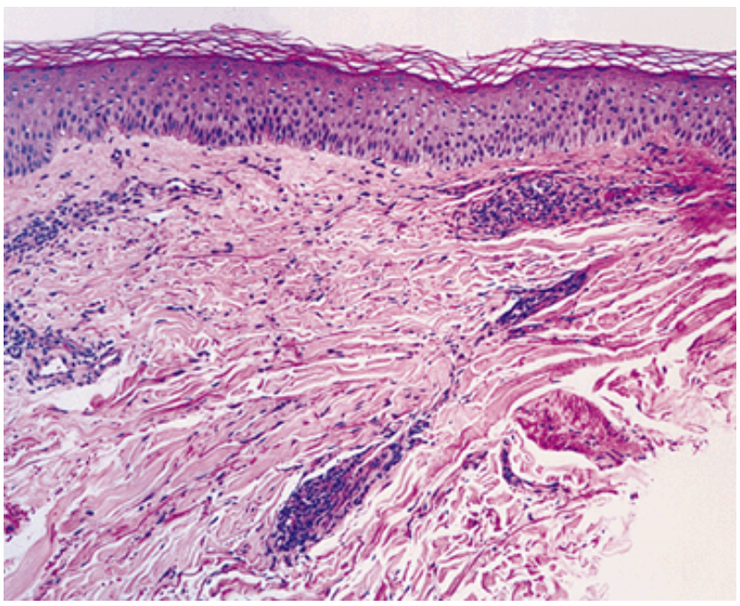


FIGURE 53.4. Cutaneous histopathologic changes in Rocky Mountain spotted fever. The skin biopsy, obtained from a petechial lesion, demonstrates superficial and mid-dermal lymphohistiocytic and leukocytoclastic vasculitis. H&E, original magnification 400×.

After the infectious agent is inoculated via tick bite into the dermis, the rickettsiae attach to endothelial cells via rickettsial

protein adhesins (147). Rickettsial phospholipase activity initiates focal host cell membrane injury that induces phagocytosis for repair (155). The engulfed rickettsia then gains access to the cytoplasm by continued membrane lysis. *R. rickettsii* initiates directional movement via intracellular actin polymerization allowing easy invasion of neighboring cells (156). While proliferating, the rickettsiae damage host cell membranes by peroxidative reactions and continued phospholipase activity (155, 157). Infection initiates an inflammatory cascade with cytokine release, including tumor necrosis factor-α (TNF-α), interleukin-1ß (IL-1ß), and interferon gamma (IFN-γ) from recruited mononuclear phagocytes and lymphocytes and from infected endothelial cells (158). Although these inflammatory and immune responses have been suspected to contribute to vascular injury, the beneficial effects of inflammation and immunity outweigh inflammatory damage since depletion of TNF-α and IFN-γ in animal models diminishes survival and increases morbidity (159). The *Rickettsia*-mediated injury associated with upregulated procoagulant activity on infected endothelial cells may promote platelet adhesion, leukocyte emigration, and coagulation factor consumption, a clinical syndrome similar to but distinct from disseminated intravascular coagulation (160).

After a median incubation period of 7 days (range 2 to 14 days), an initially nonspecific illness with headache, fever, and anorexia ensues. Symptoms that reflect gastrointestinal involvement, such as nausea, vomiting, diarrhea, or abdominal pain occur often early in the disease. The skin rash is observed usually after 3 days of illness, but the typical clinical triad of headache, fever, and rash is absent in 97% of all patients at first presentation (161). Approximately 10% of patients, predominantly adults, lack or have an atypical rash. The rash may spread rapidly over the entire body, including soles and palms, and may later become petechial. Severe myalgia and malaise may accompany the persistent fever and headache. Meningoencephalitis may lead to changes in sensorium, delirium, or coma. Additional central nervous system involvement is manifested by ataxia, meningismus, or auditory deficits. Other severe complications may include pneumonitis with noncardiogenic pulmonary edema, myocarditis, acute renal failure, and vascular collapse. Clinical findings associated with death include hepatomegaly, jaundice, stupor, acute renal failure, respiratory distress, and a disseminated intravascular coagulation-like syndrome (161, 162).

Thrombocytopenia, minimal leukopenia or the lack of leukocytosis with a marked left shift in a severely febrile patient, and hyponatremia are observed in about 50% of patients and may facilitate early consideration of the diagnosis. Most other laboratory findings are unhelpful and vary only with involvement of specific organs. The CSF in RMSF meningoencephalitis usually is characterized by a mononuclear cell pleocytosis.

Delayed diagnosis and treatment are important reasons for the high case fatality (2% to 7%) and morbidity rates of RMSF (163). Thus, the diagnosis of RMSF initially must be based upon clinical suspicion so that early therapy can be initiated in suspected cases. Prompt therapy in uncomplicated infections leads to defervescence within 1 to 3 days and recovery within 7 to 10 days.

Diagnostic Procedures

No laboratory test definitively establishes an early diagnosis of RMSF. Therefore, treatment should not be withheld while awaiting *R. rickettsii*-specific laboratory tests. With the availability of skin biopsy of a typical rash lesion in 85% of patients, RMSF can be diagnosed as early as day 3 to 4 of illness by demonstrating spotted fever group rickettsial antigens within endothelial cells using immunofluorescence or immunoenzymatic methods (Fig. 53.5) (164, 165). The sensitivity of these methods is not greater than 70% and is diminished with prior antimicrobial therapy, biopsy of suboptimal lesions, or limited tissue examination owing to the focality of infection. Evaluation of blood for *R. rickettsii* nucleic acids by PCR has been disappointing because of the very low levels of rickettsemia in most patients (166). Most diagnoses are established in convalescence by demonstration of a fourfold or greater rise in antibody titer by

indirect fluorescent antibody (IFA) assay; a single elevated convalescent IFA titer ≥ 64 supports the diagnosis in a clinically compatible case (167). Antibodies usually do not develop before day 10; thus, serologic tests in the active phase of infection will be unhelpful. A variety of older, less frequently used serologic tests have been described, but are rarely used and do not merit further discussion. Commercially available ELISA and solid-phase immunoassays have shown promise for objective, sensitive, and specific serologic confirmation (168). Weil-Felix testing should not be performed because of the lack of both sensitivity and specificity.

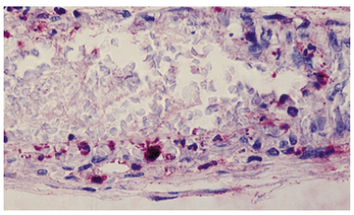


FIGURE 53.5. Immunoperoxidase demonstration of *Rickettsia rickettsii* in a large dermal vessel from a patient with Rocky Mountain spotted fever. Note the presence of small red-stained coccobacilli within endothelial cells and a background of lymphocytes and histiocytes infiltrating the vessel wall. Immunoperoxidase, original magnification 630×.

Therapy and Prevention

Effective therapy for RMSF requires either a tetracycline antibiotic or chloramphenicol. Alternatives such as fluoroquinolones or new the macrolides have not been evaluated and other broad-spectrum antibiotics such as ß-lactams or aminoglycosides are not effective. A high degree of morbidity and mortality are associated with sulfonamide use. Because both the tetracyclines and chloramphenicol are rickettsiostatic, recovery depends upon induction of appropriate immune responses. As a result, prophylactic therapy of tick bites to prevent RMSF will only serve to delay the onset and confound diagnosis. Because vaccines are not available, prevention relies on eliminating tick infestations of dogs, avoiding tick-habitats, the use of insect repellents and special protective clothing, and careful inspection of body surfaces after tick exposures.

Other Spotted Fever Rickettsioses

There are many human pathogens among the spotted fever group aside from *R. rickettsii*, including *R. conorii* (Boutonneuse fever), *R. sibericus* (North Asian tick typhus), *R. japonica* (Oriental spotted fever), *R. australis* (Queensland tick typhus), *R. honei* (Flinders Island spotted fever), the unnamed Israeli spotted fever rickettsia, and *R. africae* (African tick bite fever). As opposed to tick transmission for most members of the spotted fever group, *R. akari*, which causes rickettsialpox, is transmitted by mite bites. Boutonneuse fever (Mediterranean spotted fever, Kenya tick typhus, or Indian tick typhus), caused by *R. conorii*, occurs in much of southern Europe, south western and central Asia, the middle East, Morocco, Ethiopia, Kenya, and South Africa. The illnesses caused by other spotted fever group rickettsiae are distributed regionally and cause clinically similar infections with occasional characteristic features (169). *R. sibericus* occurs mostly in Russia, China, Mongolia, and Pakistan, *R. australis* or *R. honei* occur in Australia, *R. japonica* in Japan, and *R. africae* is found in South Africa. By antigenic and genetic analyses, these species are closely related to *R. rickettsii*. The recently described *R. felis* that is transmitted by cat fleas (*Ctenocephalides felis*) in the Southern United States, Mexico, and potentially Central America and causes a murine typhuslike illness, has attributes most like a spotted fever group rickettsia as well (170, 171).

Boutonneuse fever has been steadily increasing in incidence in southern Europe since 1980. The seroprevalence may be as high as 11% to 26% in some regions (172). *R. conorii* is transmitted by *Rhipicephalus sanguineus*, the brown dog tick, or by *Dermacentor*, *Haemaphysalis*, *Amblyomma*, *Hyalomma*, or *Ixodes* spp. ticks. As for RMSF, incidence is highest in summer months in the Mediterranean basin, but may occur at any time in warm regions with year-round tick activity. Imported infections are frequent in travelers to endemic regions, particularly those with recent exposures via safaris and bush land (173).

R. akari, the causative agent of rickettsialpox, is transmitted by *Allodermanyssus sanguineus*, the mouse mite. Despite the mouse host's wide distribution in urban environments, the disease is infrequently diagnosed (174). Recent investigations suggest that a major target of infection in humans may be macrophages as well as endothelial cells (175).

Pathology and Clinical Manifestations

Fever, headache, myalgias, and a maculopapular rash with onset 3 to 5 days after symptoms typically are observed with Boutonneuse fever (169). An eschar (*tache noire*) at the initial site of tick attachment and regional lymphadenopathy will be detected in about 70% of patients. The clinical illness is severe and similar to RMSF in up to 6% of infected individuals and death occurs in 1.4% to 5.6%. Severe manifestations include purpura, neurologic involvement, respiratory insufficiency, acute renal failure, and severe thrombocytopenia. A particularly malignant form occasionally is observed in glucose-6-phosphate dehydrogenase deficient patients and in individuals with alcoholic liver disease or diabetes mellitus. Most spotted fever group rickettsioses present with similar manifestations and laboratory findings; an eschar is present in many at the site of initial tick bite. A form of spotted fever that occurs in Israel has a more severe course including pediatric fatalities (176). It is important to recognize that diagnosis should be rendered based upon the clinical presentation, but since some patients do not recall a tick bite, and others do not have rash or an eschar, initial diagnosis may be very difficult.

Rickettsialpox, caused by *R. akari* infection, usually is mild and often recognized when a varicelliform rash is observed. Most patients with rickettsialpox experience a typical macular or maculopapular rash as seen in other rickettsioses prior to the appearance of vesicles associated with fever, headache, and chills (174). Spontaneous resolution even in the absence of antirickettsial therapy is the norm and complications are rare.

Laboratory confirmation of other spotted fever group rickettsioses is similar to that for RMSF and is achieved with immunohistologic studies that demonstrate spotted fever group rickettsiae in skin biopsies (Fig. 53.6 and Fig. 53.7) or in circulating endothelial cells (177), *in vitro* cultivation in shell vial tissue culture (178), and most often by the serologic studies that demonstrate a four-fold increase in spotted fever group antibodies during convalescence. Fortunately, many reagents used for RMSF diagnosis can be used for diagnosis of most spotted fever group infections. Therapy and prevention are similar to that for RMSF and also may include the use of fluoroquinolone antibiotics.

Typhus Group Rickettsioses

Rickettsia typhi

The causative agent of murine typhus is *Rickettsia typhi*, a rickettsia in the typhus group transmitted to man by fleas (179). The

rickettsia is not maintained in the flea by transovarial transmission; thus, flea to rat (or opossum) to flea (vertical transmission) is an important mechanism for natural maintenance. The new rickettsia, *Rickettsia felis*, that is genetically related to *R. typhi* and *R. rickettsii*, has been identified in south Texas and Mexico; it is maintained by transovarial transmission in cat fleas (*Ctenocephalides felis*) (180). Humans are accidentally infected when rickettsia-infected flea feces contaminate flea-bite wounds.

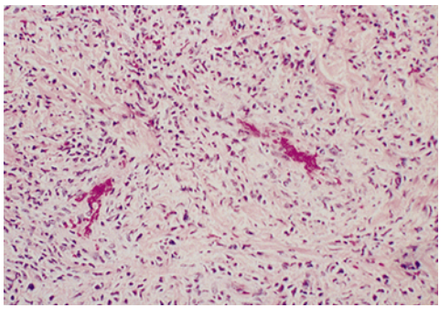


FIGURE 53.6. *Tache noire* (*eschar*) of Boutonneuse fever (*Rickettsia conorii*) acquired by a patient who traveled to Africa. Note the lymphohistiocytic infiltrate, marked edema, fibrin deposition, and vascular injury. H&E, original magnification 400×.

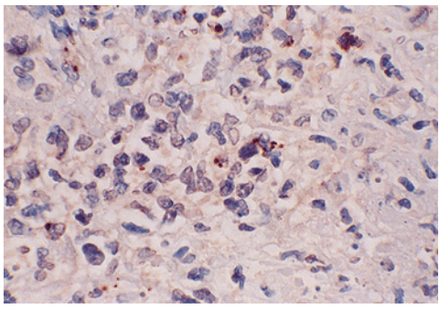


FIGURE 53.7. *Tache noire* (*eschar*) of Boutonneuse fever (same patient as in Fig. 3) with demonstration of *Rickettsia conorii* by immunoperoxidase staining. Note the intracellular location of the rickettsiae presumably limited to endothelial cells and potentially macrophages.

Epidemiology

Although murine typhus occurs worldwide in warm, coastal environments, in the United States it is most prevalent in southern Texas and southern California. Two distinct ecologies exist, one involving rat fleas (*Xenopsylla cheopis*) and rats (*Rattus* species) and another involving cat fleas and opossums (181). As expected, the disease occurs most often during spring, summer, and other times when flea vectors are active.

Pathology and Clinical Manifestations

R. typhi causes disease in a way similar to that of *R. rickettsii* in RMSF, by primary infection of endothelial cells and induction of vasculitis and vascular injury (182). Other important pathologic findings include interstitial pneumonitis, interstitial nephritis, myocarditis, meningitis, and mild hepatitis. With severe vasculitis and widespread inflammatory injury, multi-organ damage may result and death may result.

Flea feces that contain *R. typhi* are inoculated into the flea-bite wound after the flea defecation-feeding reflex. The rickettsia proliferate locally, spread to lymphatics and venous blood to infect endothelium in many tissues. Although typhus group rickettsiae infect endothelial cells, unlike spotted fever group organisms, actin is poorly polymerized and cellular injury is probably induced by mechanical lysis after large numbers of rickettsiae accumulate (156). Infection and endothelial cell injury lead to vasculitis and significant vascular barrier dysfunction.

As for other vasculotropic rickettsioses, murine typhus is moderately severe in adults and mild in children (183). After an incubation period of 1 to 2 weeks, a nonspecific febrile illness develops. Important clues for rickettsiosis include headache, rash, and myalgias; other findings that occur in under half of patients include vomiting, cough, and diarrhea or abdominal pain. Neurologic involvement is infrequent in adults, but may include confusion, stupor, coma, seizures, meningismus, and ataxia. Macules or maculopapules on the trunk, extremities, including soles and palms, are observed early, but a petechial rash is infrequent. Complications include relapse, stupor, dehydration, and CSF pleocytosis, multiorgan damage, and death.

Nonspecific laboratory findings include mild leukopenia with left shift, thrombocytopenia, hyponatremia, hypoalbuminemia, and elevations in serum hepatic transaminase levels.

Diagnostic Procedures

As for RMSF, delayed therapy increases morbidity and mortality, and early diagnosis must be based on clinical suspicion (183). Diagnostic confirmation is usually achieved by IFA serologic testing in convalescence. Other tools now being evaluated include PCR amplification of *R. typhi* nucleic acids in acute phase blood, culture by the shell vial assay, and skin biopsy with immunohistology.

Therapy and Prevention

Murine typhus is treated similar to RMSF by the use of tetracyclines or chloramphenicol. A single patient successfully treated with ciprofloxacin has been reported (184); *in vitro* data suggest that azithromycin and clarithromycin may be effective therapeutic agents as well (185). Prevention of murine typhus depends on control of mammalian and flea reservoirs in urban and suburban areas alike.

Rickettsia prowazekii

It has long been considered that humans are the only reservoir of *Rickettsia prowazekii*, the causative agent of epidemic typhus and a member of the typhus group. *R. prowazekii* represents perhaps the most virulent of all of the vasculotropic rickettsiae and is responsible for a significant degree of historical morbidity and

mortality associated with wars prior to the 20th century (186). The infection results when infected human body or head lice feed on uninfected persons. Rickettsiae that infect the lice midgut epithelial cells are defecated into the host abrasions in the skin, into the conjunctivae, or are inhaled. Once established, rickettsemia allows blood-feeding uninfected lice to acquire the infection which is then passed on to other uninfected individuals during times of poor hygiene, crowding, war, etc., and allows epidemic spread. The most recent major outbreak of typhus was discovered in the war-torn parts of Burundi, where nearly 60,000 persons were infected and nearly 6,000 people died (187). Other recent epidemic have been identified in Africa, Mexico, Central America, South America, eastern Europe, Afghanistan, northern India, and China. Epidemic typhus is not always severe; a mild form, caused by a genetic variant of *R. prowazekii* has been identified and associated with exposure to flying squirrels that harbor infected fleas or lice (188). Primary infection results in non-sterile immunity and recrudescence or Brill-Zinnser disease, a mild form of typhus, may occur years later when debility or immunocompromise supervenes.

Clinical Manifestations. Epidemic typhus may be mild or severe. After an incubation period of usually fewer than 14 days, fever, severe headache, abdominal tenderness, and rash are observed in most patients; chills, myalgias, and arthralgias are present in most, while gastrointestinal, respiratory, and central nervous system involvement are seen in a decreasing proportion of patients. The rash starts as pink or erythematous blanching macules and maculopapules and may evolve to petechiae predominantly involving the trunk. The case fatality rates ranges between 3.8% and 20% during outbreaks.

Orientia (Rickettsia) tsutsugamushi

Scrub typhus is caused by the rickettsial pathogen *Orientia tsutsugamushi* (formerly *Rickettsia tsutsugamushi*) (144). The causative agent is genetically and antigenically distinct from spotted fever and typhus group rickettsiae. These small obligate intracellular bacteria likely occupy an intra-endothelial niche; however, members of this monospecific genus lack both lipopolysaccharide and peptidoglycan in their cell envelopes (146). Scrub typhus is predominantly found in the Far East in areas bounded by Korea, Pakistan, and northern Australia. Imported infections to the United States also occur. The rickettsia is transmitted after bites of trombiculid mites in the genus Leptotrombidium, which are both vectors and reservoirs. The infectious agent is maintained by highly efficient transovarial transmission; rodent hosts of mites are probably not involved in natural maintenance. Multiple serotypes are known, many of which share antigenic cross-reactivity, but may represent unique separate species.

Clinical Manifestations. The incubation period for scrub typhus varies from 6 to 21 days before the onset of a mild to severe febrile disease. An eschar appears at the chigger (mite) bite site in less than half of cases. In addition to fever, the illness is characterized by headache, myalgia, cough, lymphadenopathy, maculopapular rash, and gastrointestinal symptoms. Infection can be complicated by meningoencephalitis or interstitial pneumonitis, and fatalities may occur in as many as 7% of untreated patients. Diagnosis is usually confirmed by serology using the indirect fluorescent antibody or immunoperoxidase serologic tests with *O. tsutsugamushi* as antigen.

Treatment

Therapy, as for other vasculotropic rickettsioses, includes doxycycline or chloramphenicol, after which defervescence occurs within 24 to 48 hours. Recently, highly virulent doxycycline-resistant strains have been reported (189). *In vitro* susceptibility tests show that azithromycin may be an effective alternative for scrub typhus.

Ehrlichioses

Although *Ehrlichia* spp are well-recognized etiologic agents of disease in dogs and ruminants, it wasn't until 1987 that human infection in the United States was identified (190). That first case was erroneously attributed to infection by *Ehrlichia canis*; however, subsequent evaluations have shown similar infections transmitted by ticks to be caused by several distinct species of *Ehrlichia* (191,192 and 193). *Ehrlichia* infections are characterized by undifferentiated fever, and peripheral blood examination may sometimes demonstrate the infectious agent as clusters of small, coccoid to coccobacillary bacteria sequestered within a vacuole in either monocytes or neutrophils. After its identification in humans in 1990, *Ehrlichia chaffeensis*, a new species, was cultivated from an infected patient and identified as the main cause of this infection (191). Since then, seroepidemiologic studies indicate that *E. chaffeensis* infections are more frequent than RMSF in some geographic regions (194).

Another *Ehrlichia* spp was implicated in 1994 as the cause of disease in the upper Midwest that was characterized by intracellular inclusions, called morulae, that were only in blood neutrophils (192, 195). Serologic and molecular studies showed that the cause was not *E. chaffeensis*, and molecular studies identified DNA sequences nearly identical to veterinary ehrlichiae known to infect horse or ruminant neutrophils. These findings were confirmed by serologic tests, and data now indicate that the causative agent and the veterinary neutrophil pathogens, *E. equi*, and *E. phagocytophila*, represent variants of a single species. The disease was called human granulocytic ehrlichiosis or HGE, to distinguish it from infection with *E. chaffeensis*, then designated human monocytic ehrlichiosis (HME), that most often involved monocytes in blood. Recently, another *Ehrlichia* spp., *E. ewingii*, previously known as a canine pathogen, has been implicated as a cause of febrile illness in humans (193). It is a serologic and close genetic relative of *E. chaffeensis*, but resides in neutrophils. It is likely that other *Ehrlichia* species yet to be discovered may contribute to fever that occurs after tick bites.

Although these small obligate intracellular bacteria have a gram negative-type cell wall by ultrastructure, gram-staining is not useful for identification. Phylogenetic analysis of 16S ribosomal RNA and groesl genes, serologic and protein investigations, as well as vector relationships can be used to divide the *Ehrlichia* genus into three distinct clades, noted here as the *E. canis*, *E. sennetsu*, and *E. phagocytophila* groups (196). These bacteria parasitize specific derivatives of mammalian bone marrow-derived cells; *E. chaffeensis* and *E. sennetsu* infect mononuclear

phagocytes, and the *E. phagocytophila* group and *E. ewingii* infect neutrophils.

Epidemiology

Distributed mostly in southeastern and south central states, *E. chaffeensis* infections are found wherever the Lone Star tick, *Amblyomma americanum* may bite humans (197). Moreover, infections identified in California, Washington, New York, Massachusetts, and Connecticut, as well as potential infections in Europe, Africa, and the Far East suggest a wider array of potential vectors or serologically similar agents. HGE has been documented mostly in the upper Midwest and northeast United States, but infected patients have been found in California, the mid-Atlantic, south Atlantic, and south central United States, and broadly across Europe (197,198 and 199). *Ixodes persulcatus* group ticks, including *I. scapularis* (black-legged or deer tick) in the eastern United States, *I. pacificus* (western black-legged tick) in the west, and probably *I. ricinus* (sheep tick) in Europe are important vectors (198). Consequently, co-transmission and co-infections with *Borrelia burgdorferi* (Lyme disease) and *Babesia microti* (human babesiosis) are well recognized. The geographic distribution of *E. ewingii* is incompletely documented, but transmission by *A. americanum* ticks would predict human infections to occur in the same regions where HME is identified (193).

The mean age of patients with HME and HGE is more than 44 years in most series, unlike the situation for RMSF or Lyme disease (198). In spite of this, severe and fatal infections have been documented in the pediatric age group. As expected, most infections are identified during May through September when the small nymphal stage ticks are at peak activity and human outdoor exposure is maximal; a second peak of HGE occurs in late October through December when the adult stages of the vector ticks are most active.

Unlike the situation for *Rickettsia* spp., *Ehrlichia* species are maintained by horizontal transmission (tick to mammal to tick). For *E. chaffeensis*, the white-tailed deer (*Odocoileus virginianus*) is a major reservoir host (200), while a major reservoir for the *E. phagocytophila* group in the eastern United States is the white-footed mouse, *Peromyscus leucopus* (201,202). The role that white-tailed deer and domestic ruminants play in the ecology of these infectious agents is not clear.

Because of this horizontal transmission, persistent mammalian infection is important and is well recognized in dogs (*E. canis*, *E. ewingii*, and *E. chaffeensis*), white-tailed deer (*E. chaffeensis*), ruminants (*E. phagocytophila*), and hosts for other ehrlichial species. However, persistent infections appear to be very infrequent in humans.

Pathology and Clinical Manifestations

Despite the clinical similarities with the vasculotropic rickettsioses, the pathogenesis of human monocytic and granulocytic ehrlichiosis does not involve endothelial cell infection or vasculitis (198). Pathologic findings show mild perivascular lymphohistiocytic infiltrates, hepatocyte apoptoses, mild lobular hepatitis, increases in mononuclear phagocytes in spleen, lymph node, liver, and bone marrow, occasional erythrophagocytosis, and granulomas of liver and bone marrow with *E. chaffeensis* infections. Despite the frequency of leukopenia, anemia, or thrombocytopenia, most bone marrow samples are normocellular or hypercellular. In HME, a mononuclear cell CSF pleocytosis can be identified. Gastrointestinal and pulmonary hemorrhages, diffuse alveolar damage with organizing pneumonia, and nosocomial or opportunistic infections may occur as severe complications of HME or HGE.

The underlying pathogenetic factor in HME and HGE appears to involve diffuse increases in mononuclear phagocytes with granuloma formation (in HME) or infiltrates of histiocytes that are activated for increased nonspecific phagocytic activity. As a result, moderate to profound leukopenia and thrombocytopenia are frequent manifestations that can not be explained on the basis of destruction of infected cells. Hepatic or other organ-specific injury probably is not related to direct infection and may result from poorly regulated release of deleterious cytokines at sites of inflammatory cell recruitment. A similar mechanism may underlie diffuse alveolar damage that occasionally is seen and that results in adult respiratory distress syndrome (ARDS) (203). Inflammation and edema with meningoencephalitis is a serious complication that seems restricted to HME.

HME, HGE, and infection by *E. ewingii* are distinct processes caused by different *Ehrlichia* species that present with similar clinical features (193, 198). After a post-tick bite incubation period of approximately 7 to 10 days, the usual presentation is undifferentiated fever, with headache, myalgia, and malaise. Other findings attributable to involvement of the gastrointestinal, respiratory, renal, or central nervous systems occur in the minority of patients, but are not rare. Rash, including macular or maculopapular, and rarely petechial lesions, occurs infrequently in HME and rarely in HGE. Ordinarily, illness lasts approximately 4 to 12 days, and hospitalization is often required. The high prevalence of *E. phagocytophila* group antibodies and asymptomatic seroconversions after intense tick exposure for *E. chaffeensis* suggest that frequent mild or subclinical infections may occur.

Patients with HME and HGE often present with leukopenia, lymphopenia, neutropenia, or thrombocytopenia during the first 7 days of illness (204, 205). Thereafter, leukocyte components normalize while platelet counts may remain depressed for 2 weeks or longer. Bone marrow examinations often reveal cellular or reactive bone marrows despite pancytopenia (206). In HME, 75% of bone marrows contain granulomas and granulomatous inflammation, a finding not identified in patients with HGE. Mild to moderate elevations in hepatic aminotransferase activities are frequently found as evidence of underlying hepatic injury. Adult patients infrequently develop acute renal failure associated with concurrent rises in serum urea nitrogen or creatinine levels. A syndrome similar to disseminated intravascular coagulation (DIC) characterized by both prolonged activated partial thromboplastin time/prothrombin time and hypofibrinogenemia can occur in HME and HGE.

Diagnostic Procedures

Many patients with HME and HGE are identified presumptively by the recognition of typical *Ehrlichia morulae* in peripheral blood leukocytes. While some view this as too infrequent to be useful for diagnosis of HME, an intense search will sometimes

identify a very small number of infected cells (Fig. 53.8). By comparison, between 20% and 80% of HGE patients have peripheral blood neutrophils with typical *Ehrlichia morulae* (Fig. 53.9) that are found in between 0.5% and 40% of cells (198). However, morphologic analyses can be difficult and misleading; thus, an experienced microscopist is required for this evaluation. A more definitive study and identification can be achieved by PCR amplification of nucleic acids specific for each species or by demonstrating an increasing titer of *E. chaffeensis* or *E. phagocytophila* group antibodies during convalescence (197, 198, 207, 208). Currently, a diagnosis is based upon a clinically compatible disease with laboratory evidence of *Ehrlichia* spp. infection by (i) seroconversion using specific *E. chaffeensis* or *E. phagocytophila* group serologic tests, (ii) PCR identification of specific *E. chaffeensis* or *E. phagocytophila* group nucleic acids and specific serologic evidence of infection by a valid serologic method, or (iii) PCR identification or specific serologic evidence of infection and identification of morulae in appropriate peripheral blood leukocytes. Patients with HGE and HME develop cross-reactive serologies in between 12% and 30% of cases; thus, serodiagnosis mandates tests using both *E. chaffeensis* and *E. phagocytophila* group antigens. PCR amplification performed during the acute phase may be as sensitive as 86% under optimal circumstances (207, 208). Culture of *Ehrlichia* spp. has been performed only in research and public health laboratories with a strong interest. However, the ease of cultivation of *E. phagocytophila* group ehrlichiae (Fig. 53.10) and recent improvements in *E. chaffeensis* cultivation (Fig. 53.11) suggest a diagnostic role in the future. No evidence exists to classify *E. chaffeensis* or *E. phagocytophila ehrlichiae* as highly biohazardous. Because both HME and HGE have caused fatalities, therapy should be instituted based upon clinical evaluation and not withheld while awaiting laboratory confirmation.

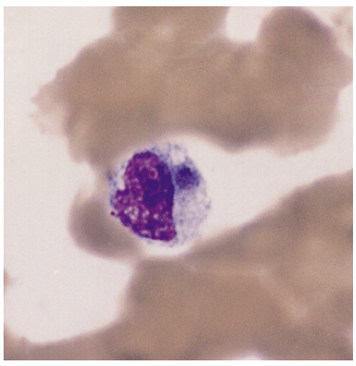


FIGURE 53.8. *Ehrlichia chaffeensis* in a peripheral blood monocyte from a patient with fatal human monocytic ehrlichiosis (HME) in Illinois. Note the round intracytoplasmic inclusion with stippling. Photograph courtesy of Joan Barenfanger, M.D.; Wright stain, original magnification 1,000×.

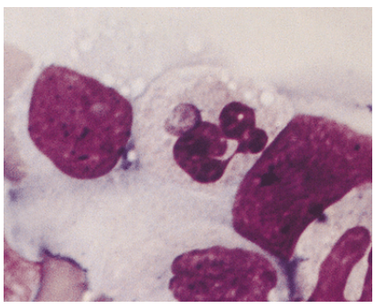


FIGURE 53.9. Peripheral blood neutrophil from a Wisconsin patient with human granulocytic ehrlichiosis (HGE). The organism that causes HGE is almost identical to *Ehrlichia phagocytophila* and *Ehrlichia equi*. Note the intracytoplasmic inclusion with stippling. Wright stain, original magnification 1,000×.

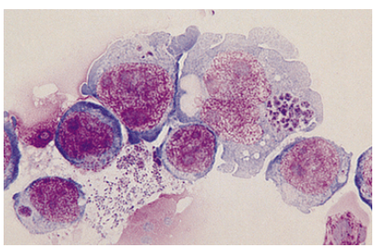


FIGURE 53.10. *Ehrlichia phagocytophila* group (HGE agent) cultivated *in vitro* in human myelomonocytic (THP-1) leukemia cells. Note the large intracellular aggregates of coccoid bacteria confined within a vacuole. Romanowsky stain, original magnification 1,000×.

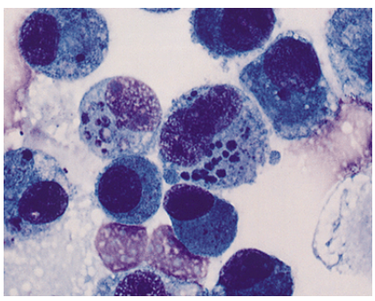


FIGURE 53.11. *Ehrlichia chaffeensis* cultivated *in vitro* in a macrophage cell line (canine DH82 cells). Note the multiple, large, round intracytoplasmic aggregates of bacteria. Romanowsky stain, original magnification 1,000×.

Fatalities with HME and HGE occur in between 0.5% and 2% of all infections documented (209). Worrisome complications include respiratory distress, meningoencephalitis, a toxic-shock or septic-shock like illness, and multiorgan failure. Long term neurologic injury may be a consequence of severe or untreated infections. Immune compromise, including infection with HIV, corticosteroid therapy, pre-existing malignancies, or transplantation seem to predispose to more severe or fulminant infection and the potential for a fatal outcome.

Therapy

As for RMSF, HME and HGE can be effectively treated with tetracyclines, and most patients improve significantly within 48 hours. Although chloramphenicol therapy was associated with a significantly shorter interval of fever and hospitalization than

non-tetracycline treated patients with HME, both *E. chaffeensis* and *E. phagocytophila* group ehrlichiae are resistant to its effects *in vitro* (204, 210, 211). Rifampin has been used with success in individual HGE cases, but no comprehensive studies have proven its clinical utility for HME or HGE. Other broad-spectrum antibiotics such as ß-lactams, aminoglycosides, and macrolides are not effective *in vitro*, and current *in vivo* data do not support a role for their therapeutic use. Antimicrobial prophylaxis after tick bite for HME and HGE has not been evaluated.

Coxiella burnetii

Coxiella burnetii is the causative agent of Q fever, a febrile disease that often lacks rash, and may present in acute or chronic forms. The acute form of the disease often begins as an influenza-like illness, occasionally with interstitial pneumonitis, or granulomatous hepatitis (212). Chronic Q fever usually is culture-negative endocarditis (213). Although prevalent on a worldwide basis, Q fever is diagnosed infrequently in the United States. The usual mechanism for transmission is via aerosols or ingestion of contaminated dairy products; arthropod vectors are rarely implicated. Persons at risk include those who are immunocompromised, those with underlying cardiac valve or vascular damage or cardiac prostheses, farmers, abbatoir workers, and others who are exposed to farm animals or products (213, 214). The relative risk for infection increases with age.

Epidemiology

It is likely that many more cases of Q fever occur each year than are diagnosed in the United States owing to the nonspecific clinical manifestations and lack of availability of specific laboratory diagnostic tests. Regardless, *C. burnetii* is implicated in 0.5% to 3% of cases of serologically investigated acute respiratory illnesses or hepatitis in some areas of the United States, and may account for 40% of atypical pneumonias in Japanese pediatric patients (215, 216). Cattle, sheep, goats, parturient cats, and wild animals such as rabbits, are reservoirs for *C. burnetii*. This bacterium is more closely related to *Legionella pneumophila* and the arthropod symbiont *Francisella* (formerly *Wolbachia*) *persica* than to *Rickettsia* and *Ehrlichia* spp. (143). The organism is considered highly contagious for humans and animals, as even a single organism may establish infection. *C. burnetii* is highly stable under a variety of environmental conditions owing to a spore-like form. It is transmitted by inhalation of aerosols generated by dust, straw, or cloth contaminated with birth tissues, by processing of contaminated animal products, or by consuming raw dairy products. Although animal exposure is the major risk factor in Europe, the United States, and Australia, urban acquisition, without concomitant animal exposure, occurs at a high rate in France (213). Transmission from infected human products of conception places obstetric health care workers at high risk in endemic regions (217). Immunocompromise after cancer therapy, HIV infection, organ transplantation, hemodialysis, cirrhosis, or chronic granulomatous disease is a risk factor in over 20% of patients with acute or chronic Q fever (218).

Pathology and Clinical Manifestations

In contrast to other rickettsial infections, humans acquire *C. burnetii* predominantly after inhalation of infectious aerosols. The most frequent presentations of acute Q fever include an influenza-like respiratory illness, a granulomatous illness with hepatitis, and an undifferentiated febrile illness. Pulmonary infection elicits a mild interstitial lymphocytic pneumonitis with dense macrophage-rich intraalveolar exudates heavily infected with *C. burnetii* (219). The infection may elicit granulomas in liver, bone marrow, and other organs, all signs of an acute and usually self-limited infection (220). Chronic Q fever endocarditis is characterized by macrophage and lymphocyte-rich infiltrates in necrotic fibrinous valvular vegetations (Fig. 53.12) and by the absence of granulomas (221).

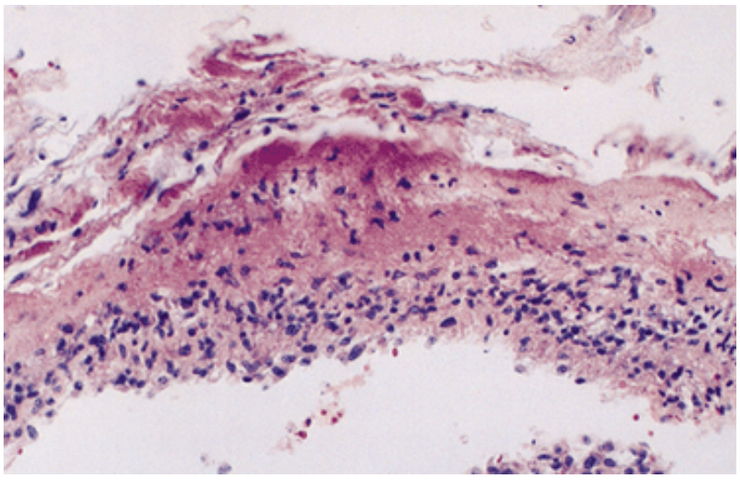


FIGURE 53.12. Histopathology of Q fever endocarditis reveals focally dense lymphohistiocytic infiltrates associated with fibrin deposition and necrosis. Note the scarcity of purulent inflammation that can also accompany *Coxiella burnetii* infection of cardiac valves. H&E, original magnification 400×.

As with many other obligate intracellular pathogens, recovery from acute infection may result in nonsterile immunity. It is

likely that development of debilitating chronic Q fever occurs after recovery from mild or inapparent acute Q fever after which the agent is not cleared (221). The persistence of *C. burnetii* in tissue macrophages at sites of pre-existing tissue damage (Fig. 53.13) causes a low-grade smoldering inflammation, which eventually leads to irreversible cardiac valve damage or persistent vascular injury.

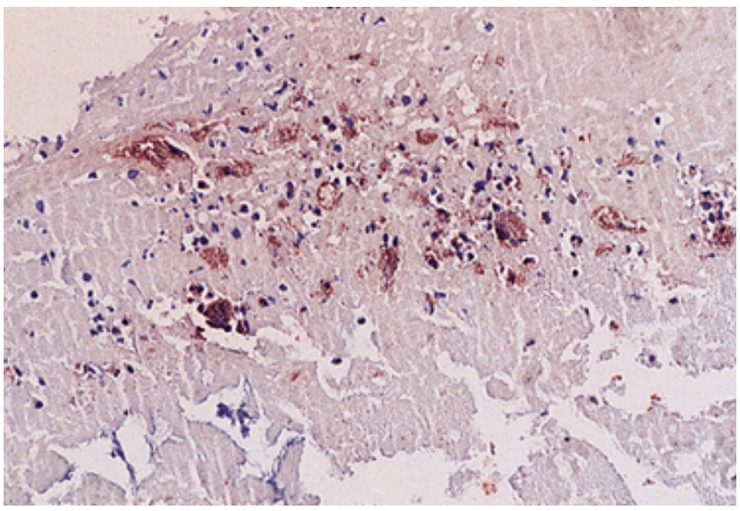


FIGURE 53.13. Immunoperoxidase demonstration of *Coxiella burnetii* in Q fever endocarditis. Note the presence of coccoid bacteria within histiocytes associated with the necrotic vegetation. Immunoperoxidase, original magnification 400×.

Acute Q fever develops approximately 3 weeks (range, 14–39 days) after exposure to the causative agent (212). The severity of illness ranges from subclinical infection to a febrile systemic illness with severe frontal headache, arthralgia, and myalgia often accompanied by respiratory symptoms. In adults, the pneumonia usually resembles primary atypical or viral pneumonitis or legionnaires' disease with a nonproductive cough. Other prominent clinical findings that may lead to diagnostic confusion include fatigue, vomiting, abdominal pain, and meningismus. Hepatomegaly and splenomegaly may be detected in some patients. Laboratory findings may reveal leukopenia with a left shift (>5%) in 50% of patients and thrombocytopenia, which is infrequent. In a review of 170 adult French patients with acute Q fever, 81.8% had fever, and of these patients 62% had elevated serum hepatic transaminase levels, 46% had respiratory involvement, 18% had cutaneous findings, and 11% had neurologic findings that required lumbar puncture for evaluation. Nearly 20% of these patients were afebrile and presented with one or more of hepatic, pulmonic, cutaneous, and/or neurologic findings (212). Other series of patients indicate that pulmonary involvement is the most frequent manifestation. Radiographically, the pulmonary consolidations become round and resolve slowly. Acute Q fever is a self-limited illness that lasts for 2 to 3 weeks. Severe infections including acute encephalopathy with impairment of consciousness and an abnormal pattern on electroencephalogram and computed tomographic scans of the brain have been reported.

Risk for development of chronic Q fever is strongly correlated with advancing age. Chronic Q fever tends to be recalcitrant to therapy and often (23% to 65%) results in death. Endocarditis, usually seen in damaged or prosthetic valves, may occur months to years after acute Q fever or in the absence of any history of Q fever (221). Chronic Q fever less frequently presents as infections of vascular prostheses and aneurysms, osteomyelitis, myocarditis, undifferentiated fever, pneumonia, hepatitis, or an isolated purpuric rash. In Q fever endocarditis, fever may be absent in up to 15% of cases, and more than 75% of all identified patients have congestive heart failure (213). Other frequently observed features include marked clubbing of the fingers, hepatomegaly, and splenomegaly.

Frequent laboratory abnormalities in chronic Q fever include an erythrocyte sedimentation rate greater than 20 mm/hour in 80% of cases, hypergammaglobulinemia in 54% of cases, and hyperfibrinogenemia in 67%. Leukocytosis, leukopenia, thrombocytopenia, or anemia are seen in a minority of patients (212, 213). The presence of rheumatoid factor in more than 50% of cases, circulating immune complexes in nearly 90%, and the frequent findings of anti-platelet antibodies, anti-smooth muscle antibodies, anti-mitochondrial antibodies, circulating anticoagulants, and positive direct Coombs test may suggest an autoimmune process (222). A variety of other syndromes have been associated with Q fever, including meningoencephalitis, inflammatory pseudotumor of the lung, glomerulonephritis, immune complex vasculitis, hemolytic anemia, and autoimmune disorders.

Diagnostic Procedures

Although infrequently diagnosed, Q fever should be considered in patients with fever of unknown origin or culture-negative endocarditis who live in rural areas or who are in close contact with domestic livestock, cats, or their products of conception. The diagnosis of Q fever can be confirmed serologically by significant increases in indirect fluorescent antibody titers to phase I and phase II antigens (223). The inability of the complement fixation test to discriminate between recent and remote infection diminishes its usefulness when acute-phase sera are not available. Elevated or rising titers of phase II antibody alone are characteristic of acute Q fever, and the appearance and persistence of elevated titers of phase I and phase II antibody are indicative of chronic Q fever. Elevated titers of phase I immunoglobulin (Ig) A antibody are reported to be diagnostic for Q fever endocarditis (224); however, one evaluation showed that a phase II IgG titer of 200 or greater is indicative of *C. burnetii* infection and that a phase I IgG titer of less than 800 is inconsistent with chronic Q fever (223).

Cultivation and antimicrobial susceptibility testing of *C. burnetii* is not difficult and requires a traditional virology laboratory (225); however, some consider *C. burnetii* to be highly biohazardous and recommend cultivation only by trained personnel in specialized biohazard facilities. Diagnosis by culture can be performed using a variety of cell lines and primary cultured cells. Alternative methods for diagnosis include PCR amplification of *C. burnetii* genomic or plasmid targets in blood, valve tissues, and perhaps other infected tissues such as placenta (226). Immunohistologic demonstration of *C. burnetii* by using polyclonal or a monoclonal antibody has a sensitivity of approximately 67% as compared with culture, and is useful when only fixed tissues are available (221).

Therapy

Most patients with Q fever experience a self-limited illness that should be empirically treated within 3 days of onset with tetracycline or doxycycline if suspected (227). Later therapy has little effect on the course of the acute infection. Ofloxacin and pefloxacin are effective and success with a combination of pefloxacin and rifampin is achieved with prolonged (16 to 21 days) therapy. ß-lactams are not effective, but some success with chloramphenicol, trimethoprim-sulfamethoxazole, and ceftriaxone have been reported for individual cases. Prednisone was reported to provide additional clinical benefit in cases of hepatitis with “autoimmune” laboratory findings.

For chronic Q fever, especially endocarditis, prolonged therapy is mandatory. This usually requires the bacteriostatic drugs tetracycline or doxycycline in combination with bactericidal drugs such as rifampin, ofloxacin, or pefloxacin. The use of lysosomotropic alkalinizing agents such as hydroxychloroquine may aid in maintaining activity of pH-sensitive antimicrobial agents in the phagolysosomal environment of *C. burnetii* (228). For patients with heart failure, valve replacement may be warranted and should be accompanied by an effective antibiotic regimen to avoid reinfection of the prosthetic valves (213). Therapy should be monitored by serologic evaluation every 4 months; phase I titers less than 200 for IgG and the absence of an IgA titer indicate cure. Even with this evaluation, cure of chronic Q fever in less than 2 years is unlikely, and thus therapy should be continued for at least 3 years.

Bartonelloses

Perhaps few other bacterial genera have undergone as much change in nomenclature, description of new species, recognition of clinical manifestations, and discovery of zoonotic potential than the genus *Bartonella*. 10 years ago, the genus *Bartonella* contained only a single species, Bartonella bacilliformis, the agent of an intriguing vector-borne infection with two clinical phases, hemolytic anemia and cutaneous vascular proliferations, and found only at certain altitudes in parts of the Andes mountains. Similarly, the genus *Rochalimaea* contained only a single human pathogen, *R. quintana*, the cause of trench fever. With the application of molecular phylogenetic methods, the emergence of AIDS, rigorous examinations of tissues from patients with cat scratch disease (CSD), and the recovery of previously reemerging and newly emerging bacterial pathogens, the number of species classified in the genus *Bartonella* now is approximately 15 (229, 232). In parallel, the clinical consequences of infection also have markedly broadened to include febrile bacteremia, chronic lymphadenopathy, cutaneous and visceral vascular lesions, multifocal abscesses and tissue inflammatory lesions, and endocarditis, among other less frequently cited abnormalities (233).

Bartonella species are members of the α_2-subgroup of *Proteobacterium*. They are closely related to the mammalian pathogens in the genera *Afipia* and *Brucella*, and to the plant pathogens *Agrobacterium* (234). Molecular studies led to the reclassification of the genera *Rochalimaea* and *Grahamella* as *Bartonella*, and the removal of the family Bartonellaceae from the order Rickettsiales (235). This genus *Bartonella* is characterized by small Gram-negative rods that are fastidious, aerobic, and in some cases, facultative intracellular organisms. Unlike members of the order Rickettsiales, *Bartonella* spp. can be cultivated axenically on agar supplemented with blood or hemin in air or 5% to 10% CO_2. Flagella have been identified in *B. bacilliformis* and *B. clarridgeiae*, but pili that may be important mediators of host cell attachment are present in both *B. henselae* and *B. quintana*. Bacteriophage-like particles have been found associated with both *B. bacilliformis* and *B. henselae*.

The pathogenetic mechanisms of these species are incompletely described. Erythrocyte invasion by *B. bacilliformis* depends upon propulsion from the flagellum and bacterial production of the protein deformin (236). Inhibition of deformin's ability to mediate attachment to and deformation of host erythrocytes does not inhibit lysis. *Bartonella henselae* apparently attaches to endothelial cells via its 43 kDa surface porin protein (237). After attachment, *B. bacilliformis* and perhaps other *Bartonella* spp., invade host cells by virtue of invasion-associated locus (ialA and ialB) proteins (149). Hemin is required for growth, and its acquisition in *B. quintana* is mediated by a tail-specific protease, encoded by ctpA that apparently renders the cell highly permeable to hemin and hemin-like molecules (238). ialA, ialB, ctpA, and filA (a gene encoding the filament-A protein) are clustered together akin to virulence gene pathogenicity islands observed among Enterobacteriaceae (239). Both *B. bacilliformis* and *B. henselae* produce proteins that are angioproliferative (240).

Bartonella spp. often occupy an intra-erythrocytic location in felids, canids, rodents, birds, fish, etc., and may establish persistent bacteremia in the absence of overt clinical signs. Arthropod vectors may be important in maintenance of naturally infected vertebrate hosts or in transmission to man. *B. bacilliformis* and *B. quintana* are transmitted to humans via sandflies (*Lutzomyia* spp.) and body lice (*Pediculus humanus*), respectively, and humans are the only known reservoirs for these bacterial pathogens (241, 242). In addition, cat fleas (*Ctenocephalides felis*) often are naturally infected and are proven vectors of *B. henselae* among cats (243). The association of some *Bartonella* spp. with cats, rodents, or other animals in the absence of clinical signs, and the detection of *Bartonella* spp. in a variety of ectoparasites such as fleas, lice, and ticks, suggest a role for horizontal (mammal-arthropod-mammal) transmission in the natural maintenance cycle (244,245,246 and 247).

Epidemiology

Domestically acquired *Bartonella* spp. infections are relatively prevalent. It is estimated that approximately 20,000 cases of CSD occur in the United States per year or approximately 10 cases per 100,000 population nationally (248). Infection in immunocompromised individuals seems to be less frequent, but good statistical data are lacking. Cat exposure, in particular kitten bites and scratches, are common variables among *Bartonella henselae* infections of immunocompetent and immunodeficient persons (249, 250).

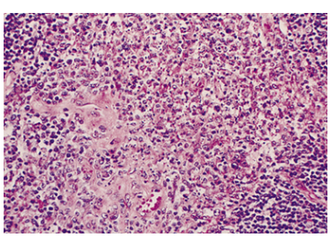


FIGURE 53.14. Early histologic lymph node changes in cat scratch disease caused by *Bartonella henselae*. Note the vague rim of granulomatous inflammation surrounding a region of karyorrhexis and necrosis. Bacilli are most likely to be detected by staining techniques in early lesions such as this. H&E, original magnification, 400×.

Specific epidemiologic associations are dependent upon the *Bartonella* spp. For example, *B. bacilliformis* has been identified only in association with sandfly bites in South America. Previously, infections had only been documented at specific altitudes

in the Peruvian, Colombian, and Ecuadorian Andes mountains; however, recent investigations have clearly shown disease acquisition in the low jungle and adjacent areas and the likely existence of mild or subclinical disease forms (251).

B. quintana long has been associated with lice and conditions of poor hygiene (246, 247, 252). Homeless persons, alcoholics, persons with HIV infection, and injection drug users are at increased risk for infection, bacteremia, and endocarditis. A role for cats has been speculated as some patients with *B. quintana*-associated bacillary angiomatosis or *peliosis hepatis* have been cat owners or had cat exposures; however, transmission of *B. quintana* to cats and fleas has not yet been demonstrated. In contrast, ample evidence exists to show that *B. henselae* is flea-borne among cats and that persons with *B. henselae*-associated CSD, bacillary angiomatosis, bacillary peliosis, febrile bacteremia, and endocarditis have significantly high rates of cat exposure, including cat bites and scratches (233, 249, 250). CSD has a bimodal seasonal distribution with peaks in the fall and early winter, perhaps corresponding to peak flea seasons for cats (253). Up to 40% of infections may occur in adults and a slight male predominance is recorded (253). Occupational exposure (veterinarians) may be a risk factor; however, the high frequency of cat ownership and exposure among the general population precludes more definitive study (254). The highest risk factors for CSD, most frequently caused by *B. henselae*, are ownership of pet kittens, a kitten scratch or bite, and ownership of at least one kitten with fleas (250). Chomel showed that kittens, seropositive cats, and flea-infested cats were more likely to be bacteremic, and prolonged bacteremia in cats has been associated with acquisition of CSD (255, 256). Human infection by other species of *Bartonella*, such as *B. elizabethae*, *B. clarridgeiae*, or *B. vinsonii*, is infrequent enough that specific epidemiological associations cannot be made.

Pathology and Clinical Manifestations

Bartonelloses present with diverse pathologic and clinical findings depending upon species and host immunocompetence. However, infections often share some features owing to similarities in pathogenetic mechanisms and niches. For example, pathogenic species have been shown capable of invasion and survival in erythrocytes of various mammalian species. The three major pathogenic species, *B. bacilliformis*, *B. quintana*, and *B. henselae* will be discussed briefly here.

Bartonella Bacilliformis, Oroya Fever, and Verruga Peruana

After inoculation into the dermis from *Lutzomyia* spp. sandflies, *B. bacilliformis* gains access to the blood. Following an interval of approximately 21 days (1 to 30 weeks) the clinical presentation of febrile hemolytic anemia (Oroyo fever) or Carrion's disease, a condition with cutaneous angioproliferative lesions (*verruga peruana*), may develop (257). Ordinarily, Oroya fever precedes the verrucous stage by about 2 to 20 weeks. Oroya fever results from the direct invasion of erythrocytes and subsequent hemolysis; systemic manifestations include malaise, somnolence, anorexia, myalgia, headache, arthralgia, chills, dyspnea, and fever. Examination of blood smears shows that up to 100% of erythrocytes may contain intracellular bacilli. Severe complications include pericardial effusion, myocarditis, coma, convulsions, delerium, acute respiratory distress, anasarca, or abortion. The case fatality rate ranges from 88% in epidemics to 8% in hospital-managed patients (257). Opportunistic infections, especially salmonellosis, account for many of the fatalities. Simultaneous with the hemolytic process, the bacteria infect endothelial cells (258). Their persistent presence may later lead to the secondary verruga peruana that result from the stimulation of vascular proliferations by a *B. bacilliformis* angioproliferative factor (240).

Bartonella Infections in Immunocompetent Hosts

Cat Scratch Disease. The most frequent Bartonella infection is cat scratch disease. This most often results from *B. henselae* infection after cat exposure. CSD generally is a benign self-limited condition, characterized by prolonged lymphadenopathy, with or without fever (233, 259). However, severe complications may occur including Parinaud's oculoglandular syndrome, encephalitis/encephalopathy and febrile seizures, myelitis, peripheral neuropathy, neuroretinitis, granulomatous hepatitis or splenitis, erythema nodosum, disseminated abscesses, and osteolytic bone lesions (233). In Arkansas children, 5% of all fevers of unknown origin were found to be caused by *Bartonella henselae*, and *B. henselae* has been implicated as a cause of status epilepticus in CSD complicated by encephalopathy (260, 261).

Approximately 1 week after inoculation of *Bartonella* into a scratch or bite wound, a papule develops and persists for 1 to 3 weeks. At about 2 weeks post-inoculation, adenopathy develops in the draining lymph nodes, most often axillary, cervical, and submandibular regions. Adenopathy may occur in other sites, but generalized adenopathy is rare. Ordinarily, adenopathy resolves within several months, but may persist for 12 to 24 months.

The pathologic manifestations of CSD include dermal and epidermal necrosis associated with a dense neutrophil and macrophage-rich infiltrate in which numerous bacilli can be demonstrated at the inoculation papule (262). Initially, lymph nodes show necrosis (Fig. 53.14), followed with infiltration by neutrophils (stellate microabscesses), and later by a rim of granulomatous inflammation surrounding a suppurative center (231). Similar pathology may be observed in other tissues and organs, occasionally leading to misdiagnoses such as tuberculosis, histoplasmosis, toxoplasmosis, tularemia, and brucellosis. Bacilli are most often observed in early lesions (Fig. 53.15).

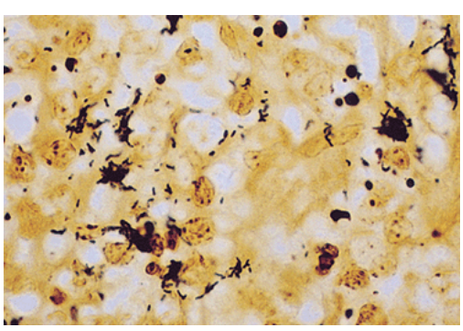


FIGURE 53.15. *Bartonella henselae* in cat scratch disease demonstrated by silver impregnation. This region corresponds to the region of necrosis and karyorrhexis shown in Fig. 11. The organisms were proven *B. henselae* by PCR. Wharthin-Starry stain, original magnification 1,000×.

Bacteremia, Febrile Bacteremia, and Endocarditis

Trench fever is a form of bacteremia caused by *B. quintana* that was first recognized during World War I (252). It occurs 9 to 25 days after inoculation of infected louse feces into skin wounds or abrasions, and leads to a relapsing fever that is accompanied by severe headache and bone pain. Patients who are bacteremic with *B. quintana* but do not have classical trench fever are often afebrile but also may have headache and bone pain. In Brouqui's report, 5 of 10 bacteremic patients had positive blood cultures for 2 to 6 weeks (247). Bacteremia with *B. quintana* and *B. henselae* may occur with or without simultaneous endocarditis.

Endocarditis caused by *B. quintana* is most often associated

with homelessness and alcoholism in patients with no history of valvular injury; however, *B. henselae* endocarditis is observed most often in patients with previously damaged valves and cat exposure (252). Echocardiographic evaluation frequently is unrevealing. The pathology of endocarditis is similar to that observed with other bacterial causes, including fibrin deposition, necrosis, and inflammatory cell infiltrates (Fig. 53.16) (263); however, most valve vegetations have large numbers of bacilli identified by Wharthin-Starry stains (Fig. 53.17). Specific bacterial identification can be achieved by immunohistologic studies (264).

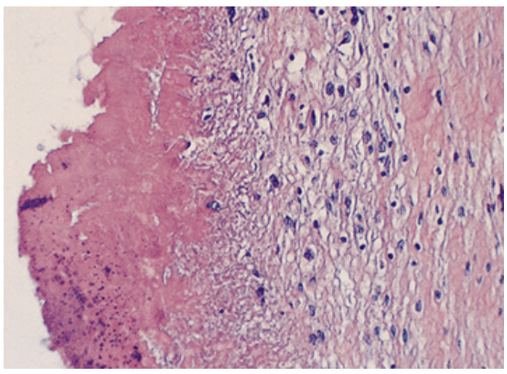


FIGURE 53.16. Histologic findings in *Bartonella quintana* endocarditis. Note the typical features of endocarditis with vegetations that contain fibrin, mixed inflammatory cells, and necrotic tissue. The majority of the acellular vegetation corresponds to the location of the bacilli by special stains. H&E, original magnification 400×.

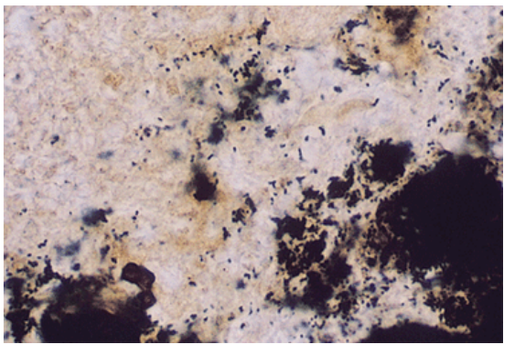


FIGURE 53.17. *Bartonella quintana* in a cardiac valve vegetation with endocarditis. Note the abundant bacilli that cluster in regions corresponding to the fibrinous necrotic acellular material in Fig. 13. Wharthin-Starry stain, original magnification 1,000×.

Bartonella Infections in Immunocompromised Hosts

Bacillary Angiomatosis (BA) and Peliosis

Both *B. quintana* and *B. henselae* may cause unique lesions when patients with HIV infection or other immunocompromising conditions exist. A frequent and dominant manifestation is the development of angioproliferative lesions similar to those in *verruga peruana*, but called bacillary angiomatosis (BA) (Fig. 53.18) (265). Vascular lacunes called peliosis may occur in visceral organs (liver, spleen) or in bone (ostelytic lesions), each of which contain abundant bacilli in the interstitium in close proximity to endothelial cells when stained by the Wharthin-Starry method (Fig. 53.19) (266, 267). BA in HIV-infected patients typically occurs when CD4 counts fall below 100 cells/mm^3 (268). Rare consequences of *Bartonella* infection in immunocompromised hosts include febrile bacteremia, endocarditis, involvement of bone marrow, parenchyma of brain, abdominal cavity, cervix, or vulva, and potentially dementia (233).

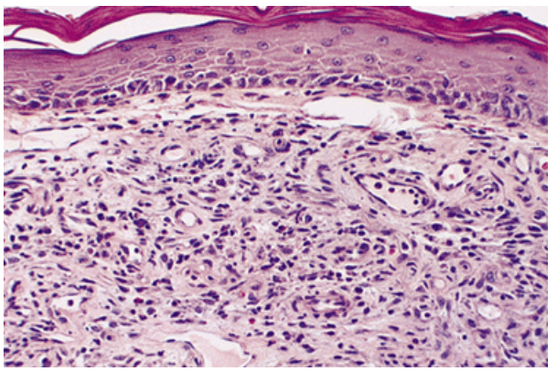


FIGURE 53.18. Histopathology of bacillary angiomatosis caused by *Bartonella henselae* infection in an HIV-infected person. Note the vascular proliferations associated with interstitial mixed inflammation and edema. The histopathologic findings are distinct from those of Kaposi's sarcoma, but do resemble pyogenic granuloma. H&E, original magnification 200×.

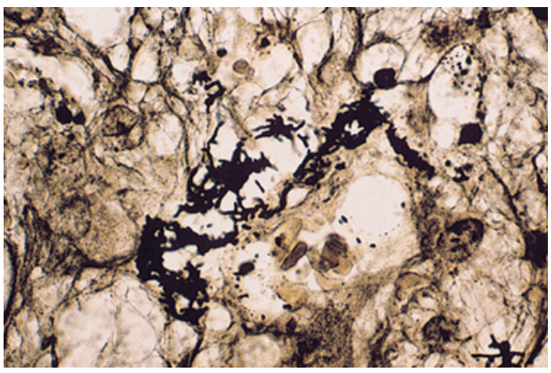


FIGURE 53.19. *Bartonella henselae* in the interstitial spaces of bacillary angiomatosis, corresponding to Fig. 53.18. Note the distinct clusters of bacilli. Wharthin-Starry stain, original magnification 1,000×.

Laboratory Findings

Laboratory evaluation in patients with CSD usually is unremarkable with the exception of an erythrocyte sedimentation

rate that is elevated in 60%. Organ-specific abnormalities (elevated liver enzymes, etc.) may belie focal bacillary involvement. The leukocyte, erythrocyte, and platelet counts are usually normal. In HIV-infected persons, CD4 lymphocyte count and hematocrit are lower than case controls, whereas alkaline phosphatase and aspartate aminotransferase are higher (268).

Diagnosis

Because of the diversity of clinical manifestations and organ involvement, the differential diagnosis for *Bartonella* spp. infections is very wide. However, bartonelloses should be suspected in several characteristic situations in immunocompetent individuals that include persistent lymphadenopathy, fever of unknown origin, encephalopathy with febrile seizures, granulomatous hepatitis, and optical neuritis. Immunocompromised or debilitated patients may develop these findings and should be evaluated for bartonellosis when evidence of cutaneous or visceral organ vascular proliferations, osteolytic bone lesions, or culture-negative endocarditis is detected. Particularly important clinical clues are cat or kitten exposure or louse infestation.

Culture

The diagnosis of *Bartonella* spp. infection is definitively confirmed with *in vitro* cultivation of the bacterium from blood or tissues. Blood samples may be inoculated directly into commercially available vials provided for automated blood culture; however, success has been achieved most frequently using the centrifugation-lysis method (Isolator tubes, Wampole, Cranbury, NJ) or by first freezing EDTA-anticoagulated blood prior to isolation attempts (230, 269,270,271,272,273,274 and 275). A variety of liquid, agar, and cell culture systems have been used for primary isolation and cultivation of *Bartonella* spp. The most frequently recommended method is direct plating of frozen or lysed blood onto agar medium that is enriched for hemin, including Columbia blood, CDC anaerobic blood, chocolate, sheep blood, heart infusion blood, and buffered charcoal yeast extract agars. A chemically defined broth-based medium supplemented with 15 mg/mL hemin has been reported to be more sensitive for primary isolation as compared with solid agar (271). Endothelial cell cultures are also efficient means for isolating *B. henselae* and *B. quintana* (267). In a comparison of methods for isolation of *Bartonella* spp., blind subculture from blood culture broth (BACTEC Plus, Becton Dickinson Microbiology Systems) into endothelial cell shell vial cultures was superior to direct agar inoculation or subcultivation from broth to agar-based medium for recovery of *Bartonella* spp. in endocarditis (269). The same authors also demonstrate that subculture of blood culture broth to agar was 98% sensitive in recovering *B. quintana* in homeless patients with endocarditis, and that recovery of *B. henselae* from lymph nodes in CSD was poor (9% to 13%) regardless of the method used. Because it is impractical to subculture all negative blood cultures, some groups have used acridine orange staining of negative or growth index-suspicious cultures to screen for those that would benefit from further subcultivation (246, 273). Regardless, patients in whom a high degree of suspicion for *Bartonella* infection exists should have cultures held for at least 21 days after which small dry colonies may appear (Fig. 53.20).

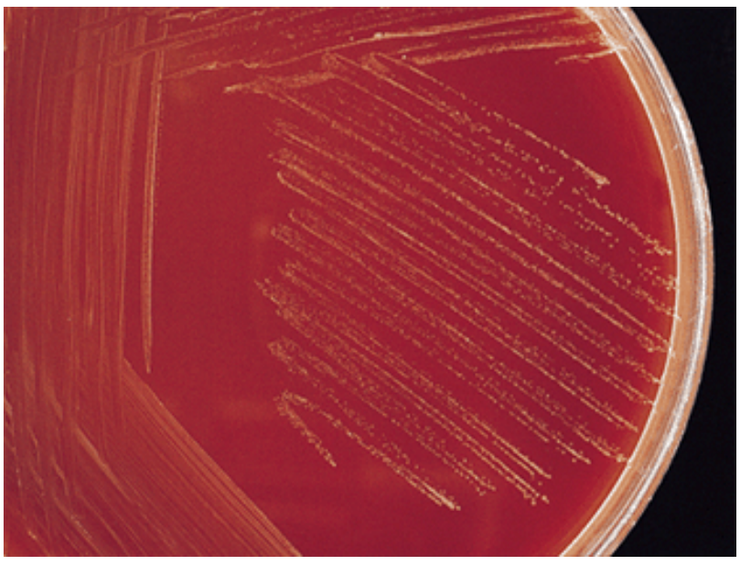


FIGURE 53.20. *Bartonella henselae* cultivated on blood agar usually requires 3 weeks or more to develop small white colonies that adhere to the plate and may be difficult to disperse.

With conventional bacterial identification systems, most *Bartonella* spp. are inert. However, supplementation of the suspension medium with 100 µg/mL hemin has allowed the biochemical differentiation of many *Bartonella* spp. in routinely used laboratory identification systems (276). Other methods for identification and differentiation of suspected *in vitro* isolates include gas-liquid chromatography analysis of cell wall fatty acids, PCR, PCR-RFLP, or sequence analysis of citrate synthase (gltA) or 16S ribosomal RNA genes, and immunofluorescence or other antigenic analyses.

Molecular Diagnosis

The most frequently used molecular method for diagnosis of the bartonelloses is PCR. In general, PCR is a more sensitive diagnostic method than culture and may identify infection in seronegative patients (252). PCR sensitivity tends to be very high for blood obtained from patients with endocarditis (approximately 80%) and in skin biopsies of patients with BA (up to 100%); however, reported sensitivity in lymph nodes from CSD

patients is variable (between 30% and 96%) (252, 269, 277). Little data exist on the utility of direct amplification of *Bartonella* spp. DNA from blood of bacteremic patients, but quantitation of CFUs from patients suggests that sufficient numbers of bacteria circulate to allow sensitive detection in many cases (230).

Serology

The preparation of specific serologic reagents has revolutionized the diagnosis of CSD and *Bartonella* infections (253). The most frequently used assay is the indirect fluorescent antibody (IFA) test, but enzyme immunoassays and western immunoblotting have also been applied for diagnosis (232, 248, 253, 277,278,279 and 280). Several studies now confirm that IFA for detection of *B. henselae* IgG antibodies is sensitive (84% to 95%) and specific (93% to 97%), and has a high positive predictive value (62% to 91%) in patients with clinically diagnosed CSD (253, 281, 282). More limited studies of *Bartonella* spp. ELISA indicate similar performance characteristics (279). In CSD, 78% of patients have IgG antibodies detected within the first week after onset of illness, and peak titers were found 8 weeks after onset (253). The IFA test is incapable of differentiation between *B. henselae* and *B. quintana* antibodies (252, 253); insufficient data are available to show an ability to differentiate among infections in humans by other *Bartonella* spp. However, significant titers of anti-*Chlamydia pneumoniae* antibodies are detected in *Bartonella* spp., endocarditis, and high titers of anti-*Bartonella* spp. antibodies can be detected in patients with Q fever endocarditis, potentially leading to misdiagnoses (283, 284).

Although many immunocompromised patients with *Bartonella* infections have detectable antibody titers, there is still controversy about the use of serodiagnosis in this group. Moreover, well-documented cases continue to show a low level of seronegativity in infected individuals.

Histopathology

The typical histopathologic findings of adenopathy in CSD, proliferative and vascular lesions in BA and peliosis, as well as in endocarditis are described above. Necrotizing, stellate microabscesses are no longer required for the diagnosis of CSD (253). When bacilli are identified in characteristic lesions by nonspecific staining methods, a definitive identification cannot be rendered; however, these morphologic features provide strong evidence of *Bartonella* spp. infections in the correct clinical and epidemiological situations. An immunohistologic method for *in situ* antigenic identification of has been described (264).

Therapy

With the exception of Oroya fever caused by *B. bacilliformis*, *Bartonella* spp. infections are rarely fatal. For Oroya fever, chloramphenicol, penicillin, and fluoroquinolones are effective and streptomycin is often used for *verruga peruana*, but *B. bacilliformis* is susceptible to a very broad range of antibiotics *in vitro* (285). Similarly, *B. henselae*, *B. quintana*, *B. elizabethae*, and *B. vinsonii* are highly susceptible to many antibiotics *in vitro*, with some variance in susceptibility to fluoroquinolones (233, 286). Whether *in vitro* susceptibility can be used to predict in vivo efficacy has yet to be determined.

Preferred antimicrobial regimens for BA and peliosis include erythromycin and doxycycline; alternative second-line drugs are tetracycline, minocycline, and azithromycin (233). It is recommended that therapy continue for at least 2 months in immunocompromised persons with BA and for at least 4 months for persons with peliosis or osteomyelitis. If relapse occurs, life-long therapy is advocated.

Despite the success of antimicrobial therapy in immunocompromised patients, the only drugs that provide clinical benefit for CSD include rifampin, ciprofloxacin, gentamicin, sulfamethoxazole/trimethoprim, and azithromycin (259, 287). The length of therapy is not clearly established, and the majority of cases of typical CSD will resolve spontaneously regardless of therapy.

For patients with bacteremia or endocarditis, too little data exist to make well-supported recommendations. Some authors prefer to use doxycycline, erythromycin, azithromycin, or tetracycline for 14 and 6 months in patients with febrile bacteremia and endocarditis, respectively (233). Often, sufficient injury is present to require valve replacement in those with endocarditis.

REFERENCES

1. Moulder JW. A primer for chlamydiae. In: Mardh PA, Holmes KK, Oriel JD, et al., eds. *Chlamydial Infections.* Amsterdam: Elsevier Biomedical, 1982:3–14.

2. Everett KDE, Bush RM, Andersen AA. Emended description of the order *Chlamydiales*, proposal of *Parachlamydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotype genus, revised taxonomy of the family Chlamydiaceae, including a new genus and five new species, and standards for the identification of organisms. *Int J System Bacteriol* 1999;49:415–440.

3. Dhir SP, Hakamori S, Kenny GE, et al. Immunochemical studies on chlamydial group antigen. (Presence of a 2-keto-3-deoxycarbohydrate as immunodominant group). *J Immunol* 1972;109:116–122.

4. Grayston JT, Kuo CC, Wang SP, et al. A new Chlamydia psittaci strain, TWAR, isolated in acute respiratory tract infections. *N Engl J Med* 1986;315:161–168.

5. Campbell LA, Kuo CC, Grayston JT. Characterization of the new *Chlamydia* agent, TWAR, a unique organism by restriction endonuclease analysis and DNA-DNA hybridization. *J Clin Microbiol* 1987;25:1911–1916.

6. Chi EY, Kuo CC, Grayston JT. Unique ultrastructure in the elementary body of *Chlamydia* sp. strain TWAR. *J Bacteriol* 1987;169:3757–3763.

7. Thompson SE, Washington AE. Epidemiology of sexually transmitted *Chlamydia trachomatis* infections. *Epidemiol Rev* 1983;5:96–123.

8. Washington AE, Johnson RE, Sanders LL Jr. *Chlamydia trachomatis* infections in the United States: what are they costing us? *JAMA* 1987;257:2070–2072.

9. Bowie WR, Holmes KK. *Chlamydia trachomatis* (trachoma, perinatal infections, lymphogranuloma venereum, and other genital infections). In: Mandell GL, Douglas RG Jr., Bennett JE, eds. *Principles and Practice of Infectious Diseases*. 3rd ed. New York: Churchill Livingstone, 1990:1426–1440.

10. Ronnerstam R, Perrson K. Chlamydial conjunctivitis in a Swedish population. In: Mardh PA, Holmes KK, Oriel JD, et al., eds. *Chlamydial Infections*. Amsterdam: Elsevier Biomedical, 1982:87–90.

11. Barnes RC. Laboratory diagnosis of human chlamydial infections. *Clin Microbiol Rev* 1989;2:119–136.

12. Tarizzo M, ed. *Field Methods for the Control of Trachoma*. Geneva: World Health Organization, 1973:9.

13. Rubin SJ. Chlamydiae. In: Howard BJ, Klaas J, Rubin SJ, et al., eds. *Clinical and Pathogenic Microbiology*. St. Louis: CV Mosby, 1987:835–842.

14. Wilfert C, Gutman L. Chlamydial infections. In: Feigin RD, Cherry JD, eds. *Textbook of Pediatric Infectious Diseases*. 2nd ed. Philadelphia: WB Saunders, 1987:1867–1877.

15. Berger RE, Alexander ER, Monda GD, et al. Chlamydia trachomatis as a cause of “idiopathic” epididymitis. *N Engl J Med* 1978;298:301–304.

16. Keat A, Thomas BJ, Taylor-Robinson D. Chlamydial infection in the aetiology of arthritis. *Br Med Bull*. 1983;39:168–174.

17. Paavonen J, Kiviat N, Brunham RC, et al. Prevalence and manifestations of endometritis among women with cervicitis. *Am J Obstet Gynecol* 1985;152:280–286.

18. Centers for Disease Control. *Chlamydia trachomatis* infections: policy guidelines for prevention and control. *MMWR* 1985;34(suppl 3):53–74.

19. Washington AE, Gove S, Schachter J, et al. Oral contraceptives, Chlamydia trachomatis infection, and pelvic inflammatory disease: a word of caution. *JAMA* 1985;253:2246–2250.

20. Brunham RC, Paavonen J, Stevens CE, et al. Mucopurulent cervicitis: the ignored counterpart of urethritis in men. *N Engl J Med* 1984;311:1–6.

21. Wang SP, Eschenbach DA, Holmes KK, et al. Chlamydia trachomatis in the Fitz-Hugh-Curtis syndrome. *Am J Obstet Gynecol* 1980;138:1034–1038.

22. Henry-Suchet J, Catalan F, Loffredo V, et al. *Chlamydia trachomatis* associated with chronic inflammation in abdominal specimens from women selected for tuboplasty. *Fertil Steril* 1981;36:599–605.

23. Stamm WE, Wagner KF, Amsel R, et al. Causes of the acute urethral syndrome in women. *N Engl J Med* 1980;303:409–415.

24. Beem MO, Saxon EM. Respiratory-tract colonization and a distinctive pneumonia syndrome in infants infected with *Chlamydia trachomatis*. *N Engl J Med* 1977;296:306–310.

25. Tipple MA, Beem MO, Saxon EM. Clinical characteristics of the afebrile pneumonia associated with *Chlamydia trachomatis* infection in infants less than 6 months of age. *Pediatrics* 1979;63:192–197.

26. Thygeson P, Stone W. Epidemiology of inclusion conjunctivitis. *Arch Ophthalmol* 1942;27:91–122.

27. Bernstein DI, Hubbard T, Wenman WM, et al. Mediastinal and supraclavicular lymphadenitis and pneumonitis due to Chlamydia trachomatis serovars L1 and L2. *N Engl J Med* 1984;311:1543–1546.

28. Clyde WA, Kenny GE, Schachter J. Laboratory diagnosis of chlamydial and mycoplasmal infections. In: Drew WL, ed. *Cumitech 19*. Washington, DC: American Society for Microbiology, 1984.

29. Cles LD, Bruch K, Stamm WE. Staining characteristics of six commercially available monoclonal immunofluorescent reagents for direct diagnosis of Chlamydia trachomatis infections. *J Clin Microbiol* 1988;26:1735–1737.

30. Amortegui AJ, Meyer MP. Enzyme immunoassay for detection of *Chlamydia trachomatis* from the cervix. *Obstet Gynecol* 1985;65:523–526.

31. Chernesky MA, Mahony JB, Castriciano M, et al. Detection of *Chlamydia trachomatis* antigens by enzyme immunoassay and immunofluorescence in genital specimens from symptomatic and asymptomatic men and women. *J Infect Dis* 1986;154:141–148.

32. Howard LV, Coleman PF, England BJ, et al. Evaluation of Chlamydiazyme for the detection of genital infections caused by *Chlamydia trachomatis*. *J Clin Microbiol* 1986;23:329–332.

33. Van Ulsen J, Van Zuuren-van der Horst A, Tjiam KH, et al. Solid-phase enzyme immunoassay for detection of *Chlamydia trachomatis*. *Eur J Clin Microbiol* 1985;4:397–399.

34. Taylor-Robinson D, Thomas BJ, Osborn MF. Evaluation of enzyme immunoassay (Chlamydiazyme) for detecting *C. trachomatis* in genital tract specimens. *J Clin Pathol* 1987;40:194–199.

35. Hipp SS, Yangsook H, Murphy D. Assessment of enzyme immunoassay and immunofluorescence test for detection of *Chlamydia trachomatis*. *J Clin Microbiol* 1987;25:1938–1943.

36. Pothier P, Kazmierczak A. Comparison of cell culture with two different chlamydia tests using immunofluorescence or enzyme-linked immunosorbent assay. *Eur J Clin Microbiol* 1986;5:569–572.

37. Ryan RW, Kwasnik I, Steingrimsson O, et al. Rapid detection of *Chlamydia trachomatis* by an enzyme immunoassay method. *Diagn Microbiol Infect Dis* 1986;5:225–234.

38. Mabey DCW, Robertson JN, Ward ME. Detection of *Chlamydia trachomatis* by enzyme immunoassay in patients with trachoma. *Lancet* 1987;2:1491–1492.

39. Hammerschlag MR, Roblin PM, Cummings C, et al. Comparison of enzyme immunoassay and culture for diagnosis of chlamydial conjunctivitis and respiratory infections in infants. *J Clin Microbiol* 1987;25:2306–2308.

40. Blanding J, Hirsch L, Stranton N, et al. Comparison of the Clearview Chlamydia, the PACE 2 assay, and culture for detection of *Chlamydia trachomatis* from cervical specimens in a low-prevalence population. *J Clin Microbiol* 1993;31:1622.

41. Clarke LM, Sierra MF, Daidone BJ, et al. Comparison of the Syva Microtrak enzyme immunoassay and Gen-probe PACE 2 with cell culture for diagnosis of cervical *Chlamydia trachomatis* infection in a high-prevalence female population. *J Clin Microbiol* 1993;31:968–971.

42. Iwen PC, Blair TMH, Woods GL. Comparison of the Gen-probe Pace 2TM system direct fluorescent-antibody, and cell culture for detecting *Chlamydia trachomatis* in cervical specimens. *J Clin Path* 1991;95:578.

43. Kluytmans JAJW, Niesters HGM, Mouton JW, et al. Performance of nonisotopic DNA probe for detection of *Chlamydia trachomatis* in urogenital specimens. *J Clin Microbiol* 1991;29:2685.

44. Kluytmans JAJW, Goessens WHF, vanRijsoort-Vos JH, et al. Improved performance of PACE 2 with modified collection system in combination with probe competition assay for detection of Chlamydia trachomatis in urethral specimens from males. *J Clin Microbiol* 1994;32:568.

45. Warren R, Dwyer B, Plackett M, et al. Comparative evaluation of detection assays for *Chlamydia trachomatis*. *J Clin Microbiol* 1993;31:1663.

46. Woods GL, Garza DM. Use of Gen-Probe probe competition assay as a supplement to probes for direct detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in urogenital specimens. *J Clin Microbiol* 1996;34:177.

47. Carroll KC, Aldeen WE, Morrison M, et al. Evaluation of the Abbott LCx ligase chain reaction assay for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in urine and genital swab specimens from sexually transmitted disease clinic population. *J Clin Microbiol* 1998;36:1630.

48. Ferrero DV, Myers HN, Schultz DE, et al. Performance of the Gen-Probe amplified *Chlamydia trachomatis* assay in detecting *Chlamydia trachomatis* in endocervical and urine specimens from women and urethral and urine specimens from men attending sexually transmitted disease and family planning clinics. *J Clin Microbiol* 1998;36:3230.

49. Goessens WHF, Mouton JW, vanDer Meijden WI, et al. Comparison of three commercially available amplification assays, AMP CT, LCx, and COBAS AMPLICOR, for detection of *Chlamydia trachomatis* in first-void urine. *J Clin Microbiol* 1997;35:2628.

50. Mahony J, Chong S, Jang D, et al. Urine specimens from pregnant and nonpregnant women inhibitory to amplification of *Chlamydia trachomatis* nucleic acid by PCR ligase chain reaction, and transcription-mediated amplification; identification of urinary substances associated with inhibition and removal of inhibitory activity. *J Clin Microbiol* 1998;36:3122.

51. Puolakkainen M, Hiltunen-Back E, Reunala T, et al. Comparison of performances of two commercially available tests, a PCR assay and a ligase chain reaction test, in detection of urogenital Chlamydia trachomatis infection. *J Clin Microbiol* 1998;36:1489.

52. Wylie JL, Moses S, Babcock R, et al. Comparative evaluation of Chlamydiazyme, PACE 2, and AMP-CT assays for detection of *Chlamydia trachomatis* in endocervical specimens. *J Clin Microbiol* 1998;36:3488.

53. Frei W. On the skin test in lymphogranuloma inguinale. *J Invest Dermatol* 1938;1:367–377.

54. Rake G, McKee CM, Shaffer MF. Agent of lymphogranuloma venereum in the yolk sac of the developing chick embryo. *Proc Soc Exp Biol Med* 1940;43:332–334.

55. Schachter J, Smith DE, Dawson CR, et al. Lymphogranuloma venereum. I. Comparison of the Frei test, complement fixation test, and isolation of the agent. *J Infect Dis* 1969;120:372–375.

56. Wang SP, Grayston JT. Immunologic relationship between genital TRIC, lymphogranuloma venereum, and related organisms in a new microtiter indirect immunofluorescence test. *Am J Ophthalmol* 1970;70:367–374.

57. Wang SP, Grayston JT, Alexander ER, et al. Simplified microimmunofluorescence test with trachoma-lymphogranuloma venereum (*Chlamydia trachomatis*) antigens for use as a screening test for antibody. *J Clin Microbiol* 1975;1:250–255.

58. Yong EC, Chinn SJ, Caldwell HD, et al. Reticulate bodies as single antigen in *Chlamydia trachomatis* serology with microimmunofluorescence. *J Clin Microbiol* 1979;10:351–356.

59. Nettleman MD, Jones RB, Roberts SD. Cost-effectiveness of culturing for Chlamydia trachomatis. A study in a clinic for sexually transmitted diseases. *Ann Intern Med* 1986;105:189–196.

60. Phillips RS, Aronson MD, Taylor WC, et al. Should tests for *C. trachomatis* cervical infection be done during routine gynecologic visits? An analysis of the costs of alternative strategies. *Ann Intern Med* 1987;107:188–194.

61. Trachtenberg AL, Washington AE, Halldorson S. A cost-based decision analysis for chlamydia screening in California family planning clinics. *Obstet Gynecol* 1988;71:101–108.

62. Hammerschlag MR, Chandler JW, Alexander ER, et al. Erythromycin ointment for ocular prophylaxis of neonatal chlamydial infection. *JAMA* 1980;244:2291–2293.

63. Greaves AB, Hilleman MR, Taggart SRT, et al. Chemotherapy in bubonic lymphogranuloma venereum. Clinical and serological evaluation. *Bull WHO* 1957;16:277–289.

64. Bowie WR. Treatment of chlamydial infections. In: Mårdh PA, Holmes KK, Oriel JD, et al., eds. *Chlamydial Infections*. Amsterdam: Elsevier Biomedical Press, 1982:231–244.

65. Handsfield HH, Alexander ER, Wang SP, et al. Differences in the therapeutic response of chlamydia-positive and chlamydia-negative forms of nongonococcal urethritis. *J Am Vener Dis Assoc* 1975;2:5–9.

66. Bowie WR, Alexander ER, Floyd JF, et al. Differential response of chlamydial and ureaplasma-associated urethritis to sulfafurazole (sulfixoxazole) and aminocyclitols. *Lancet* 1976;2:1276–1278.

67. Meyer KF. The ecology of psittacosis and ornithosis. *Medicine* 1942;21:175–206.

68. Potter ME, Kaufmann AF, Plikaytis BD. Psittacosis in the United States 1979. CDC Surveillance Summaries. *MMWR* 1983;32(suppl 1):27–31.

69. Centers for Disease Control. Psittacosis associated with turkey processing-Ohio. *MMWR* 1982;30:638–640.

70. Meyer KF, Eddie B. The influence of tetracycline compounds on the development of antibodies in psittacosis. *Am Rev Tuberc* 1956;74:566–571.

71. Schachter J. Chlamydiae. In: Rose NR, Friedman H, eds. *Manual of Clinical Immunology*, 2nd ed. Washingon, DC: American Society for Microbiology, 1980;701–706.

72. Augenbraun MH, Roblin PM, Chirgwin K, et al. Isolation of Chlamydia pneumoniae from the lungs of patients infected with the human immunodeficiency virus. *J. Clin Microbiol* 1991;29:401–402.

73. Campbell LA, Kuo CC, Grayston JT. Chlamydia pneumoniae and cardiovascular disease. *Emerg Infect Dis* 1998;4:571–579.

74. Danesh J, Collins R, Peto R. Chronic infections and coronary heart disease: is there a link? *Lancet* 1997;350:430–436.

75. Li DK, Daling JR, Wang SP, et al. Evidence that Chlamydia pneumoniae, strain TWAR, is not sexually transmitted. *J Infect Dis* 1989;160:328–331.

76. Grayston JT, Diwan VK, Cooney M, et al. Community- and hospital-acquired pneumonia associated with *Chlamydia* TWAR infection demonstrated serologically. *Arch Intern Med* 1989;149:169–173.

77. Lo SC. A newly identified human *Mycoplasma* disease. *Infect Dis News* 1990;9:73–75.

78. McMillan JA, Sandstorm C, Weiner LB, et al. Viral and bacterial organisms associated with acute pharyngitis in a school-aged population. *J Pediatr* 1986;109:747–752.

79. Gerber MA, Randolph MF, Chanatry J, et al. Role of *Chlamydia trachomatis* and *Mycoplasma pneumoniae* in acute pharyngitis in children. *Diagn Microbiol Infect Dis* 1987;6:263–265.

80. Huovinen P, Lahtonen R, Ziegler T, et al. Pharyngitis in adults: the presence and coexistence of viruses and bacterial organisms. *Ann Intern Med* 1989;110:612–616.

81. Cassell GH, Cole BC. Mycoplasmas as agents of human disease. *N Engl J Med* 1981;304:80–89.

82. Ponka A. The occurrence and clinical picture of serologically verified Mycoplasma pneumoniae infections with emphasis on central nervous system, cardiac, and joint manifestations. *Ann Clin Res* 1979;24:1–60.

83. Sköldenberg B. On the role of viruses in acute infectious diseases of the central nervous system. Clinical and laboratory studies on hospitalized patients. *Scand J Infect Dis* 1972;3(suppl):1–96.

84. MacDonald NE, Wolfish N, McLaine P, et al. Role of respiratory viruses in exacerbations of primary nephrotic syndrome. *J Pediatr* 1986;108:378–382.

85. Foy HM, Kenny GE, Cooney MK, et al. Long-term epidemiology of infections with *Mycoplasma pneumoniae*. *J Infect Dis* 1979;139:681–687.

86. Chapman RS, Henderson FW, Clyde WA Jr., et al. The epidemiology of tracheobronchitis in pediatric practice. *Am J Epidemiol* 1981;114:786–797.

87. Henderson FW, Clyde WA Jr, Collier AM, et al. The etiologic and epidemiologic spectrum of bronchiolitis in pediatric practice. *J Pediatr* 1979;95:183–190.

88. Foy HM, Grayston JT, Kenny GE, et al. Epidemiology of *Mycoplasma pneumoniae* infection in families. *JAMA* 1966;197:859–866.

89. Fernald GW, Collier AM, Clyde WA Jr. Respiratory infections due to *Mycoplasma pneumoniae* in infants and children. *Pediatrics* 1975;55:327–335.

90. Hanukoglu A, Hebroni S, Fried D. Pulmonary involvement in *Mycoplasma pneumoniae* infection in families. *Infection* 1986;14:1–6.

91. Steinberg P, White RJ, Fuld SL, et al. Ecology of *Mycoplasma pneumoniae* infections in marine recruits at Parris Island, South Carolina. *Am J Epidemiol* 1969;89:62–73.

92. Evatt BL, Dowdle WR, Johnson M, et al. Epidemic mycoplasma pneumonia. *N Engl J Med* 1971;285:374–378.

93. Cordero L, Cuadrado R, Hall CB, et al. Primary atypical pneumonia: an epidemic caused by *Mycoplasma pneumoniae*. *J Pediatr* 1967;71:1–12.

94. Clyde WA Jr. Recovery of mycoplasmas from the respiratory tract. In: Tully JG, Razin S, eds. *Methods in Mycoplasmology*, Vol. 2. New York: Academic Press, 1983:9–17.

95. Embree JE, Embil JA. Mycoplasmas in diseases in humans. *Can Med Assoc J* 1980;123:105–111.

96. George RB, Ziskind MM, Rasch JR, et al. Mycoplasma and adenovirus pneumonias. Comparison with other atypical pneumonias in a military population. *Ann Intern Med* 1966;65:931–942.

97. Kleemola M, Kayhty H. Increase in titers of antibodies to *Mycoplasma pneumoniae* in patients with purulent meningitis. *J Infect Dis* 1983;146:284–288.

98. Pasculle AW. Recognition of *Mycoplasma hominis* in routine bacteriology specimens. *Clin Microbiol News* 1988;10:145–148.

99. Mufson MA, Ludwig WM, Purcell RH, et al. Exudative pharyngitis following experimental *Mycoplasma hominis* type 1 infection. *JAMA* 1965;192:1146–1152.

100. Moffet HL, Siegel AC, Doyle HK. Nonstreptococcal pharyngitis. *J Pediatr* 1968;73:51–60.

101. Neu HC, Ellner PD. Role of mycoplasma in exudative pharyngitis. *NY State J Med* 1969;69:3026–3028.

102. Taylor-Robinson D. *Ureaplasma urealyticum* (T-strain Mycoplasma) and *Mycoplasma hominis*. In: Mandell GE, Douglas RG Jr, Bennett JE, eds. *Principles and Practice of Infectious Diseases*. 3rd ed. New York: Churchill Livingstone, 1990:1458–1463.

103. Sacker I, Walker M, Brunell PA. Abscess in newborn infants caused by mycoplasma. *Pediatrics* 1970;46:303–304.

104. Powell DA, Miller K, Clyde WA Jr. Submandibular adenitis in a newborn, caused by *Mycoplasma hominis*. *Pediatrics* 1979;63:798–799.

105. Brunell PA, Dische RM, Walker MB. Mycoplasma, amnionitis, and respiratory distress syndrome. *JAMA* 1969;207:2097–2099.

106. Glaser JB, Engleberg M, Hammerschlag M. Scalp abscess associated with Mycoplasma hominis infection complicating intrapartum monitoring. *Pediatr Infect Dis* 1983;2:468–470.

107. Mizutani H, Mizutani H. Immunologic responses in patients with *Mycoplasma pneumoniae* infections. *Am Rev Resp Dis* 1983;127:175–179.

108. Jalil N, Doble A, Gilchrist C, et al. Infection of the epididymis by *Ureaplasma urealyticum*. *Genitourin Med* 1988;64:367–368.

109. Birch DF, Fairley KF, Pavillard RE. Unconventional bacteria in urinary tract disease. *Ureaplasma urealyticum*. *Kidney Int* 1981;19:58–64.

110. Cherry JD. *Mycoplasma* and *Ureaplasma* infections. In: Feigin RD, Cherry JD, eds. *Textbook of Pediatric Infectious Diseases*. 2nd ed. Philadelphia: WB Saunders, 1987:1896–1924.

111. Klein JO, Buckland D, Finland M. Colonization of newborn infants by mycoplasmas. *N Engl J Med* 1969;280:1025–1030.

112. Braun P, Lee YH, Klein JO, et al. Birth weight and genital mycoplasma in pregnancy. *N Engl J Med* 1971;284:167–171.

113. Shurin PA, Alpert S, Rosner B, et al. Chorioamnionitis and colonization of the newborn infant with genital mycoplasmas. *N Engl J Med* 1975;293:5–8.

114. Stagno S, Brasfield DM, Brown MB, et al. Infant pneumonitis associated with cytomegalovirus, *Chlamydia*, *Pneumocystis*, and *Ureaplasma*: a prospective study. *Pediatrics* 1981;68:322–329.

115. Brasfield DM, Stagno S, Whitley RJ, et al. Infant pneumonitis associated with cytomegalovirus, *Chlamydia*, *Pneumocystis*, and *Ureaplasma*: follow-up. *Pediatrics* 1987;79:76–83.

116. Quinn PA, Gillan JE, Markestad T, et al. Intrauterine infection with Ureaplasma urealyticum as a cause of fatal neonatal pneumonia. *Pediatr Infect Dis* 1985;4:538–543.

117. Kraus VB, Baraniuk JN, Hill GB, et al. Ureaplasma urealyticum septic arthritis in hypogammaglobulinemia. *J Rheumatol* 1988;15:369–371.

118. McDonald JC, Moore DL. Mycoplasma hominis meningitis in a premature infant. *Pediatr Infect Dis* 1988;7:795–798.

119. Gilbert GL, Law F, Macinnes SJ, et al. Chronic *Mycoplasma hominis* infection complicating severe intraventricular hemorrhage in premature neonate. *Pediatr Infect Dis* 1988;7:817–818.

120. Waites KB, Rudd PT, Crouse DT, et al. Chronic *Ureaplasma urealyticum* and *Mycoplasma hominis* infection of the central nervous systems in preterm infants. *Lancet* 1988;1:17–21.

121. Waites KB, Duffy LB, Crouse DT, et al. Mycoplasmal infections of cerebrospinal fluid in newborn infants from a community hospital population. *Pediatr Infect Dis* 1990;9:241–245.

122. Kundsen RB, Poulin SA. Laboratory diagnosis of genital *Mycoplasma* infection. *Ann NY Acad Sci* 1988;549:72–80.

123. Cunningham CK. The role of genital mycoplasmas in neonatal disease. *Clin Microbiol News* 1990;12:147–149.

124. Taylor-Robinson D. The male reservoir of *Ureaplasma urealyticum*. *Pediatr Infect Dis* 1986;5:S234–S235.

125. Cassell GH, Waites KB, and Davis JK. Role of *Ureaplasma urealyticum* in amnionitis. *Pediatr Infect Dis* 1986;5:S247–S252.

126. Dinsmoor MJ, Ramamurthy RS, Gibbs RS. Transmission of genital mycoplasmas from mother to neonate in women with PROM. *Pediatr Infect Dis* 1989;8:483–487.

127. Embree J. *Mycopasma hominis* in maternal and fetal infections. *Ann NY Acad Sci* 1988;549:56–64.

128. Foy HM, Kenny GE, Levinsohn EM, et al. Acquisition of mycoplasmata and T-strains during infancy. *J Infect Dis* 1970;121:579–587.

129. Hammerschlag MR, Doraiswamy B, Cox P, et al. Colonization of sexually abused children with genital mycoplasmas. *Sex Transm Dis* 1987;14:23–25.

130. Smaron MF, Boonlayangoor S, Zierdt CH. Detection of *Mycoplasma hominis* septicemia by radiometric blood culture. *J Clin Microbiol* 1985;21:298–301.

131. Robertson JA. Bromthymol blue broth: improved medium for detection of *Ureaplasma urealyticum* (T-strain mycoplasma). *J Clin Microbiol* 1978;7:127–132.

132. Shepard MC. Culture medium for ureaplasmas. In: Razin S, Tully JG, eds. *Methods in Mycoplasmology*. New York: Academic Press, 1983;1:137–146.

133. Taylor-Robinson D. Recovery of mycoplasmas from the genitourinary tract. In: Tully JG, Razin S, eds. *Methods in Mycoplasmology*. New York: Academic Press, 1983;1:19–26.

134. Weissfeld AS, Howard BJ, Almazan RD, et al. Miscellaneous pathogenic organisms. In: Howard BJ, Klaas J, Rubin SJ, et al., eds. *Clinical and Pathogenic Microbiology*. St. Louis: CV Mosby, 1987:457.

135. Shepard MC. Culture medium for ureaplasmas. In: Razin S, Tully JG, eds. *Methods in Mycoplasmology*. New York: Academic Press, 1983;1:137–146.

136. Shepard MC, Masover GK. Special features of ureaplasmas. In: Barile MF, Razin S, eds. *The Mycoplasmas*. New York: Academic Press, 1979;1:451–494.

137. Brown MB, Cassell GH, Taylor-Robinson D, et al. Measurement of antibody to *Ureaplasma urealyticum* by an enzyme-linked immunoassay the detection of antibody responses in patients with nongonococcal urethritis. *J Clin Microbiol* 1983;17:288–295.

138. Brown MB, Cassell GH, McCormack WM, et al. Measurement of antibody to *Mycoplasma hominis* by an enzyme-linked immunosorbent assay and detection of class-specific responses in women with postpartum fever. *Am J Obstet Gynecol* 1987;156:701–708.

139. Lo SC, Dawson MS, Wong DM, et al. Identification of *Mycoplasma incognitos* infection in patients with AIDS: an immunohistochemical, *in situ* hybridization and ultrastructural study. *Am J Trop Med Hyg* 1989;41:601–616.

140. Baseman JB, Quackenbush RL. Preliminary assessment of AIDS-associated Mycoplasma. *ASM News* 1990;56:319–323.

141. Lemaître M, Guètard D, Hènin Y, et al. Protective activity of tetracycline analogs against the cytopathic effect of the human immunodeficiency viruses in CEM cells. *Res Virol* 1990;141:5–16.

142. Stratton CW. AIDS-associated Mycoplasma: the rest of the story. *Infect Dis News* 1990;9:75–78.

143. Weisburg WG. Polyphyletic origin of bacterial parasites. In: Moulder JW, ed. *Intracellular Parasitism*. Boca Raton, FL: CRC Press; 1989:2

144. Tamura A, Ohashi N, Urakami H, et al. Classification of *Rickettsia tsutsugamushi* in a new genus, *Orientia* gen. nov., as *Orientia tsutusgamushi* comb. nov. *Int J Syst Bacteriol* 1995;45:589

145. Anacker RL, Mann RE, Gonzales C. Reactivity of monoclonal antibodies to Rickettsia rickettsii with spotted fever and typhus group rickettsiae. *J Clin Microbiol* 1987;25:167.

146. Amano K, Tamura A, Ohashi N, et al. Deficiency of peptidoglycan and lipopolysaccharide components in *Rickettsia tsutsugamushi*. *Infect Immun* 1987;55:2290.

147. Li H, Walker DH. Characterization of rickettsial attachment to host cells by flow cyotmetry. *Infect Immun* 1992;60:2030.

148. Hackstadt T, Messer R, Cieplak W, et al. Evidence for proteolytic cleavage of the 120-kilodalton outer membrane protein of rickettsiae: identification of an avirulent mutant deficient in processing. *Infect Immun* 1992;60:159.

149. Mitchell SJ, Minnick MF. Characterization of a two-gene locus from Bartonella bacilliformis associated with the ability to invade human erythrocytes. *Infect Immun* 1995;63:1552.

150. Walker DH. Rocky Mountain spotted fever: a seasonal alert. *Clin Infect Dis* 1995;20:1111.

151. Salgo MP, Telzak EE, Currie B, et al. A focus of Rocky Mountain spotted fever within New York City. *N Engl J Med* 1988;318:1345.

152. McDade JE, Newhouse VF. Natural history of *Rickettsia rickettsii*. *Ann Rev Microbiol* 1986;40:287.

153. Wells GM, Woodward TE, Fiset P, et al. Rocky Mountain spotted fever caused by a blood transfusion. *JAMA* 1978;239:2763.

154. Kao GF, Evancho DC, Barskaya-Ioffe O, et al. Cutaneous histopathology of Rocky Mountain spotted fever. *J Cutaneous Pathol* 1997;24:614.

155. Silverman DJ, Santucci L, Meyers N, et al. Penetration of host cells by *Rickettsia rickettsii* appears to be mediated by a phospholipase of rickettsial origin. *Infect Immun* 1992;60:2733.

156. Heinzen RA, Hayes SF, Peacock MG, et al. Directional actin polymerization associated with spotted fever group rickettsia infection of vero cells. *Infect Immun* 1993;61:1926.

157. Silverman DJ, Santucci L. Potential for free radical-induced lipid peroxidation as a cause of endothelial cell injury in Rocky Mountain spotted fever. *Infect Immun* 1988;56:3110.

158. Walker DH, Popov VL, Wen J, et al. *Rickettsia conorii* infection of C3H/HeN mice: a model of endothelial-target rickettsiosis. *Lab Invest* 1994;70:358.

159. Feng H-M, Walker DH. Interferon-γ and tumor necrosis factor-α exert their antirickettsial effect via induction of synthesis of nitric oxide. *Am J Pathol* 1993;143:1016.

160. Sporn LA, Haidaris PJ, Rui-Jin S, et al. *Rickettsia rickettsii* infection of cultured human endothelial cells induces tissue factor expression. *Blood* 1994;83:1527.

161. Helmick CG, Bernard KW, D'Angelo LJ. Rocky Mountain spotted fever: clinical, laboratory, and epidemiological features of 262 cases. *J Infect Dis* 1984;150:480.

162. Sexton DJ, Corey GR. Rocky Mountain “spotless” and “almost spotless” fever: a wolf in sheep's clothing. *Clin Infect Dis* 1992;15:439.

163. Kirkland KB, Wilkinson WE, Sexton DJ. Therapeutic delay and mortality in cases of Rocky Mountain spotted fever. *Clin Infect Dis* 1995;20:1118.

164. Walker DH, Cain BG, Olmstead PM. Laboratory diagnosis of Rocky Mountain spotted fever by immunofluorescent demonstration of *Rickettsia rickettsii* in cutaneous lesions. *Am J Clin Pathol* 1978;69:619.

165. Dumler JS, Gage WR, Pettis GL, et al. Rapid immunoperoxidase demonstration of *Rickettsia rickettsii* in fixed cutaneous specimens from patients with Rocky Mountain spotted fever. *Am J Clin Pathol* 1990;93:410.

166. Sexton DJ, Kanj SS, Wilson K, et al. The use of a polymerase chain reaction as a diagnostic test for Rocky Mountain spotted fever. *Am J Trop Med Hyg* 1994;50:59.

167. Clements ML, Dumler JS, Fiset P, et al. Serodiagnosis of Rocky Mountain spotted fever: comparison of IgM and IgG enzyme-linked immunosorbent assays and indirect fluorescent antibody test. *J Infect Dis* 1983;148:876.

168. Kelly DJ, Chan CT, Paxton H, et al. Comparative evaluation of a commercial enzyme immunoassay for the detection of human antibody to *Rickettsia typhi*. *Clin Diagn Lab Immunol* 1995;2:356.

169. Raoult D, Zuchelli P, Weiller PJ, et al. Incidence, clinical observations and risk factors in the severe form of Mediterranean spotted fever among patients admitted to hospital in Marseilles 1983–1984. *J Infect* 1986;12:111.

170. Schriefer ME, Sacci JB Jr, Dumler JS, et al. Identification of a novel rickettsial infection in a patient diagnosed with murine typhus. *J Clin Microbiol* 1994;32:949.

171. Zavala-Velasquez JE, Ruiz-Sosa JA, Jimenez B, et al. *Rickettsia felis*: the etiologic agent of a case of rickettsiosis in the Yucatan. In: *Program and Abstracts of the European Working Group and American Society for Rickettsiology Joint Meeting*, Marseille, France; June 14–16, 1999; abstact 147A.

172. Raoult D, Weiller PJ, Chagnon A, et al. Mediterranean spotted fever: clinical, laboratory, and epidemiological features of 199 cases. *Am J Trop Med Hyg* 1986;35:845.

173. Sexton DJ, Corey GR, Greenfield JC Jr, et al. Imported African tick bite fever: a case report. *Am J Trop Med Hyg* 1999;60:865.

174. Kass EM, Szaniawski WK, Levy H, et al. Rickettsialpox in a New York City hospital, 1980 to 1989. *N Engl J Med* 1994;331:1612.

175. Walker DH, Hudnall SD, Szaniawski WK, et al. Monoclonal antibody-based immunohistochemical diagnosis of rickettsialpox: the macrophage is the principal target. *Mod Pathol* 1999;12:529.

176. Yagupsky P, Wolach B. Fatal Israeli spotted fever in children. *Clin Infect Dis* 1993;17:850.

177. Drancourt M, George F, Brouqui P, et al. Diagnosis of Mediterranean spotted fever by indirect immunofluorescence of *Rickettsia conorii* in circulating endothelial cells isolated with monoclonal antibody-coated immunomagnetic beads. *J Infect Dis* 1992;166:660.

178. Marrero M, Raoult D. Centrifugation-shell vial technique for rapid detection of Mediterranean spotted fever rickettsiae in blood culture. *Am J Trop Med Hyg* 1989;40:197.

179. Traub R, Wisseman CL Jr, Farhang-Azad A. The ecology of murine typhus: a critical review. *Trop Dis Bull* 1978;75:237.

180. Azad AF, Sacci JB, Nelson WM, et al. Genetic characterization and transovarial transmission of a typhus-like rickettsia found in cat fleas. *Proc Natl Acad Sci USA* 1992;89:43.

181. Betz TG, Rawlings JA, Taylor JP, et al. Endemic typhus in Texas. *Tex Med* 1983;79:48.

182. Walker DH, Parks FM, Betz TG, et al. Histopathology and immunohistologic demonstration of the distribution of *Rickettsia typhi* in fatal murine typhus. *Am J Clin Pathol* 1989;91:720.

183. Dumler JS, Taylor JP, Walker DH. Clinical and laboratory features of murine typhus in south Texas, 1980–1987. *JAMA* 1991;266:1365.

184. Strand Ö, Strömberg A. Ciprofloxacin treatment of murine typhus. *Scand J Infect Dis* 1990;22:503.

185. Keren G, Itzhaki A, Oron C, et al. Evaluation of the anti-rickettsial activity of fluoroquinolones. *Drugs* 1995;49(suppl 2):208.

186. Zinsser H. Rats, Lice, and History. Boston, MA: Little, Brown and Company, 1934.

187. Raoult D, Ndihokubwayo JB, Tissot-Dupont H, et al. Outbreak of epidemic typhus associated with trench fever in Burundi. *Lancet* 1998;352:353.

188. Walker DH, Barbour AG, Oliver JH, et al. Emerging bacterial zoonotic and vector-borne diseases. Ecological and epidemiological factors. *JAMA* 1996;275:463.

189. Strickman D, Sheer T, Salata K, et al. *In vitro* effectiveness of azithromycin against doxycycline-resistant and -susceptible strains of *Rickettsia tsutsugamushi*, etiologic agent of scrubtyphus. *Antimicrob Agents Chemother* 1995;39:2406.

190. Maeda K, Markowitz N, Hawley RC, et al. Human infection with *Ehrlichia canis*, a leukocytic *Rickettsia*. *N Engl J Med* 1987;316:853.

191. Anderson BE, Dawson JE, Jones DC, et al. *Ehrlichia chaffeensis*, a new species associated with human ehrlichiosis. *J Clin Microbiol* 1991;29:2838.

192. Chen S-M, Dumler JS, Bakken JS, et al. Identification of a granulocytotropic *Ehrlichia* species as the etiologic agent of human disease. *J Clin Microbiol* 1994;32:589.

193. Buller RS, Arens M, Hmiel SP, et al. *Ehrlichia ewingii*, a newly recognized agent of human ehrlichiosis. *N Engl J Med* 1999;341:148.

194. Fishbein DB, Kemp A, Dawson JE, et al. Human ehrlichiosis: Prospective active surveillance in febrile hospitalized patients. *J Infect Dis* 1989;160:803.

195. Bakken JS, Dumler JS, Chen SM, et al. Human granulocytic ehrlichiosis in the upper midwest United States. A new species emerging? *JAMA* 1994;272:212.

196. Walker DH, Dumler JS. The emergence of human ehrlichioses as human health problems. *Emerg Infect Dis* 1996;2:18.

197. Comer JA, Nicholson WL, Olson JG, et al. Serologic testing for human granulocytic ehrlichiosis at a national referral center. *J Clin Microbiol* 1999;37:558.

198. Dumler JS, Bakken JS. Human ehrlichioses. Newly recognized infections transmitted by ticks. *Ann Rev Med* 1998;49:201.

199. Petrovec M, Furlan SL, Zupanc TA, et al. Human disease in Europe caused by a granulocytic Ehrlichia. *J Clin Microbiol* 1997;35:1556.

200. Lockhart JM, Davidson WR, Stallknecht DE, et al. Isolation of *Ehrlichia chaffeensis* from wild white-tailed deer (*Odocoileus virginianus*) confirms their role as natural reservoir hosts. *J Clin Microbiol* 1997;35:1681.

201. Telford III SR, Dawson JE, Katavolos P, et al. Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle. *Proc Natl Acad Sci USA* 1996,93:6209.

202. Walls JJ, Greig B, Neitzel DF, et al. Natural infection of small mammal species in Minnesota with the agent of human granulocytic ehrlichiosis. *J Clin Microbiol* 1997;35:853.

203. Sulkowski MS, Dumler JS. The clinical spectrum of disease due to *Ehrlichia chaffeensis*: report of two cases and review. *Infect Dis Clin Prac* 1998;7:252.

204. Fishbein DB, Dawson JE, Robinson LE. Human ehrlichiosis in the United States, 1985 to 1990. *Ann Intern Med* 1994;120:736.

205. Bakken JS, Krueth J, Wilson-Nordskog C, et al. Clinical and laboratory characteristics of human granulocytic ehrlichiosis. *JAMA* 1996;275:199.

206. Dumler JS, Dawson JE, Walker DH. Human ehrlichiosis: Hematopathology and immunohistologic detection of *Ehrlichia chaffeensis*. *Hum Pathol*. 1993;24:391.

207. Everett ED, Evans KA, Henry RB, et al. Human ehrlichiosis in adults after tick exposure: Diagnosis using polymerase chain reaction. *Ann Intern Med* 1994;120:730.

208. Edelman DC, Dumler JS. Evaluation of an improved PCR diagnostic assay for human granulocytic ehrlichiosis. *Molecular Diagnosis* 1996;1:41.

209. Walker DH, Dumler JS. Human monocytic and granulocytic ehrlichioses. Discovery and diagnosis of emerging tick-borne infections and the critical role of the pathologist. *Arch Pathol Lab Med* 1997;121:785.

210. Brouqui P, Raoult D. *In vitro* antibiotic susceptibility of the newly recognized agent of ehrlichiosis in humans, *Ehrlichia chaffeensis*. *Antimicrob Agents Chemother*. 1992;36:2799.

211. Klein MB, Nelson CM, Goodman JL. Antibiotic susceptibility of the newly cultivated agent of human granulocytic ehrlichiosis: promising activity of quinolones and rifamycins. *Antimicrob Agents Chemother* 1997;41:76.

212. Tissot Dupont H, Raoult D, Brouqui P, et al. Epidemiologic features and clinical presentation of acute Q fever in hospitalized patients: 323 French cases. *Am J Med* 1992;93:427.

213. Brouqui P, Tissot Dupont H, Drancourt M, et al. Chronic Q fever. Ninety-two cases from France, including 27 cases without endocarditis. *Arch Intern Med* 1993;153:642.

214. Langley JM, Marrie TJ, Covert A, et al. Poker player's pneumonia. An urban outbreak of Q fever following exposure to a parturient cat. *N Engl J Med* 1988;319:354.

215. Sawyer LA, Fishbein DB, McDade JE. Q fever in patients with hepatitis and pneumonia: results of laboratory-based surveillance in the United States. *J Infect Dis* 1988;158:497.

216. To H, Kako N, Zhang GQ, et al. Q fever pneumonia in children in Japan. *J Clin Microbiol* 1996;34:647.

217. Raoult D, Stein A. Q fever during pregnancy — a risk for women, fetuses, and obstetricians. *N Engl J Med* 1994;330:371.

218. Raoult D, Brouqui P, Marchou B, et al. Acute and chronic Q fever in patients with cancer. *Clin Infect Dis* 1992;14:127.

219. Janigan DT, Marrie TJ. An inflammatory pseudotumor of the lung in Q fever pneumonia. *N Engl J Med* 1983;308:86.

220. Walker DH. Pathology of Q fever. In: Walker DH, ed. *Biology of Rickettsial Diseases*. Boca Raton, Florida: CRC Press, 1988:17.

221. Brouqui PB, Dumler JS, Raoult D. Immunohistologic demonstration of *Coxiella burnetii* in the valves from patients with Q fever endocarditis. *Am J Med* 1994;97:451.

222. Levy P, Raoult D, Razongles JJ. Q-fever and autoimmunity. *Eur J Epidemiol* 1989;5:447.

223. Tissot Dupont H, Thirion X, Raoult D. Q fever serology: cutoff determination for microimmunofluorescence. *Clin Diagn Lab Immunol* 1994;1:189.

224. Peacock MG, Philip RN, Williams JC, et al. Serological evaluation of Q fever in humans: enhanced phase I titers of immunoglobulins G and A are diagnostic for Q fever endocarditis. *Infect Immun* 1983;41:1089.

225. Musso D, Raoult D. Coxiella burnetii blood cultures from acute and chronic Q-fever patients. *J Clin Microbiol* 1995;33:3129.

226. Stein A, Raoult D. Detection of Coxiella burnetii by DNA amplification using polymerase chain reaction. *J Clin Microbiol* 1992;30:2462.

227. Raoult D. Treatment of Q fever. *Antimicrob Agents Chemotherap* 1993;37:1733.

228. Raoult D, Houpikian P, Tissot Dupont H, et al. Treatment of Q fever endocarditis: comparison of 2 regimens containing doxycycline and ofloxacin or hydroxychloroquine. *Arch Intern Med* 1999;159:167.

229. Relman DA, Loutit JS, Schmith TM, et al. The agent of bacillary angiomatosis. An approach to the identification of uncultured pathogens. *N Engl J Med* 1990;323:1573.

230. Slater L, Welch DF, Hensel D, et al. A newly recognized fastidious gram-negative pathogen as a cause of fever and bacteremia. *N Engl J Med* 1990;323:1587.

231. Wear DJ, Margileth AM, Hadfield TL, et al. Cat scratch disease: a bacterial infection. *Science* 1983;221:1403.

232. Regnery RL, Olson JG, Perkins BA, et al. Serological response to “Rochalimaea henselae” antigen in suspected cat-scratch disease. *Lancet* 1992;339:1443.

233. Spach DH, Koehler JE. Bartonella-associated infections. *Infect Dis Clin NA* 1998;12:137.

234. Birtles RJ, Harrison TG, Saunders NA, et al. Proposals to unify the genera *Grahamella* and *Bartonella*, with descriptions of *Bartonella talpae* comb. nov., *Bartonella peromysci* comb. nov., and three new species, *Bartonella grahamii* sp. nov., *Bartonella taylorii* sp. nov., and *Bartonella doshiae* sp. nov. *Int J Syst Bacteriol* 1995;45:1.

235. Brenner DJ, O'Connor SP, Winkler HH, et al. Proposals to unify the genera *Bartonella* and *Rochalimaea*, with descriptions of *Bartonella quintana* comb. nov., *Bartonella vinsonii* comb. nov., *Bartonella henselae* comb. nov., and *Bartonella elizabethae* comb. nov., and to remove the family Bartonellaceae from the order Rickettsiales. *Int J Syst Bacteriol* 1993;43:777.

236. Mernaugh G, Ihler GM. Deformation factor: an extracellular protein synthesized by *Bartonella bacilliformis* that deforms erythrocyte membranes. *Infect Immun* 1992;60:937.

237. Burgess AW, Anderson BE. Outer membrane proteins of *Bartonella henselae* and their interaction with human endothelial cells. *Microb Pathog* 1998;25:157.

238. Minnick MF, Smitherman LS, Carroll JA. A tail-specific protease from Bartonella quintana may be involved in hemin acquisition. In: Program and Abstracts of the European Working Group and American Society for Rickettsiology Joint Meeting, Marseille, France, June 14–16, 1999; abstact 146A.

239. Schmiederer M, Anderson BE. Characterization of a virulence gene homologue in *Bartonella henselae*. In: Program and Abstracts of the European Working Group and American Society for Rickettsiology Joint Meeting, Marseille, France, June 14–16, 1999; abstract 39.

240. Garcia FU, Wojta J, Broadley KN, et al. *Bartonella bacilliformis* stimulates endothelial cells *in vitro* and is angiogenic *in vivo*. *Am J Pathol* 1990;136:1125.

241. Hertig M. Phlebotomus and Carrion's disease. *Am J Trop Med* 1942;22:637.

242. Higgins JA, Radulovic S, Jaworski DC, et al. Acquisition of the cat scratch disease agent *Bartonella henselae* by cat fleas (Siphonaptera: pulicidae). *J Med Entomol* 1996;33:490.

243. Chomel BB, Kasten RW, Floyd-Hawkins K, et al. Experimental transmission of Bartonella henselae by the cat flea. *J Clin Microbiol* 1996;34:1952.

244. Ellis BA, Regnery RL, Beati L, et al. Rats of the genus *Rattus* are reservoir hosts for pathogenic *Bartonella* species: an old world origin for a new world disease? *J Infect Dis* 1999;180:220.

245. Koehler JE, Glaser CA, Tappero JW. *Rochalimaea henselae* infection: a new zoonosis with the domestic cat as reservoir. *JAMA* 1994;271–531.

246. Spach DH, Kanter AS, Dougherty MJ, et al. *Bartonella* (*Rochalimaea*) quintana bacteremia in inner city patients with chronic alcoholism. *N Engl J Med* 1995;332:424.

247. Brouqui P, Lascola B, Roux V, et al. Chronic *Bartonella quintana* bacteremia in homeless patients. *N Engl J Med* 1999;340:184.

248. Jackson LA, Perkins BA, Wenger JD. Cat scratch disease in the United States: an analysis of three national databases. *Am J Public Health* 1993;83:1707.

249. Tappero JW, Koehler JE, Berger TG, et al. Bacillary angiomatosis and bacillary splenitis in immunocompetent adults. *Ann Int Med* 1993;118:363.

250. Zangwill KM, Hamilton DH, Perkins BA, et al. Cat scratch disease in Connecticut. Epidemiology, risk factors, and evaluation of a new diagnostic test. *N Engl J Med* 1993;329:8.

251. Amano Y, Rumbea J, Knobloch J, et al. Bartonellosis in Ecuador: serosurvey and current status of cutaneous verrucous disease. *Am J Trop Med Hyg* 1997;57:174.

252. Maurin M, Raoult D. *Bartonella (Rochalimaea) quintana* infections. *Clin Microbiol Rev* 1996;9:273.

253. Dalton MJ, Robinson LE, Cooper J, et al. Use of Bartonella antigens for serologic diagnosis of cat-scratch disease at a national referral center. *Arch Intern Med* 1995;155:1670.

254. Noah DA, Kramer CM, Verbsky MP, et al. Survey of veterinary professionals and other veterinary conference attendees for antibodies to *Bartonella henselae* and *B. quintana*. *JAVMA* 1997;210:342.

255. Chomel BB, Abbott RC, Kasten RW, et al. Bartonella henselae prevalence in domestic cats in California: risk factors and association between bacteremia and antibody titers. *J Clin Microbiol* 1995;33:2445.

256. Kordick DL, Wilson KH, Sexton DJ, et al. Prolonged *Bartonella* bacteremia in cat associated with cat-scratch disease patients. *J Clin Microbiol* 1995;33:3245.

257. Gray GC, Johnson AA, Thornton SA, et al. An epidemic of Oroya fever in the Peruvian Andes. *Am J Trop Med Hyg* 1990;42:215.

258. Garcia FU, Wojta J, Hoover RL. Interactions between live *Bartonella bacilliformis* and endothelial cells. *J Infect Dis* 1992;165:1138.

259. Margileth AM. Antibiotic therapy for cat-scratch disease: clinical study of therapeutic outcome in 268 patients and a review of the literature. *Pediatr Infect Dis J* 1992;11:474.

260. Jacobs RF, Schutze GE. *Bartonella henselae* as a cause of prolonged fever and fever of unknown origin in children. *Clin Infect Dis* 1998;26:80.

261. Armengol CE, Hendley JO. Cat-scratch disease encephalopathy: a cause of status epilepticus in school-aged children. *J Pediatr* 1999;134:365.

262. Margileth AM, Wear DJ, Hadfield TL, et al. Cat-scratch disease. Bacteria in skin at the primary inoculation site. *JAMA* 1984;252:928.

263. Drancourt M, Mainardi JL, Brouqui P, et al. *Bartonella* (*Rochalimaea*) quintana endocarditis in three homeless men. *N Engl J Med* 1995;332:419.

264. Reed JA, Brigati DJ, Flynn SD, et al. Immunocytochemical identification of *Rochalimaea henselae* in bacillary (epithelioid) angiomatosis, parenchymal bacillary peliosis, and persistent fever with bacteremia. *Am J Surg Pathol* 1992;16:650.

265. LeBoit PE, Berger TG, Egbert BM, et al. Bacillary angiomatosis. The histopathology and differential diagnosis of a pseudoneoplastic infection in patients with human immunodeficiency virus disease. *Am J Surg Pathol* 1989;13:909.

266. Perkocha LA, Geaghan SM, Benedict Yen TS, et al. Clinical and pathological features of bacillary peliosis hepatis in association with human immunodeficiency virus infection. *N Engl J Med* 1990;323:1581.

267. Koehler JE, Quinn FD, Berger TG, et al. Isolation of *Rochalimaea* species from cutaneous and osseous lesions of bacillary angiomatosis. *N Engl J Med* 1992;327:1625.

268. Tappero JW, Mohle-Boetani J, Koehler JE, et al. The epidemiology of bacillary angiomatosis and bacillary peliosis. *JAMA* 1993;269:770.

269. LaScola B, Raoult D. Culture of *Bartonella quintana* and *Bartonella henselae* from human samples: a 5-year experience (1993–1998). *J Clin Microbiol* 1999;37:1899.

270. Brenner SA, Rooney JA, Manzewitsch P, Regnery RL. Isolation of *Bartonella (Rochalimaea) henselae*: effects of methods of blood collection and handling. *J Clin Microbiol* 1997;35:544.

271. Wong MT, Thornton DC, Kennedy RC, et al. A chemically defined liquid medium that supports primary isolation of *Rochalimaea (Bartonella) henselae* from blood and tissue specimens. *J Clin Microbiol* 1995;33:742.

272. Kordick DL, Hilyard EJ, Hadfield TL, et al. *Bartonella clarridgeiae*, a newly recognized zoonotic pathogen causing inoculation papules, fever, and lymphadenopathy (cat scratch disease). *J Clin Microbiol* 1997;35:1813.

273. Larson AM, Dougherty MJ, Nowowiejski DJ, et al. Detection of *Bartonella (Rochalimaea) quintana* by routine acridine orange staining of broth blood cultures. *J Clin Microbiol* 1994;32;1492.

274. Daly JS, Worthington MG, Brenner DJ, et al. *Rochalimaea elizabethae* sp. nov. isolated from a patient with endocarditis. *J Clin Microbiol* 1993;31:872.

275. Welch DF, Carroll KC, Hofmeister EK, et al. Isolation of a new subspecies, *Bartonella vinsonii* subsp. *arupensis*, from a cattle rancher: identity with isolates found in conjunction with *Borrelia burgdorferi* and *Babesia microti* among naturally infected mice. *J Clin Microbiol* 1999;37:2598.

276. Drancourt M, Raoult D. Proposed tests for the routine identification of Rochalimaea species. *Eur J Clin Microbiol Infect Dis* 1993;12:710.

277. Bergmans AMC, Groothedde J-M, Schellekens JFP, et al. Etiology of cat scratch disease: comparison of polymerase chain reaction detection of *Bartonella* (formerly *Rochalimaea*) and *Afipia felis* DNA with serology and skin tests. *J Infect Dis* 1995;171:916.

278. Not T, Canciani M, Buratti E, et al. Serologic response to *Bartonella henselae* in patients with cat scratch disease and in sick and healthy children. *Acta Paediatr* 1999;88:284.

279. Szelc-Kelly CM, Goral S, Perez-Perez GI, et al. Serologic responses to *Bartonella* and *Afipia* antigens in patients with cat scratch disease. *Pediatrics* 1995;96:1137.

280. McGill SL, Regnery RL, Karem KL. Characterization of human immunoglobulin (Ig) isotype and IgG subclass response to *Bartonella henselae* infection. *Infect Immun* 1998;66:5915.

281. Patnaik M, Peter JB. Cat-scratch disease, *Bartonella henselae*, and the usefulness of routine serological testing for *Afipia felis*. *Clin Infect Dis* 1995;21:1064.

282. Raoult D, Tissot Dupont H, Enea-Mutillod M. Positive predictive value of Rochalimaea henselae antibodies in the diagnosis of cat-scratch disease. *Clin Infect Dis* 1994;19:355.

283. Maurin M, Eb F, Etienne J, et al. Serological cross-reactions between *Bartonella* and *Chlamydia* species: implications for diagnosis. *J Clin Microbiol* 1997;35:2283.

284. La Scola B, Raoult D. Serologic cross-reactions between *Bartonella quintana*, *Bartonella henselae*, and *Coxiella burnetii*. *J Clin Microbiol* 1996;34:2270.

285. Sobraquès M, Maurin M, Birtles RJ, et al. *In vitro* suseptibilities of four *Bartonella bacilliformis* strains to 30 antibiotic compounds. *Antimicrob Agents Chemotherap* 1999;43:2090

286. Maurin M, Raoult D. Antimicrobial susceptibility of *Rochalimaea quintana*, *Rochalimaea vinsonii*, and the newly recognized *Rochalimaea henselae*. *J Antimicrob Chemother* 1993;32:587.

287. Bass JW, Freitas BC, Freitas AD, et al. Prospective randomized double blind placebo-controlled evaluation of azithromycin for treatment of cat-scratch disease. *Pediatr Infect Dis J* 1998;17:447

54
Aerobic Actinomycetes

Michael A. Saubolle

The order *Actinomycetales* includes a vast array of taxonomically heterogeneous genera and species, of which the majority are saprophytic and only a small minority pathogenic. Human infection by this group of organisms is uncommon, and when it occurs, it is most often “opportunistic.” Nevertheless, there has been a recent increase in recognizable infections resulting from aerobic actinomycetes. This increase may in part be the result of shifts in population dynamics, including increasing age and centralization in urban areas, and to the increasing prevalence of chronic obstructive lung disease and immunocompromising conditions such as the acquired immune deficiency syndrome (AIDS).

The aerobic actinomycetes cause a wide spectrum of diseases in humans; clinical presentations and histopathology may differ greatly between taxa, although there often may be significant overlap. In some instances, such as with tuberculosis and leprosy, the role of the aerobic actinomycetes is well known. In others, their role is increasing in scope but is still commonly under-appreciated; their incidence probably is greater than suspected (1, 2, 3, 4, 5 and 6). Initial infection almost always occurs from exogenous sources because these organisms are found rarely as normal microbial flora of humans.

Therapeutic approaches may vary widely depending on specific etiologies.

With the exception of *Mycobacterium leprae*, *Mycobacterium tuberculosis* complex, and the genus *Dermatophilus*, most aerobic actinomycetes exist predominantly as saprophytes. Their presence has been noted in a wide variety of natural as well as man-made environments. Wide variation in the geographic prevalence between species contributes to their diverse epidemiology.

The large number of species present within *Actinomycetales*, together with the confusion and difficulty often associated with their identification, add importance to the laboratory's familiarization with more commonly isolated species. Of the over 40 genera of aerobic actinomycetes, only approximately 18 seem to be of clinical significance (1). The laboratory should be able to differentiate quickly between probable saprophytes and possible pathogens, and to provide guidance to the clinician in interpreting findings and in choosing therapeutic modalities.

GENERAL ASPECTS

Part of "54 - Aerobic Actinomycetes"

Microbiology and Classification

Actinomycetes are prokaryotic bacteria that often elongate or form filaments with tendency toward some degree of true branching; reproduction by formation of spores or fragmentation of the hyphae is normal. Recently, more systematic approaches to classification using chemical and molecular techniques for evaluation, as well as numerical taxonomic means, have provided a better understanding of the relationships between organisms within this group (1, 2). At the general laboratory level, however, morphologic and chemical characteristics help separate these organisms into broad groups and are useful for identification purposes.

Aerobic actinomycetes that produce hyphae, which later fragment into coccal elements to renew the growth cycle, can be considered as “nocardioform.” The term does not, however, imply a close relationship between members of the group. Specifically, the nocardioform actinomycetes show a tendency to produce mycelia that often are transitory or fleeting and that break up quickly into coccoid or bacillary forms. Some genera in this group (*Nocardia*, *Streptomyces*, *Oerskovia*) produce thin substrates or aerial mycelia and may resemble fungi, while other genera (*Mycobacterium*, *Rhodococcus*) have rudimentary or absent mycelia, remaining most commonly in coccoid or short bacillary form (1, 2, 3, 4, 5, 6 and 7). A second group that includes the actinomadurae and streptomycetes, produces spores in or on mycelia. The genus *Dermatophilus* belongs to yet another group of actinomycetes that form mycelia and show division both transversely and longitudinally, producing multilocular, primitive sporangia (3, 4 and 5).

The amino acids, carbohydrates, and lipids found in the cell wall seem to be stable in many genera, and their analysis often is helpful for organism identification. Diaminopimelic acid (DAP) and its stereoisomeric configuration (meso or L-form) together with cell-wall chemotypes further differentiates the clinically significant actinomycetes (3, 4 and 5). However, although useful for taxomic purposes, such studies are normally beyond the capabilities of most clinical laboratories.

Mycolic acids (mycolates) are present in all the genera within the families *Mycobacteriaceae* (*Mycobacterium* and *Tsukamurella*), *Nocardiaceae* (*Nocardia*, *Gordona*, *Rhodococcus*, and *Skermania*), and Corynebacteriaceae (*Corynebacterium*, *Dietzia*, *Turicella*) except for *Turicella* (1). The mycolic acid containing families can be differentiated further by the number of carbons in their mycolates (22–38 in corynomycolates, 34–78 in nocardomycolates, and 60–90 in mycobacterial mycolates). Absence of tuberculostearic acid in the genus *Corynebacterium* separates it from all other genera within the mycolate-containing families (4, 5, 6 and 7).

Traditionally, initial identification of the aerobic actinomycetes

requires recognition of their individual colonial and microscopic morphologies and staining characteristics. Unfortunately, some of the genera frequently resemble each other, and at times it may be hard to differentiate them. Evaluation of growth rates and patterns, delineation of antibiograms, and limited biochemical studies may provide further information as to an isolate's identity (Table 54.1). Newer techniques such as nucleic acid probes, high-performance liquid chromatography (HPLC), macromolecular analysis (e.g., of 16S rRNA and the RNA polymerase gene rpoB), as well as the polymerase chain reaction (PCR) combined with restriction endonuclease analysis have been applied with variable success for the identification of many genera and species (1, 2, 9, 10, 11, 12, 13 and 14).

TABLE 54.1. SELECTED DIFFERENTIAL CHARACTERISTICS OF AEROBIC ACTINOMYCETES ENCOUNTERED IN CLINICAL SPECIMENS[a,b]

Characteristic or Feature	Genus							
	Actinomadura	*Gordona*	*Nocardia*	*Nocardiopsis*	*Oerskovia*	*Rhodococcus*	*Streptomyces*	*Tsukamurella*
DAP isomer type	*meso*	*meso*	*meso*	*meso*	None	*meso*	L	*meso*
Cell wall sugars[c]	Mad	Arab, Gal	Arab, Gal	None	Gal	Arab, Gal	None	Arab, Gal
Size of mycolic acid (No. of carbons)	None	48–66	44–60	None	None	34–64	None	64–78
Acid fastness (weak)[d]	No	Yes	Yes	No	No	Yes	No	Yes
Lysozyme resistance	No	Variable	Yes	No	Yes	Variable	No	Yes
Aerial filament growth	Variable	No	Yes	Yes	No	No	Yes	No
Hyphal fragmentation	No		Yes	No	Yes (motile)	Yes	Variable	

[a] *Mycobacterium* and *Dermatophilus* are not listed because the former is characterized in Table 47.4, and the latter is easily differentiated and rarely encountered in clinical specimens.

[b] Compiled from references 1, 4.

[c] Arab: arabinose; Gal: galactose; Mad: madurose.

[d] Acid-fastness: weak, 0.5% to 1% sulfuric acid in decolorizer; strong, 3% hydrochloric acid in decolorizer.

Diagnostic Procedures

Laboratory Safety Considerations

Laboratories processing specimens for the detection, isolation, and identification of aerobic actinomycetes (especially mycobacteria) should ensure that safety measures are in place and that safety procedures are appropriately implemented, taught and enforced. Classification of practices is based on the potential risk for each laboratory, which in turn depends on factors such as prevalence of mycobacteriosis, number of specimens processed, and prevalence of multi-drug resistant tuberculosis (2).

Normal Biosafety level 2 practices, equipment, and facilities are needed for laboratories at low risk (i.e., those that limit themselves to preparing acid-fast smears and setting up cultures). Class I or II level biological cabinets and aerosol preventive centrifuge bucket covers are necessary for aerosol-generating procedures at the minimum.

Biosafety level 3 practices and Biosafety level 2 facilities are necessary for laboratories that work with *M. tuberculosis* cultures, while Biosafety level 3 practices and facilities are necessary for laboratories at highest risk, including those that have high prevalence of resistant *M. tuberculosis* (2, 15, 16, 17 and 18). Appropriate facilities include a separate, dedicated, controlled-access room maintained at negative pressure relative to the rest of the laboratory. The room should contain a class II laminar-flow biological safety cabinet equipped with a high-efficiency particulate air (HEPA) filtration system, an adequate air draw across the front, and an ultraviolet light for surface decontamination. The cabinet should be inspected and certified at least annually and preferably every 6 months. An autoclave should be present in an adjacent area to decontaminate infectious waste prior to transport to disposal areas. High-speed centrifuges should be equipped with bucket covers and domes to contain possible broken tubes. Electric incinerators rather than Bunsen burners should be utilized to avoid splatters. A container of sand and 95% alcohol or a 5% phenol solution also should be available to clean wires, spades, or loops prior to their insertion in the incinerator.

Personal protective equipment including gloves, National Institute for Occupational Safety and Health (NIOSH)-certified fit-masks or respirators, caps, and gowns should be worn when working with cultures. A protocol should be available regarding the use of a respirator mask, how to conduct fit-testing of the mask, and training of personnel in the use of the respirator. All procedures should minimize aerosol production while maximizing protection from infection by droplet nuclei. All personnel working in the laboratory should be given tuberculin skin tests at least annually for those at low risk and biannually for those at highest risk to elucidate any conversions.

Laboratory Proficiency Considerations

Not all laboratories can be expected to have the facilities, volume, or expertise to work with the aerobic actinomycetes, and especially with the mycobacteria, within adequate safety or proficiency parameters. The College of American Pathologists (CAP) and the American Thoracic Society (ATS) have recommended progressive extents (CAP) and levels (ATS) of proficiency of services for the diagnosis of mycobacterial diseases (2, 15). In general, laboratories with minimum volume of specimens (<10–15 for smears, <20 for culture) requiring mycobacterial workup should utilize the services of reference laboratories capable of documenting proficiency at isolating and identifying mycobacteria. Laboratories performing mycobacterial studies

should be capable of performing smears daily and using modern culture techniques minimizing delay in recovery or susceptibility testing of mycobacteria (17).

Primary Isolation

In most instances, isolation of an agent causing disease is necessary to provide definitive identification and to perform antimicrobial susceptibility studies when such are available or necessary. Specific isolation methods for individual groups of aerobic actinomycetes vary widely. Methods that are optimal for the recovery of groups such as the mycobacteria may not be suitable for groups such as the actinomadura or streptomyces. For some actinomycetes (e.g., *Mycobacterium leprae*), isolation methods are not routinely available, while for others, isolation methods may require selective concentration, decontamination, and/or culture onto specialized media for optimal recovery.

Methods of specimen collection and transport to the laboratory directly influence the detection and isolation of the aerobic actinomycetes. Not all laboratories should be capable of isolation or detection of all genera of actinomycetes. Still, they should be able to provide information to enhance collection techniques and to expedite specimen transportation with minimum delay and maximum quality assurance.

The majority of specimens from mycobacterial or nocardial infections are from the respiratory tract, followed by lymph node tissue, urine, other tissue, and normally sterile body fluids. Specimens from patients with AIDS may include blood and stool. Specimens from actinomycotic mycetomas may include pus, serosanguinous fluid, sinus scrapings, or other biopsy material.

The slower growth of the actinomycetes becomes a problem when other rapidly growing organisms are present in the specimen; such normal or contaminating flora may overgrow the etiologic agent. Collection methods should therefore bypass areas of contamination. In some instances, specimens containing mixed flora should be processed to eliminate the contamination or be inoculated onto selective media to allow the pathogens a chance to grow. All specimens should be refrigerated if not processed within a few hours.

Mailing

The packaging of specimens or cultures to be mailed to reference laboratories must adhere to the Code of Federal Regulations. Screw-cap, watertight tubes must be used. These must be placed within a second watertight container (usually metallic) with enough absorbent material capable of taking up liquid in case of breakage. One or more of these can be packed into containers made of paperboard wood and mailed only after appropriate warning labels are attached (15).

Direct Detection

Microscopic evaluation of specimens can provide rapid information on the presence or absence of: (i) the possible etiologic agent; (ii) infected material characterized by polymorphonuclear or mononuclear leukocytes; or (iii) contamination characterized by squamous epithelial cells. A variety of wet-mount, Gram-stain, or other special stains may be used to detect the aerobic actinomycetes either directly from specimens (sputum, pus, tissue, stool, etc.) or after their decontamination and/or concentration (e.g., specimens for mycobacteria, normally sterile body fluids).

Molecular-based amplification as well as fluorescent-column chromatographic technologies now offer fairly sensitive direct detection methods and provide supplemental capabilities useful in some clinical situations (19, 20, 21, 22, 23, 24 and 25).

Serologic Procedures

The complexity, heterogeneity, and cross-reactivity of antigens among the aerobic actinomycetes have historically precluded the routine use of serologic techniques for diagnosis; in some instances, however, there seems some hope for the future (2, 28). In mycobacterial studies, antigens evaluated have included the 38-kDa antigen, lipoarabinomannan, antigen 60 and the antigen 85 complex (2). Although some reagents currently are available commercially, none have been adequately documented to be sensitive or specific enough for routine clinical use. Antigen capture assays also have been evaluated but cannot be recommended at this time (2, 27, 28).

Skin tests may be of value, both in the mycobacterioses and in actinomycotic mycetoma. Reviews of skin tests and their interpretation and significance can be found elsewhere and are not discussed here (26, 27, 29).

Interpretation of Data

The presence of some genera of aerobic actinomycetes in clinical material is highly associated with an active disease process. Thus, identification of *M. tuberculosis*, *M. leprae*, or *Dermatophilus congolensis* in patients always is considered significant (Table 54.2). Correlation of other groups of actinomycetes, such as the nontuberculous mycobacteria, and nocardia, with disease is more difficult. The environmental prevalence of many actinomycetes, coupled with their low virulence, increases the difficulty of interpreting their causal role in an infectious process.

TABLE 54.2. CLINICALLY SIGNIFICANT AEROBIC ACTINOMYCETES

Genus	Species	More Commonly Associated Infections
Actinomadura	A. madurae and *A. pelletieri* most significant; soil organisms	Mycetoma
Dermatophilus	*D. congolensis* the only spp.; obligate animal parasite	Pustular, exudative dermatitis
Mycobacterium	*M. leprae*, *M. tuberculosis*, *M. bovis* obligate pathogens;	*M. leprae*: leprosy; *M. tuberculosis*, *M. bovis*: tuberculosis
	Nontuberculous mycobacteria (NTM) include a large number of species with potential capability to invade the immunocompromised host; NTM found widely in water, environment	*M. avium* complex, *M. kansasii*, *M. abscessus*, *M. chelonae*, *M. fortuitum* most commonly isolated: associated with pulmonary or infrequently with disseminated disease; rapid growers also associated with localized abscesses post–trauma, foreign body implantation, including lines
Nocardia	Of 11 recognized species, only eight are clinically significant and only six are more frequently isolated in the United States, found in soil, organic debris, and water	*N. asteroides*, *N. farcinica*, *N. pseudobrasiliensis*, *N. nova*: pulmonary, disseminated. *N. nova* is less common and less invasive than others; *N. brasiliensis*, *N. otitidiscaviarum*: localized soft tissue, trauma, but can disseminate
Nocardiopsis	*N. dassonvillei* significant but rare	Mycetoma
Rhodococcus	*R. equi* most commonly associated with disease	Pneumonia, bacteremia most common; endophthalmitis, peritonitis reported
Gordona	*G. bronchialis*, *G. rubroperctinctus*, *G. sputi*, *G. terrae*, and *G. aichiensis* rarely associated with humans; soil, sewage treatment plants	Implicated in primary cutaneous, chronic pulmonary, catheter associated, and brain abscess infections
Tsukamurella	*T. paurometabola* primary spp. of clinical importance	Bacteremias, synovitis in immunocompromised patients, associated also with those undergoing dialysis or with indwelling catheters
Streptomyces	Of a large number of species in this genus, *S. somaliensis* and *S. anulatus (griseus)* most commonly associated with human infections	Mycetoma, abscesses

Therapeutic Considerations

Therapeutic regimens for aerobic actinomycetes vary considerably according to the infecting species, the disease process, and the immunocompetency of the patient. Surgical intervention supplements or supplants antimicrobial therapy in some instances. Although standardized procedures for susceptibility studies do not exist for all genera or species, methods for a few have been developed adequately enough to provide guidance. In other cases, the disease is self-limited, and therapeutic intervention is not normally required (e.g., dermatophilosis).

This chapter focuses on the characteristics of the genera most commonly causing human infections (e.g., *Mycobacterium*, *Nocardia*, and *Rhodococcus*). The genera *Actinomadura*, *Nocardiopsis*, *Streptomyces*, and *Dermatophilus* are only briefly reviewed, while the corynebacteria and the anaerobic and thermophilic actinomycetes are not discussed.

MYCOBACTERIUM

Microbiology and Classification

The genus *Mycobacterium* consists of a diverse group of strongly acid-fast bacilli (AFB) with high lipid content. The lipids in the cell and its wall include waxes with characteristic long-chain (60–90 carbon) mycolic acids.

The diversity in the genus *Mycobacterium* is manifested by variability in morphology, growth rate characteristics, nutritional and temperature requirements, as well as pathogenicity. There are easily over 80 species of mycobacteria currently recognized, but few are more commonly associated with human infections (6, 7, 24). *M. tuberculosis* (MTB), *M. bovis*, *M. africanum*, and *M. leprae* have a high propensity for causing disease and are always considered as pathogens. With the exception of *M. leprae*, the causative agent of leprosy, they all cause classic tuberculosis and form the *M. tuberculosis* complex (Table 54.3). There is presently some doubt as to the individual status of the genus *M. africanum*, which biochemically and by DNA relatedness studies may be indistinct from MTB (2). Mycobacteria that do not belong to the MTB complex have been designated in the past by several names, but the term “nontuberculous mycobacteria” (NTM) is most appropriate.

TABLE 54.3. MYCOBACTERIAL SPECIES MORE COMMONLY ENCOUNTERED IN CLINICAL SPECIMENS IN THE UNITED STATES[a]

Group (Characteristic)	Obligately Pathogenic	Probably or Potentially Pathogenic[b]	Probably Saprophytic Rarely Pathogenic[c]
Tuberculosis complex (slow growth; nonpigmented)	*M. tuberculosis* *M. bovis* *M. leprae*		
Photochromogens (slow growth; light stimulated pigment)		*M. kansasii*, *M. simiae*, *M. marinum*	
Scotochromogens (slow growth; pigmented irrespective of light)		*M. scrofulaceum*, *M. xenopi*, *M. szulgai*[d]	*M. flavescens*, *M. gordonae*
Nonphotochromogens (slow growth, nonpigmented)		*M. avium*, *M. intracellulare*, *M. haemophilum*	M. gastri, M. terrae, M. triviale
Rapid growers (growth within 7 days; nonpigmented)		*M. fortuitum*, *M. chelonae*, *M. abscessus*	*M. smegmatis*

[a] Compiled from 2, 15, 31.

[b] Other potentially pathogenic species include *M. malmoense*, *M. asiaticum*, *M. ulcerans* (obligate pathogen), *M. celatum*, *M. genavense*.

[c] Other rarely pathogenic, commonly saprophytic species: *M. neoaurum*, *M. nonchromogenicum*, *M. parafortuitum* complex, *M. paratuberculosis*, *M. phlei*, *M. thermoresistibile*, and *M. vaccae*.

[d] Photochromogenic at 25°C, but scotochromogenic at 35°C.

Originally, colonial morphology and chromogenicity (carotenoid pigmentation) were used to categorize species of NTM. Additionally, growth rates of species were incorporated into the categorization method to delineate divisional units referred to as Runyon groups. The Runyon groups included the pigmented photochromogens (group I: yellow to orange pigment development after exposure to light; Fig. 54.1) and scotochromogens (group II: pigment development in the dark), the normally colorless nonphotochromogens (group III; Fig. 54.2) and the rapid growers (group IV). This classification was useful, but not completely accurate, because some species' characteristic chromogenicity and growth rates are influenced by other environmental factors such as temperature (2, 15, 29, 30 and 31).

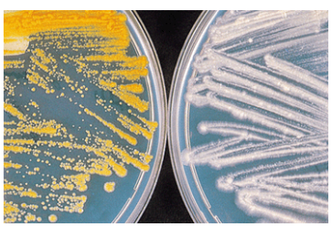


FIGURE 54.1. Colonies of the photochromogen *Mycobacterium kansasii* showing typical formation of carotenoid pigmentation only after exposure to light **(left)** and not when kept in the dark **(right)**.

FIGURE 54.2. Nonpigmented colonies of members of the *Mycobacterium avium* complex.

A second approach to delineation of species of NTM chosen by the ATS uses a clinical classification by type of disease presentation such as pulmonary, lymphadenitis, or cutaneous (30).

The spectrum of biochemical activity of culturable mycobacterial species also provides a tool for classification and identification (Table 54.4). However, results cannot always be depended on because of variability within individual species being studied. Such phenotypic variability becomes a larger problem when the number of biochemical determinations diminishes or the isolate being characterized is rarely isolated (31).

Applications of molecular hybridization techniques (genetic probes) for determination of species relatedness and high-performance liquid chromatography for mycolic acid patterns have facilitated the differentiation of species. Such methods have become more commonly available to clinical laboratories and their routine application has become feasible for cost-effective identification in clinical practice (2, 31). Other molecularly based methods such as evaluation of the hypervariable region of the 16S rRNA gene have provided an understanding to the numerical taxonomy of the mycobacteria. But such methods require molecular sophistication and at present are limited to research laboratories (2, 31).

Esoteric identification methods such as bacteriophage typing and serotyping may have been useful for epidemiologic studies,

but their poor reproducibility and specificity, together with the difficulty of standardizing test systems, have made them unsuccessful as routine tools for evaluation of mycobacterial epidemiology. Such methods have been supplanted by molecular techniques such as RFLP analysis of repetitive insertion sequence IS6110 for *M. tuberculosis* (2), and plasmid typing, rRNA spacer sequencing, and large-restriction-fragment analysis using PFGE for the NTM and other aerobic actinomycetes (32).

TABLE 54.4. MAJOR DIFFERENTIATING TESTS FOR MORE COMMON NONTUBERCULOUS MYCOBACTERIAL ISOLATES[a]

Mycobacterium Species	Nitrate Reduction	Catalase: Semiquantitative	Catalase: Heat Stable	Tween 80 Hydrolysis (5 day)	Tellurite Reduction	Urease	Arylsulfatase (3 day)
M. kansasii	+	+	+	+	V(-)	V(-)	-
M. simiae	-	+	+	-	+	V(+)	-
M. marinum	-	-	-	+	V(-)	+	V(-)
M. scrofulacium	-	+	+	-	V(-)	V(+)	-
M. xenopi	-	-	V(-)	-	V(-)	-	V(-)
M. szulgai	+	+	+	V(-)	V(+)	+	V(-)
M. gordonae	-	+	+	+	-	V(-)	-
M. flavescens	+	+	+	+	V(-)	+	-
M. avium-complex	-	-	V(+)	-	+	-	-
M. gastri	-	-	-	+	V(+)	V(-)	-
M. terrae-complex	+	+	+	+	V(-)	-	-
M. fortuitum	+	+	+	V	V(+)	+	+
M. chelonae	-	+	V(+)	V(-)	V(+)	+	+
M. abscessus	-	+		V		+	+

[a] Compiled from references 2, 15, 31.

[b] Plus signs and minus signs designate the presence or absence, respectively, of a feature or capability in the majority (>80 to 85%) of strains; V indicates variability, while the (+) or (-) signs following the V indicate a greater or lesser percent, respectively, of strains having a feature or capability.

Spectrum of Disease

M. leprae, a pathogenic member of the mycobacteria, is the causative agent of leprosy. The clinical phases of leprosy are broad, ranging from the primary silent (replication of organisms in skin macrophage) and indeterminate (replication in peripheral nerves) phases to the more severe tuberculoid and lepromatous states. Erythematous or hypopigmented skin lesions and damage to peripheral nerves are characteristic and may be progressive, depending on the stage of the disease process (33).

Members of the MTB complex, primarily transmitted person to person via droplet nuclei and the respiratory tract, produce the well-known clinical entity of tuberculosis. In the normal host, MTB remains the primary etiology of pulmonary (85% of cases) as well as of extrapulmonary (15% of cases) tuberculosis in the United States. There is an increased risk of dissemination in the immunocompromised patient. The risk of patients developing clinical disease within the first year of being infected is approximately 5%, but declines thereafter. However, risk of reactivation increases with the decrease of cellular immunity in the elderly.

Primary tuberculosis may be considered an infection of the reticuloendothelial system. The organisms are ingested but not inactivated by alveolar macrophage and polymorphonuclear cells, pass through the lymphatics, and finally enter the bloodstream to reseed the lungs or other sites. Extrapulmonary sites may include lymph nodes (28%), spinal column, as well as meninges (5%), genitourinary tract (13%), bone and joint (10%), peritoneal cavity (4%), and bone marrow. Miliary tuberculosis includes the simultaneous infection of several sites and is seen in approximately 10% of extrapulmonary cases. Pulmonary

disease may be acute and rapidly progressive or chronic with cavitation.

The recent use of the bacille Calmette-Guerin (BCG) vaccine strain of *M. bovis* to treat melanoma or carcinoma of the bladder has increased its recognition as a cause of significant infections in patients undergoing such therapy. Infection is usually limited to the urogenital tract although lymph node involvement has also been noted (31).

The NTM are far less invasive and require an underlying factor in the patient to be able to successfully colonize and cause progressive disease. Members of the *Mycobacterium avium* complex (MAC) and, to a lesser extent, *Mycobacterium kansasii* have a predilection for patients having been immunocompromised by AIDS or malignancy, or for those having chronic pulmonary disease with parenchymal lung damage. Percutaneous trauma (surgery, catheterization, accidental puncture) increases the risk for infection with *M. fortuitum, M. chelonae*, and *M. abscessus* (29, 30, 31 and 32).

Clinical presentation with infection from the NTM primarily includes pulmonary, lymph, and cutaneous/soft tissue sites, but also may involve peritoneal, joint/bone, ocular, gastrointestinal, heart valve, meningeal, and urogenital sites. Dissemination also may occur, especially in the severely immunocompromised host; members of MAC account for the majority of disseminated cases, but other species also may disseminate (29, 30, 31 and 32).

Members of MAC and *M. kansasii* are the most commonly encountered mycobacterial pathogens causing chronic pulmonary disease in the United States. Indistinguishable, indolent but potentially progressive lung disease may occasionally be caused by rapidly growing NTM, *M. xenopi*, *M. simiae*, *M. malmoense*, and *M. szulgai*. Members of MAC have become implicated in the invasion of the gastrointestinal mucosa and dissemination into blood of patients with AIDS; in such patients, the respiratory tract may not be as commonly involved. Interestingly, *M. avium* seems to be more virulent than *M. intacellulare*, accounting for the majority of infections.

Cervical lymphadenopathy, often encountered in immunocompetent pediatric patients, is commonly caused by NTM. Members of MAC are again the most common cause, although *Mycobacterium scrofulaceum* also may be encountered in this setting. Other reported etiologies include *M. kansasii*, *M. haemophilum*, and rarely other NTM (2, 32).

Several species of NTM favor reduced growth temperatures and characteristically infect cooler cutaneous or subcutaneous areas. Localized papulonodular skin lesions at sites contacting unchlorinated fresh or salt water are caused by *Mycobacterium marinum*. *Mycobacterium ulcerans*, encountered primarily in Australia, Africa, and Mexico, causes a subcutaneous infection that can progress to ulceration. *M. haemophilum* has been associated with chronic skin lesions in immunocompromised patients as well as with cervical lymphadenopathy in immunocompetent patients (29, 31, 32).

M. abscessus, *M. chelonei*, and *M. fortuitum* have emerged as major opportunistic pathogens infecting previously traumatized soft tissue such as that associated with penetrating wounds, sternotomies, mammoplasties, percutaneous catheter insertions, and keratitis/corneal ulcerations. These species also have been reported in association with cases of prosthetic valve endocarditis, osteomyelitis, synovitis, meningitis, and disseminated disease (2, 30, 31 and 32).

Other species of NTM (e.g., *Mycobacterium gordonae*, *Mycobacterium flavescens*, *Mycobacterium terrae*, *Mycobacterium asiaticum*, *Mycobacterium gastri*, *Mycobacterium thermoresistibile*, and *Mycobacterium neoaurum*) are commonly considered contaminants in culture and rarely are associated with human disease. Occasional associations with infections in immunocompromised patients, however, have been documented (31).

Epidemiology

The mycobacteria commonly associated with disease in humans and considered always to be pathogenic, are not found free-living in nature, but remain obligate animal parasites. Leprosy and tuberculosis occur worldwide. The former, caused by *M. leprae*, afflicts over 10 million people, especially in third-world or underdeveloped

countries. Person-to-person transmission through contact with intact or abraded skin, as well as by deposition on the nasal mucosa through inhalation is increased by crowded conditions and close contact (33).

In the United States, tuberculosis is caused primarily by *M. tuberculosis*. Control of the prevalence of *M. bovis* in cattle and pasteurization of milk have nearly eradicated infection in humans except as noted above for BCG. Control of *M. tuberculosis* is more problematic, although after a resurgence in the 1980s, the rate of tuberculosis is again declining in most of states. However, continued vigilance and focus by the Public Health infrastructure is necessary to keep the disease on the decline.

Unlike members of the *M. tuberculosis* complex, the NTM are found in the environment, being isolated from soil, water, sewage, house dust, air, and raw milk. Infection of humans occurs by normally unpreventable ingestion, inhalation, or inoculation from environmental sources, and person-to-person spread does not usually occur. Control of the acquisition of infection is extremely difficult (30, 32).

The NTM inhabit a wide variety of ecological niches. Sometimes, their presence in some locations may bring them in closer proximity to a human population and to increased isolation from clinical specimens. Thus, the isolation of members of MAC from water supplied to large metropolitan areas and specific sources in hospitals as well as homes may have clinical implications. These organisms also have been noted in air, water, dust, soil, plants, and animals.

Other members of NTM have been isolated from environmental sources such as soil, water, and dust (rapidly growing NTM), soil, water, milk, and oysters (*M. scrofulaceum*), untreated fresh or salt water (*M. marinum*), and water and animals (*M. xenopi*, *M. simiae*). *M. xenopi* also has been isolated from birds, coastal areas, and human tonsils, while *M. simiae* has been isolated from monkeys. *M. kansasii* has on several occasions been isolated from water, and the fastidious *M. genavense* has been associated with some psittacine birds. Environmental sources for *M. haemophilum*, *Mycobacterium szulgai*, *Mycobacterium malmoense*, *Mycobacterium ulcerans*, and the more fastidious mycobacteria are unknown (2, 30, 31).

In terms of isolation rates, members of MAC have the greatest incidence, followed by *M. kansasii*. Together, the two genera may make up 75% to 85% of all significant isolates of NTM in the clinical laboratory. Members of MAC are less likely to be considered clinically significant than *M. kansasii*. Isolation rates of other NTM drop off sharply with the rapidly growing NTM making up 12% to 15% while the remaining species, in aggregate, making up less than 10% of all isolates.

Specimen Collection

Respiratory Tract

Traditionally, three to five expectorated early-morning sputum samples collected on several different days were considered sufficient to detect mycobacteria. With newer recovery methods such as the BACTEC system (Becton Dickinson Diagnostic Instrument Systems, Cockeysville, MD), limiting sputum specimens to two on separate days is adequate if smear positive. Specimens pooled over a 24-hour period are not acceptable because of their increased contamination and slower recovery rate. Test of cure sputum specimens should be collected for microscopic analysis weekly beginning three weeks after initiation of the therapeutic protocol.

Sputum should originate from deep in the lung, and the patient should be advised on how to cough appropriately. Rinsing the mouth with water may decrease contamination. Screening for upper respiratory tract contamination is not necessary, but there should be evidence of lower tract secretions in the specimen provided (2, 15).

Specimens from patients unable to expectorate sputum may be collected after inducing coughing by nebulization of 5% to 10% saline or by bronchoscopy. Frequently, several expectorated sputum specimens collected a day or two after bronchoscopy enhance the detection of mycobacteria. In some difficult patients, transthoracic needle or open-lung biopsies may be attempted, with the latter method having the best yield. In all cases, the method of collection should be noted on the requisition.

Each specimen should be 5 to 10 mL and may be collected or transported in 50-mL centrifuge tubes. If specimens are to be mailed to a reference laboratory, 1% cetylpyridinium chloride with 2% sodium chloride may be added to the sputum as an initial decontamination step prior to shipping; AFB may survive up to 8 days in transit using this technique.

Gastric Washings

Aspiration of swallowed sputum from the stomach is reserved for patients unable to provide sputum by other means, usually the obtunded or the young. The procedure is performed early in the morning and on an empty stomach. Gastric aspirations should be processed rapidly, or 100 mg of sodium bicarbonate should be added as a buffer.

Normally Sterile Body Fluids

Specimens in this category may include 2 to 5 mL of cerebrospinal fluid (CSF), pericardial or synovial fluids, 10 to 15 mL of peritoneal, bile, paracentesis, dialysis fluid, or pleural fluid, and 3 to 5 mL of exudates. Such specimens should be collected in sterile screw-cap containers. Body fluids may be diluted in saline or a buffer to allow the relatively buoyant acid-fast bacilli to sediment adequately during centrifugation.

Tissue

Surgical biopsies should be submitted in sterile nonbacteriostatic saline to keep them from dehydrating. Larger pieces can be cut up and homogenized in a mechanized tissue grinder or ground in a mortar and pestle. Specimens that cannot be processed immediately may be frozen at -20°C for transport.

Urine

Appropriate cleansing of genitalia should precede collection of the first-morning voided urine specimens. Each specimen should have a minimum of 40 mL. Three to five specimens collected over as many days are normally adequate. 24-hour pooled specimens are not acceptable because acid-fast bacilli may be harmed by long-term exposure to urine. Specimens may be stored for short periods at 4°C.

Sinus Tracts

Caution must be used in collecting drainage material from sinus tracts, which are notorious for being easily contaminated or colonized. Attempts at collecting by needle aspiration or biopsy are preferable. Swabs are poor substitutes for exudates or biopsies, and often will decrease yields. When used, swabs should be inoculated directly to solid or broth media.

Blood

Isolation of mycobacteria from blood is possible, primarily in immunocompromised patients (especially those with AIDS) and such attempts should be reserved for such patients. Blood may be collected in the Isolator lysis-centrifugation system (Merck & Co., Inc., West Point, PA) or inoculated directly into the BACTEC 13A or Myco F Lytic blood culture bottle (Becton-Dickinson, Franklin Lakes, NJ) for mycobacteria. Both of the latter bottles are incubated without further manipulation, while blood in the Isolator tube must be processed according to the manufacturer's instructions and the sediment cultured onto egg or other-based agar media. The sediment should not be cultured in BACTEC 12B medium as this combination may be deleterious to some MAC. Mycobacterium are stable up to 24 hours in these tubes. Blood specimens also may be submitted in tubes containing the anticoagulants heparin or sodium polyanetholsulfonate but not EDTA (2).

Stool

Stool specimens are recommended only for the detection of mycobacterial involvement of the gastrointestinal tracts of patients with AIDS and are not generally recommended for routine screening. Swabs are not acceptable. Stool specimens (>1 g) should be submitted in clean containers or in stool transport systems. The stool can be emulsified in 5 mL of Dubos albumin broth and processed like a sputum sample after overnight incubation at 35°C. Fresh specimens not processed immediately may be frozen at -20°C.

Primary Isolation

Specimens considered to contain normal or superinfecting bacterial flora that can overgrow the much slower growing mycobacteria must be processed to remove or decrease the contaminants. Methods for decontamination include strong alkali, acid, or other chemicals to which the lipid-rich mycobacteria are more resistant than other bacteria. However, mycobacteria are also influenced by the decontamination procedure, with only 10% to 80% of their numbers surviving in clinical specimens. Therefore, the most gentle decontamination methods are used as necessary, with harsher methods being reserved for highly contaminated specimens containing organisms such as the pseudomonads, which are more resistant to normal procedures. Procedural steps must be followed carefully to prevent decreased survival of mycobacteria. The high lipid content of mycobacteria is responsible for their greater buoyancy in clinical specimens and especially in liquids and body fluids. Therefore, mucoid specimens such as sputum, or those with high specific gravity or density, must be digested, liquefied, and/or diluted prior to concentration of AFB. Centrifugation must be performed for 15 minutes at a minimum relative centrifugal force of 3,000 × g and preferably at a force approaching 4,000 × g (2, 15).

Most concentration procedures employ combinations of mucolytic agents together with decontaminating agents. The most commonly used agents include *N*-acetyl-L-cysteine or alternatively, dithiothreitol (Sputolysin) as mucolytic agents, combined with 2% sodium hydroxide as a decontaminant. Another acceptable decontamination method incorporates benzalkonium chloride (Zephiran) with trisodium phosphate. Use of oxalic acid combined with a mucolytic agent is beneficial for specimens that are heavily contaminated with Gram-negative rods.

Specimens requiring decontamination and concentration normally include sputum (inclusive of specimens collected by bronchoscopy), urine, gastric aspirates, stool, and contaminated tissue or swabs of pus. Heavy contamination with Gram-negative rods and especially pseudomonads may be expected in sputum from patients with cystic fibrosis or bronchiectasis due to other etiologies.

Decontamination is unnecessary for specimens such as normally sterile body fluids or tissue biopsies that are not expected to be contaminated by normal bacterial flora. If any question as to sterility status of a specimen exists, it may be advisable to first culture a portion of the specimen overnight in nutrient broth or agar to detect bacterial overgrowth prior to processing for AFB.

The laboratory's capability to recover and identify mycobacteria has changed extensively over the past 15 years. The use of newer formulations of liquid media allow smaller amounts of mycobacteria to be detected earlier than by conventional solid medium-based systems. The more rapid recovery rates are enhanced by more sensitive radiometric and fluorescent detection techniques.

Traditional primary isolation media may include nonselective egg-based media, agar-base media, and liquid media or modifications of these to formulate selective media. Egg-based media such as the preferred Lowenstein-Jensen medium usually contain products that make them opaque. Agar-base media such as Middlebrook, 7H10, and 7H11, are clear and therefore help in visualization of early mycobacterial growth; such early detection of colonies is enhanced by using a dissecting microscope. Liquid media include formulations such as the 7H9 broth (2, 15). Selective media normally contain antimicrobial agents to inhibit bacterial and fungal contamination and may include modi-fications of Lowenstein-Jensen media (Gruft modification and mycobactosel) or of 7H10 (Middlebrook medium) and 7H11 (Mitchison's medium).

The BACTEC 12B vial system was the first commercially introduced system that outperformed traditional media. That system had liquid 7H12 broth medium with similar characteristics to the modified, selective 7H11 medium, but with an added C^{14} radiolabeled fatty acid component. Growth in this medium is detected radiometrically by measuring the release of CO_2 by the BACTEC 460 instrument using needles inserted through the top of the vial. The system reduced the time to recovery of mycobacteria by 1 to 3 weeks and increased overall recovery rates of most mycobacteria including many of the fastidious ones (2, 15, 31).

Since the introduction of the BACTEC 460, many other commercial systems have become available, including the Septi-Chek

(Becton-Dickinson), the Mycobacteria Growth Indicator Tube (MGIT; Becton-Dickinson), and the automated, continuously monitoring instruments of the BACTEC 9000 MB system, the BACTEC MGIT 960, the ESP II system (Trek Diagnostics, Westlake, OH), the MB/BacT system (Organon Teknika, Durham, NC) and (presently only in Europe) the MB-Redox system (Biotest, Dreiech, Germany; 2, 34, 35, 36, 37). The newer methods have eliminated the radioisotopes and the needles of the BACTEC 460 system, relying for detection on subculture onto triple media on a paddle (Septi-Chek), fluorescence (MGIT, BACTEC 9000 MB), pressure change in the head space of the bottle (ESP II), a colorimetric CO_2 sensor (MB/BacT), and a redox indicator (MB Redox). Except for the MB Redox system, which has not yet been well evaluated in the United States and the Septi-Chek, the newer systems perform similarly to the BACTEC 460 system, but seem to have slightly higher contamination rates. To maximize recovery, detect mixed cultures, and confirm isolates of *M. tuberculosis* by niacin studies after initial probe identification, the BACTEC 460 or a comparable broth system is recommended together with a solid medium.

Isolation of *M. haemophilum* requires the additional use of chocolate agar or supplementation of media with hemin or 1% ferric ammonium citrate, as well as incubation at 30°C. Isolation of *M. genavense* from blood of patients with AIDS may require inoculation of BACTEC 13A media, extension of incubation to longer than 8 weeks, and supplementation of media with Mycobactin J. The more recently recognized *M. conspicuum* seems to require solid medium incubated at 22°C to 23°C, whereas it can grow in BACTEC broth medium at 35°C to 37°C.

Cultures should be incubated at 35°C to 37°C in 5% to 10% CO_2 in the dark, although the new media systems have their own CO_2 atmospheres and don't require additional supplementation. The radiolabeled BACTEC vials should be incubated at 36 to 38°C for best results. Glass bottles or tubes initially should be kept at a slant with caps loosened to allow CO_2 entry. Plates may be placed in CO_2-permeable polyethylene bags and incubated inverted. LJ slants may be incubated in ambient air after the first 7 to 10 days. Specimens from topical skin lesions, lymph nodes, and the environment should be incubated at both 30°C and 35°C to 37°C to enhance growth of mycobacteria preferring lower temperatures (*M. marinum*, *M. haemophilum*, *M. ulcerans*, and *M. chelonae*). Cultures should be held for 6 to 8 weeks, although 6 weeks is adequate if a BACTEC or comparable system also is used. Examination for growth should occur twice weekly for the first 4 weeks and once weekly thereafter (2, 15, 31).

Identification Methods

The mycobacteria are generally straight or slightly curved rods measuring 0.3 to 0.6 µm × 1 to 4 µm. Mycelia are observed rarely, and when produced, fragment easily into coccobacillary units.

On subculture to traditional agar media, growth rate differentiates the mycobacteria: rapid growers require less than 7 days for visual recognition under optimal conditions; slow growers require more than 7 days. With additional use of pigmentation and some key biochemical studies, it is possible to identify most of the more common clinical isolates.

Members of the strictly pathogenic MTB complex are slow growers and, having buff-colored colonies, are considered nonpigmented. Selection of key tests for this group of organisms should include production of niacin and heat-stable catalase as well as evaluation for nitrate reduction. Susceptibility to thiophen-2-carboxylic acid hydrazide (TCH) further differentiates *M. bovis* from MTB (2, 15).

Nontuberculous mycobacteria may be initially characterized by growth rate and to some degree by their pattern of pigment production (old Runyon groups). *M. chelonae* and *M. fortuitum* are not considered pigmented, but may have carotenoid pigments. A few species of NTM may have varying characteristics of pigment production, thereby limiting the usefulness of the Runyon classification. Thus, on occasion, *M. szulgai* may show scotochromogenicity at 35°C and photochromogenicity at 25°C; rarely, scotochromogenic strains of the normally photochromogenic *M. kansasii*, as well as slightly pigmented strains of nonphotochromogens, may be encountered.

Identification of species may be accomplished using the following key biochemical tests (2):

I. Photochromogens: pigment, nitrate reduction, Tween 80 hydrolysis, arylsulfatase (3 day, 28°C), growth at 28°C, niacin, and semiquantitative catalase;
II. Scotochromogens: nitrate reduction, semiquantitative catalase, Tween 80 hydrolysis, urease, arylsulfatase (3 day, 42°C), 5% NaCl tolerance, growth at 52°C, and photoactive pigment at 25°C;
III. Nonphotochromogens: niacin, nitrate reduction, semiquantitative catalase, Tween 80 hydrolysis, urease, NaCl tolerance, arylsulfatase (3 day);
IV. Rapid growers: nitrate reduction, 5% NaCl tolerance (28°C), arylsulfatase (3 day), semiquantitative catalase, 68°C catalase, MacConkey agar with crystal violet, iron uptake (28°C), and utilization of sodium citrate, inositol, and mannitol.

Thorough descriptions of test methods and extensive biochemical reaction keys have been published (2, 15, 31).

Infrequently used at present, the BACTEC 460 system also has been used to successfully differentiate between members of the MTB complex and the NTM. The NAP (*p*-nitro-α-acetylamino-ß-hydroxy-propiophenone) differentiation test uses the characteristic inhibitory capacity of NAP against isolates of MTB complex and its incapability to inhibit the NTM. Using this system the two groups can be differentiated within 5 days of isolation.

Introduction of genetic probes has allowed for rapid and specific identification of many mycobacteria. These probes are directed at mycobacterial ribosomal RNA and are available for specific identification of members of the MTB complex as a whole, of *M. avium*, *M. intracellulare*, *M. kansasii* and *M. gordonae* (AccuProbe; Gen-Probe, San Diego, CA). Sensitivities and specificities of the probes for isolate identifications are greater than 99%, although rare cross-reactions have been reported between the *M. tuberculosis* complex probe and *M. terrae* and *M. celatum* complex (2). They cannot distinguish between *M. tuberculosis*, *M. bovis*, *M. africanum*, and *M. microti*, although this does not normally present a dilemma except with *M. bovis* (inclusive of the BCG

strains), which may be isolated in the laboratory. The probes can be used on isolates grown in broth or on solid culture media and their use in conjunction with the newer liquid broth based isolation systems makes them rapid, cost-effective and a practical means for identification of the most commonly isolated, clinically relevant mycobacteria. Such applications have significantly reduced the time to detection and identification of the mycobacteria. However, poor sensitivity precludes their use for direct mycobacterial detection in specimens at this time (2, 15, 31).

Other molecular means for identification of the mycobacteria have included sequencing using the hypervariable regions of the 16S rRNA gene (16S rDNA), 16S rRNA, and RNA polymerase gene (*rpo*B; 2, 14). A number of studies have used PCR amplification of a variety of genes followed by enzymatic restriction fragment length polymorphism analysis (RFLP) to identify the mycobacteria (11, 12, 13). PCR of the 65-kDa heat shock protein-encoding gene followed by RFLP looks promising in its ability to be clinically applicable in the near future (2).

Immunologic probes using radioimmunoassay, ELISA, and immunoblot technology have not been applied extensively in clinical laboratories but have in the past shown some promise as inexpensive alternatives for rapid mycobacterial identification.

Gas-liquid chromatography (GLC) for short-chain fatty acids has shown to be a promising tool for mycobacterial identification but is limited by its presently limited database and by its inability in many circumstances to provide species level identification without biochemical test supplementation. Reverse-phase HPLC for analysis of mycolic acids as bromophenacyl esters provides the most distinct patterns for individual species identification (31). The latter method is especially useful for the identification of the NTM. However, GLC and HPLC methods require organisms grown in culture, are more complex and require extensive equipment set-up and training of personnel; they are useful alternatives for larger reference laboratories that may see a plethora of unusual isolates (2, 15, 31).

Direct Detection

Microscopy

For direct microscopic detection of mycobacteria, centrifugation of specimens such as CSF, urine, and homogenized sputum at a minimal 3,000 × g is recommended to increase yield. Pellicles found on CSF specimens should be cut and a portion smeared for staining. However, routine direct staining of urine, CSF, and bone marrow specimens may not be necessary and is not recommended (unless perhaps if the patient is suspected of having AIDS), because the yield in such specimens is negligible (2, 15).

All smears may be air dried and heat fixed on an electric hot plate at 50°C for approximately 30 minutes. In most instances, light (thin) smears should be prepared, as material from these has a lower tendency to flake off during staining. Care must be exercised with all unstained slides because any AFB present may still be viable.

Rapid and sensitive stain results can be achieved through concentration of sputum by means of cytocentrifugation of sputum which is first liquefied and decontaminated by addition of 5% bleach (equal volume). These specimens, however, cannot be then used for culture (2).

The high lipid content of the mycobacteria allows them to retain certain basic dyes better than other bacteria even after treatment with acid-alcohol. This characteristic provides the mycobacteria with the trait of being “acid-fast” (Fig. 54.3). Although they share acid-fastness to some degree with nocardia and rhodococci, the latter two show much less ability to retain the dyes.

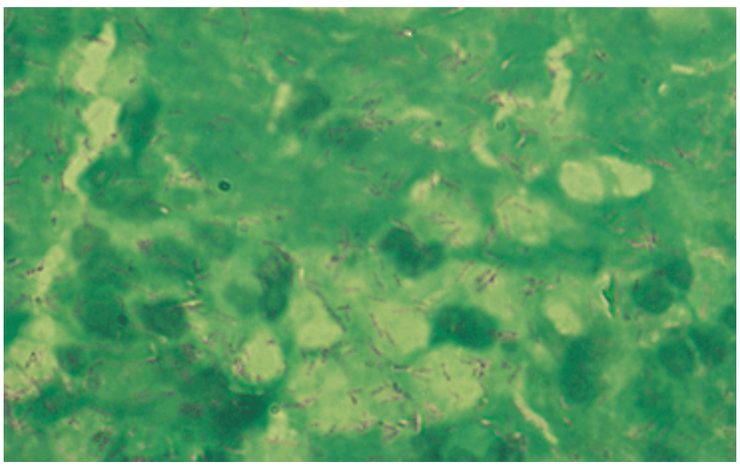


FIGURE 54.3. Acid-fast bacilli in a direct Kinyoun stain of synovial fluid yielding *Mycobacterium haemophilum*.

Several staining methods are available, including those using the aniline dye carbolfuchsin (Ziehl-Neelsen and Kinyoun stains) or fluorochrome dyes auramine O in conjunction with a counterstaining rhodamine. The auramine fluorochome stain has the same specificity as those using aniline dyes, but is a more sensitive method. Smears may be scanned for brightly fluorescing rods against a dark background at lower power (25× or 40× objectives). A mercury vapor or a halogen lamp (in an epifluorescent microscope) may be used as light sources in conjunction with appropriate filters. Although useful for scanning smears, the fluorochrome method has several problems, including inability to provide morphologic detail or to stain some isolates of rapidly growing mycobacteria (2, 15).

Of the two methods using carbolfuchsin, Kinyoun's method is simpler to perform because constant heating of the smear is not required. Otherwise, Kinyoun's cold stain and the Ziehl-Neelsen stain are comparable, decolorizing with a strong acid-alcohol (e.g., 95% ethanol and 3% hydrochloric acid). Examination of smears prepared by both methods requires use of 100× objectives and oil immersion.

The same characteristics that afford the mycobacteria acid-fast properties contribute to the difficulty of staining them by other stains, especially by the Gram method. Gram-stained preparations usually do not detect mycobacteria. However, in smears with heavy concentration of organisms, the rods may appear as poorly stained, beaded Gram-positive bacilli, or as “ghost” outlines that have taken up neither crystal violet nor safranin.

When scanning smears stained by carbolfuchsin (high magnification), at least 300 fields should be examined. Smears stained by the fluorochrome method may be scanned at a lower magnification, viewing at least 30 fields (2). Positive fluorochrome stained smears may be restained by the Kinyoun

method to better evaluate morphology if necessary. The immersion oil objective lens should be wiped clean after viewing a positive slide to avoid the chance of contamination of negative smears by carryover of AFB in the oil from the positive slide. Members of the MTB complex will typically be 2 to 8 µm long and may be slightly curved, whereas members of the NTM may have a banded appearance (*M. kansasii*) or have a pleomorphic morphology (*M. avium*). However, definitive identification cannot be made on microscopic morphology alone. Reporting of positive slides should include some form of quantitation to allow interpretation of the smears (e.g., using the fluorochrome stain at a magnification of 250X: if no AFB seen, report = no AFB seen; 1 to 2 AFB/30 fields or 1 sweep = doubtful, repeat specimen needed; 1 to 9 AFB/10 fields = 1+; 1 to 9 AFB/field = 2+; 10 to 90 AFB/field = 3+; >90 AFB/field = 4+) (2, 15). Full descriptions of methods and reporting formats may be found in references 2 and 15.

Other Means of Detection

Newer methods for the direct recognition of mycobacteria in specimens include HPLC analysis using a more sensitive fluorescent detection column and genomic amplification strategies using PCR or other nucleic acid amplification assays (2, 19, 38).

Of the PCR assays described, the majority are “home-brews” and not available commercially. Many of these use the IS6110 DNA insertion sequence which separates *M. tuberculosis* complex from other mycobacteria. Although home brew testing may be available through research facilities or private reference laboratories, care must be exercised when using such facilities because no standards exist and methods may vary widely.

A few amplification assays have been Food and Drug Administration (FDA)- approved and are commercially available as packaged kits; the first to become approved were the *Mycobacterium Tuberculosis* Direct Test (MTD test by GenProbe) and the AMPLICOR MTB assay (Roche Molecular Systems, Branchburg, NJ). The MTD test is based on transcription-mediated amplification of rRNA while the AMPLICOR assay relies on PCR methodology (2, 21, 24). Additionally, strand-displacement amplification, ligase chain reaction, Q-beta replicase-amplified assay, and nucleic acid sequence-based amplification assay have been introduced and are in the midst of evaluations (23).

Sensitivity of PCR and other nucleic acid amplification methods, both commercial and home brew, approaches 95% for smear positive sputum specimens, but varies between 40% to 80% for smear negative ones. Although not FDA approved for nonrespiratory specimens, some of the direct amplification methods have been evaluated on other specimen types such as CSF for tuberculosis with promising success (20). The clinical utility of direct amplification assays depends on the probability or heightened clinical suspicion for tuberculosis with each individual case. Studies should be limited to cases in which results would directly affect antituberculosis therapy, isolation procedures, or would limit further diagnostic testing (25).

Detection of mycobacterial antigens by immunoassay or of species-specific DNA or RNA sequences by gene probes presently is not adequately sensitive to recommend their routine application. These methods are no more sensitive than direct microscopy.

Interpretation of Results

Care must be exercised in interpreting results of direct smears of specimens and culture. Sensitivity and specificity of procedures are less important on the whole than are their accuracy for positive and negative prediction. The overall sensitivity and specificity of direct detection of AFB in sputum of patients with tuberculosis has been reported to be 22% to 80% and greater than 99%, respectively. Although good, the predictive value of a positive smear depends on the prevalence of disease. Furthermore, it is often difficult to microscopically differentiate between organisms belonging to the MTB complex and the NTM. Sensitivity and the capability of acid-fast smears to predict the absence of AFB in specimens are notoriously poor. Sensitivity can be increased by using auramine-rhodamine, by concentrating specimens (especially by cytocentrifugation), and by examining multiple specimens (2, 15).

Cross-contamination of culture tubes also has been known to occur, and care must be exercised to reduce or discover such a possibility. Inoculation order of tubes should be documented and positivity of consecutive (or closely following) specimens evaluated carefully for the possibility of carry-over contamination. All information derived from laboratory studies must be correlated to patient and clinical considerations (2, 15)

Multiple species may be present in a few infectious cases. Great care must be taken to ascertain that multiple morphologies are not present in primary broth cultures. Confirmation of the identity of an isolate identified as *M. tuberculosis* complex by genetic probe directly from a broth culture is still essential; this can be readily accomplished by documentation of niacin production using growth from the primary LJ slant.

Because the NTM are so commonly associated with the environment, their isolation from clinical specimens, especially sputum, does not immediately equate them with the etiologic agent of disease. The American Thoracic Society has proposed diagnostic criteria for relevance of *M. avium*, *M. kansasii*, and *M. abscessus* in respiratory tract secretions which are summarized below (30):

I. If three sputum or bronchoscopic washing specimens are available over the previous 12 months,
   a. 3 positive cultures (even if smears are negative), or
   b. 2 positive cultures plus one positive smear.

II. If only one bronchoscopic washing is available for evaluation,
   a. Positive culture (any quantity) with few (2+) or more organisms seen on direct smear, or
   b. Few (2+) or more colonies grow on primary solid media.

III. If sputum or bronchial washings are non-diagnostic or another process cannot be excluded,
   a. Positive biopsy (transbronchial or lung), or
   b. Compatible histopathology (granuloma or AFB on smear) and positive sputum culture.

Risk factors in the host, together with failure to identify other pathogenic etiologies, may provide additional clues. The usefulness of the histopathological evaluation of a specimen for documentation and correlation of the infectious process cannot be overstressed. Although these criteria were created to facilitate diagnosis of infections caused by the NTM, they also are useful for

suspected nocardial and other actinomycotic infections (2, 15, 30).

Therapeutic Considerations

Leprosy, tuberculosis, and the nontuberculous mycobacterioses require different treatment protocols. On one side of the spectrum, disease may be localized and self-limited (e.g., localized granuloma caused by *M. marinum*) or may require surgical excision alone (e.g., cervical lymphadenitis caused by *M. avium*, *M. scrofulacium*, or *M. haemophilum*). On the other side, they may require multiple drug regimens over an extended period of time (e.g., extensive pulmonary disease caused by resistant *M. tuberculosis* or *M. avium*).

Treatment of leprosy has depended on the stage of the disease. Dapsone and rifampin remained the primary agents for the tuberculoid stage (paucibacillary), while addition of clofazimine was recommended during the lepromatous stage (multibacillary). Ethionamide or prothionamide may be used in place of clofazimine (33). Newer agents that seem bactericidal to *M. leprae* but for which less experience has accumulated include minocycline, clarithromycin, and ofloxacin.

Antituberculosis therapy is dependent on susceptibility of the isolate and on the state of the disease process. Resistance, especially to isoniazid (INH), but also to multiple agents (multi-drug resistant *M. tuberculosis*, MDRTB) has increased. Variables associated with resistance include prior antituberculous therapy, patient's original country of residence, and his or her duration of stay in the United States. An increased incidence of resistance has been especially prominent in asiatics and hispanics and in several inner-city locations in the United States (2, 15, 39). Therapeutic protocols for tuberculosis have been recommended by the ATS in reference 39. In all cases initial treatment should include at least isoniazid, rifampin and pyrazinamide. When resistance is suspected ethambutol and/or streptomycin should be added. If resistance or adverse drug reaction is noted, the second-line of antituberculous agents can be considered and consists of para-aminosalicylic acid, ethionamide, cycloserine, capreomycin, kanamycin, amikacin, ciprofloxacin, ofloxacin, and rifabutin (2, 39, 40).

The final choice of agents of *M. tuberculosis* is aided by *in vitro* susceptibility studies which should be performed on every initial isolate and repeated if the sputum continues to grow MTB after 3 months of therapy (2, 40). Methods for such studies have been well established and determine the percent of bacilli in a population that are resistant to an agent. A poor clinical outcome is predictable with an agent to which more than 1% of bacilli in any single population are resistant. Thus, the proportional susceptibility method determines the percent of resistant organisms within a population being studied. Middlebrook 7H10 or 7H11 agar, impregnated with antituberculous agents at specific concentrations, is overlaid with inocula of known size. The plates are read quantitatively after an appropriate incubation period, with resistance being defined as growth of more than 1% of the initial inoculum as measured on the control plate (40).

The BACTEC radiometric system also has been applied to study *in vitro* susceptibilities of *M. tuberculosis* against INH, rifampin, ethambutol, and streptomycin with good results. Testing with pyrazinamide is more problematic because of the need to maintain a lower pH in the media. The BACTEC vial susceptibility testing system is a modification of the conventional proportion method, using radiolabeled Middlebrook 7H12 broth and radiometric measurements of growth. The amount of growth (as indicated by changes in the growth index) in the media with known drug concentrations compared with that in the control bottle has been correlated to the presence or absence of resistance in 1% of the inoculum (2, 40).

The conventional studies, as well as those using the radiometric system, may be performed on culture isolates (indirect method) or directly on concentrated specimens (direct method) in which AFB have been seen on smear. Care must be used in interpreting direct tests because of variables present, and results should be verified by indirect studies (40).

Many of the nontuberculous mycobacterioses are not amenable to antimicrobial therapy with antituberculous agents. Most clinically significant NTM are resistant to the first-line antituberculous agents. *M. kansasii*, alone, being commonly treatable with protocols using rifampin together with isoniazid and ethambutol or streptomycin, is only slightly more difficult to treat than *M. tuberculosis* (30). Members of MAC, however, are notoriously resistant and often require regimens combining more than four agents as described in the ATS guidelines (30). Surgical intervention is often recommended for localized disease, especially lymphadenitis.

The rapidly growing mycobacteria are also difficult to treat, being primarily resistant to conventional antituberculous agents. Approaches to therapy often include surgical intervention and debridement of localized infection together with antimicrobial therapy consisting of multiple agents. *M. fortuitum* usually is susceptible to amikacin, ciprofloxacin, sulfonamides, imipenem, cefoxitin, and clarithromycin, and variably susceptible to doxycycline. *M. abscessus* generally is susceptible to clarithromycin, amikacin, and cefoxitin, and has variable susceptibility to imipenem and ciprofloxacin. *M. chelonae* usually is susceptible to tobramycin, clarithromycin, and amikacin, and is variably susceptible to imipenem, doxycycline, and ciprofloxacin.

Historically, conventional *in vitro* susceptibility studies of NTM have had poor predictive value because antimicrobial concentrations for the evaluation of clinical usefulness have not been established. Recent evidence, however, indicates that such studies using microbroth dilution or the gradient strip (Etest, AB BioStrip, Sweden) can be helpful with certain drug selections noted below in the ATS guidelines (30):

I. *M. kansasii*: test rifampin only, and only if patient fails empiric or prophylactic therapy with that agent.
II. *M. avium*: test clarithromycin only, and only if patient fails empiric or prophylactic therapy with that agent.
III. *M. abscessus*, *M. chelonae*, *M. fortuitum*: test for susceptibility to amikacin, doxycycline, imipenem, fluoroquinolones, sulfonamides, cefoxitin, clarithromycin, and tobramycin (*M. chelonae* only).

Suceptibility studies of other NTM-drug combinations are not standardized and not normally recommended.

NOCARDIA

Part of "54 - Aerobic Actinomycetes"

Microbiology and Classification

The genus *Nocardia* historically has contained a very heterogeneous group of organisms, primarily because species identification was based solely on morphologic criteria. Presently, the genus has been restructured with many species being reassigned to new or other genera (*Gordona*, *Actinomadura*, *Nocardiopsis*, *Oerskovia*, and *Rhodococcus*), leaving it more homogeneous (1, 3, 4 and 5, 41). The nocardia possess a type IV cell wall, meso-DAP, arabinose and galactose, but may be differentiated from the mycobacteria by the former's shorter chained (40- to 60-carbon) mycolic acids.

The genus *Nocardia* presently has eleven well defined species of which eight are of clinical importance (*N. asteroides*, *N. farcinica*, *N. brasiliensis*, *N. pseudobrasiliensis*, *N. otitidiscaviarum*, *N. nova*, *N. transvalensis*, and *N. brevicatena* complex). Characterization of the recognized nocardial species is based on a multitude of morphologic and physiological determinants that usually can provide an identification to the species level. *N. asteroides* contains two subgroups that differ in several characteristics and DNA homologies, whereas *N. transvalensis* contains four such subgroup taxa. Because much of the earlier literature did not diffentiate between *N. asteroides* (sensu stricto), *N. farcinica*, and *N. nova*, these three species have often been described as part of the *N. asteroides* complex (2).

Spectrum of Disease

In the United States, the nocardia are commonly considered opportunistic pathogens seen primarily in immunocompromised patients, in those receiving long-term corticosteroid therapy, or in those who have undergone traumatic percutaneous abrasions. Nocardial infections are especially prevalent in patients with parenchymal lung damage resulting from chronic lung disease. Unlike the mycobacteria, nocardial lesions are typically suppurative in nature, with progression to acute necrosis and abscess formation. Granulomatous lesions are rarely formed.

N. asteroides, *N. farcinica*, *N. nova*, *N. brasiliensis*, *N. pseudobrasiliensis* and *N. otitidiscaviarum* are the most common species associated with human infection. *N. transvalensis* and *N. brevicatena* are seldom encountered. Primary manifestations include pneumonia, localized abscess, brain abscess, bacteremia, peritonitis, synovitis, and other localized or disseminated infections.

N. asteroides, *N. farcinica*, *N. pseudobrasiliensis*, and *N. nova* are most commonly associated with pulmonary and (in the case of the first three species) disseminated disease, and less frequently with direct soft-tissue infection. *N. transvalensis*, which is encountered rarely in the United States, has a tendency to cause localized infection, although it can disseminate. In contrast, *N. brasiliensis* and *N. otitidiscaviarum* frequently involve soft tissue. However, both species occasionally may cause systemic and pulmonary disease. Lymphocutaneous sporotrichoid infections also may be encountered, especially with *N. brasiliensis*. Outside the United States, *N. brasiliensis*, *N. otitidiscaviarum*, and *N. transvalensis*, are associated with actinomycotic mycetoma, a localized, progressive, yet chronic infection of skin and soft tissue (1, 3, 8, 42, 43). The organisms are probably introduced into subcutaneous tissue through traumatic inoculation via thorns or splinters, with extremities being involved most frequently. Infected sites swell, and lesions may progress to form draining sinus tracts. Small white granules often may be present in the draining material, representing masses of nocardial cells compacted together. Such mycetomas are seen characteristically in older soft tissue infections.

Epidemiology

The clinically significant nocardial species normally reside as saprophytes in soil and may be associated with plant material as well as water, air, and dust. Their geographic distribution is probably responsible for the variation in species' prevalence in association with disease. The most common species, *N. asteroides* complex, are found throughout the United States, although it is the author's personal observation that numbers seem higher in the southwest. *N. brasiliensis* is commonly associated with tropical areas, although in the United States it also seems to have a higher prevalence in the southwest and southeast. *Nocardia otitidiscaviarum* is an infrequent isolate from soil throughout the world, while the environmental source of *N. transvalensis* has yet to be elucidated (1).

The majority of *N. asteroides* complex and *N. pseudobrasiliensis* infections in the United States are acquired via the respiratory tract, although traumatic inoculation may also occur. Person-to-person transmission of nocardiosis is uncharacteristic, but acquisition in small clusters of patients in close proximity to each other suggests that possibility, but does not rule out environmental exposure in a specific locale (44, 45). *N. brasiliensis* is primarily acquired percutaneously, although respiratory acquisition is possible. Infection with this genus, as with *N. otitidiscaviarum*, is usually primary and is most frequently seen in the normal host.

Diagnostic Procedures

Specimen Collection

The same principles used for mycobacteria specimen collection can be used for nocardia specimens. Respiratory specimens are the most common, but tissues, other normally sterile body fluids, as well as pus, exudates, and even blood may yield nocardia in the right clinical setting. Storage at colder temperatures may be deleterious to some nocardia (1).

Primary Isolation

Nocardia are able to grow on most readily available routine media such as 5% sheep blood and chocolate agars, Mueller-Hinton agar, tryptic soy agar, various broths, Sabouraud dextrose agar, brain-heart infusion agar, and Lowenstein-Jensen agar. Unfortunately, overgrowth of the nocardia by more rapidly growing bacteria and fungi in heavily contaminated clinical specimens is a common problem (1, 3, 4 and 5).

TABLE 54.5. PHENOTYPIC DIFFERENTIATION BETWEEN THE CLINICALLY SIGNIFICANT SPECIES OF *NOCARDIA*

Characteristic	*N. asteroides (sensu stricto)*	N. farcinica	N. nova	N. brasiliensis	N. pseudobrasiliensis	N. otitidiscaviarum	N. transvalensis
Decomposition of:							
Casein	No	No	No	Yes	Yes	No	No
Xanthine	No	No	No	-	-	Yes	No
Hypoxanthine	No	No	No	-	-	-	Yes
Adenine	-	-	-	No	Yes	-	-
Arylsulfatase (14 day)	No	No	Yes	-	-	-	-
Acid from L-Rhamnose	No	Yes	-	-	-	-	-
Growth at 45°C	Variable	Yes	No	No	No	Variable	No

Compiled in part from reference 1.

Because of the overgrowth problem, several primary plating media, including selective media, should be used to optimize isolation.

Selective buffered charcoal yeast extract agar (BCYE) used for legionellosis, as well as Thayer-Martin, Lowenstein-Jensen, and Sabouraud dextrose agars may be used in conjunction with routine blood or chocolate agar plates to isolate nocardia selectively. Selective media may be inhibitory to some nocardia and should not be used alone.

A paraffin-baiting technique, using sterile carbon-free broth and a paraffin-coated glass rod, has been described and is reported to almost double isolation rates for nocardia. The method relies on the propensity of nocardia to invade and use the paraffin on the glass rod as a growth medium. However, the technique has not been popular in clinical laboratories because of its cumbersome nature and longer time to recovery. Pretreatment and processing of contaminated specimens with N-acetyl-L-cysteine or other decontaminating methods used for the mycobacteria are deleterious to the nocardia. Such treatment can significantly decrease viability and should not be used as a primary isolation method. However, isolates can survive in some instances, and nocardia can be isolated in the mycobacterial section of the laboratory (1).

Primary media should be incubated aerobically at 35° for up to 2 to 3 weeks. Cultures also may be set up simultaneously at 25°C, especially if the identity of the potential pathogen is not suspected. Although CO_2 enhancement of nocardial growth has been reported, experience has shown that CO_2 is not necessary for initial isolation of fresh isolates. Cultures should be examined every 2 days for the first week and at least twice weekly thereafter. Clear media are superior for early microscopic recognition of nocardial colonies, which often can be distinguished by their filamentous appearance and possible aerial mycelia. Colonies most frequently are chalky white, turning yellow-orange with age.

Identification

Nocardial isolates may be recognized once the colonies reach a size that will show their chalky aerial mycelia. These are best observed under a dissecting microscope. Commonly, young colonies are pure white, turning to an orange color with age. Staining will show Gram-positive to stippled, branching rods, or occasionally coccobacilli indicative of hyphal fragmentation. A certain percentage of the population of organisms from a colony will stain acid-fast when a weak acid decolorizing solution (0.5% to 1% sulfuric acid) is used with the modified Kinyoun stain.

Resistance of nocardial isolates to lysozyme differentiates them from other streptomycetes. Although the rhodococci may have a variable response to lysozyme, they are normally less acid-fast than nocardia, and their colonies are commonly salmon-pink. The nocardia and rhodococci can further be presumptively differentiated by their susceptibility to vancomycin and erythromycin. Presence of aerial hyphae in the *Nocardia* also differentiates that genus from all other related species (2).

Clinically significant species of nocardia may be differentiated by testing hydrolysis of casein, xanthine, and hypoxanthine; arylsulfatase production; growth in L-rhamnose; and adenine decomposition (Table 54.5 and Table 54.6). Schematic flow charts and a full review of characteristics have been described recently by Brown, McNeil, and Desmond (2).

TABLE 54.6. ADDITIONAL STUDIES TO DIFFERENTIATE THE SPECIES *N. ASTEROIDES*, *N. FARCINICA*, AND *N. NOVA*

Test	% Positivity (P) or Susceptibility (s)		
	N. asteroides	*N. farcinica*	*N. nova*
Growth at 45°C (3 days)	43 P	100 P	5 P
Rhamnose	10 P	80 P	5 P
Acetamide	17 P	80 P	0 P
Arylsulfatase	5 P	0 P	75 P
Cefamandole (S:>20 mm)	95 S	7 S	100 S
Tobramycin (S:>20 mm)	83 S	0 S	17 S
Erythromycin (s:>30 mm)	2 S	0 S	100 S

[a] Compiled from Wallace RJ, Brown BA, Tsukamura M, et al. Clinical and laboratory features of *Nocardia nova*. *J. Clin Microbiol* 1991; 29:2407–2411, and Schiff TA, McNeil MM, Brown JM. Cutaneous *Nocardia farcinica* infection in nonimmunocompromised host. *Clin Infect Dis* 1993; 16:756–760.

More definitive identification sometimes requires evaluation of whole-cell hydrolysates and the isolates' utilization of a variety of organic compounds. These more definitive studies are not commonly available and are not normally needed for most identifications. Identification at the molecular level also has helped define species using PCR, genetic hydridization, and sequence of 16S rRNA studies. These are beyond the capability of the normal laboratory and appropriate reference laboratories should be sought when clinically indicated.

Direct Detection

The nocardia may frequently be detected in routine Gram-stained smears of specimens as long, delicate, branched, beaded to Gram-positive filaments (Fig. 54.4). Shorter rods and coccal

forms are also possible. The Gram stain is probably the most sensitive stain for the detection of these organisms. Although the nocardia have “weak” acid-fast properties, they do not complex the basic dyes as tenaciously as the mycobacteria, and they require a milder acid-alcohol wash. Thus, a modified Kinyoun stain using 0.5% sulfuric acid decolorizer will stain 60% to 90% of nocardia in specimens. Not all nocardia within a single population or specimen will stain acid-fast (partial acid-fastness), and the property may be lost through aging of cultures. A modified auramine-rhodamine stain also may be used to provide an opportunity for detection by fluorescent microscopy.

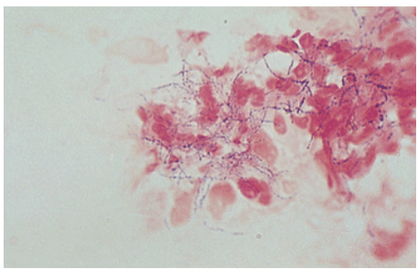


FIGURE 54.4. Direct Gram stain of sputum showing delicate, branched, beaded to Gram-positive filaments characteristic of *Nocardia* species (original magnification of 1,000 X).

Unlike material from actinomycetomas, specimens from respiratory, cutaneous, and systemic nocardiosis usually do not contain actinomycotic sulfur granules. When present, granules may be crushed between two slides and observed by wet mount or by Gram stain for typical branching filaments. A modified acid-fast stain may then be performed for identification. Negative stains, however, cannot rule out nocardial organisms. The Gram-Weigert, Brown and Brenn, and modified acid-fast stains may all detect actinomycetes in tissue and sulfur granule preparations, whereas the hematoxylin and eosin stain will not (1, 3, 41).

Therapeutic Considerations

Sulfamethoxazole-trimethroprim (SXT) and the sulfonamides together with minocycline, amikacin, and imipenem remain the drugs of choice for nocardiosis. Combination therapy with either amikacin or imipenem plus SXT often is recommended for disseminated and central nervous system infections. Treatment should be prolonged for a minimum of 2 to 12 months depending on the severity and progress of the infection. Recovery or improvement is normally best in pulmonary or localized soft tissue infections and worst in disseminated disease (8). Described alternative agents include the third-generation cephalosporins, amoxicillin-clavulanic acid combination, and the fluoroquinolones, although their activity *in vitro* against the nocardia has been variable.

The variable response of *N. asteroides* to a number of antimicrobials suggests that isolates be tested *in vitro* for specific guidelines (1, 8). Although methods for susceptibility studies of the nocardia have not officially been standardized, microbroth dilution, disk agar diffusion, gradient strip agar diffusion, and radiometric broth bottle determinations of MICs or susceptibilities have been well described (46, 47 and 48).

The broth-microdilution test system using cation-supplemented Mueller-Hinton broth and the radiometric broth BACTEC vial system correlate fairly well the disk diffusion method (46, 47 and 48). The broth microdilution and gradient strip methods are practical for application in many laboratories. Results must be noted as being nonstandardized. Further study is needed to confirm clinical applications and correlations with outcomes.

RHODOCOCCUS, *GORDONA*, AND *TSUKAMURELLA*

Part of "54 - Aerobic Actinomycetes"

Microbiology and Classification

Members of this group of organisms are diverse in their morphology and growth patterns. Their grouping and identification are based primarily on cell wall constituents (being those of type IV, with mycolic acid sizes ranging from 32 to 66 carbons in length) as well as biochemical criteria.

Considered originally as rhodochrous strains of the genus *Mycobacterium*, the rhodoccocci were placed in a separate genus in the 1970s. A redefining of the genus *Rhodococcus*, with removal of several species to the genus *Gordona* and renaming of one species as *Tsukamurella*, has left nine species intact. The rhodococci contain 34 to 52 carbon mycolic acids, no mycobactins, and eight isoprene dehydrogenated menaquinones. Only one species, *Rhodococcus equi*, has been isolated from humans and is associated with human infection (1, 49, 50). Other species not associated with disease in humans include *R. coprophilus*, *R. erythropolis*, *R. fascians*, *R. globerulus*, *R. marinonascens*, *R. rhodnii*, *R. rhodocrous*, and *R. ruber*.

The genus *Gordona* was reintroduced in 1988 to accept transfer of several species from the genus *Rhodococcus* based on molecular studies and characterized by differences in mycolates (48 to 66 carbon atoms) as well as presence of nine isoprene units in the dihydogenated menaquinoes. Presently, the species in this genus associated with human disease include *G. bronchialis*, *G. rubroperctinctus*, *G. sputi*, *G. terrae*, and *G. aichiensis*. Species not associated with humans include *G. amarae*, *G. hydrophobica*, and *G. hirsuta* (1, 49, 50).

Tsukamurella paurometabola is the primary species in the genus named after Tsukamura (in 1988), who in 1971 first studied the organism isolated from sputum. Four other species alloted to this genus include *T. inchonensis*, *T. pulmonis*, *T. tyrosinosolvens*, and *T. wratislaviensis* (1).

Spectrum of Disease

Rhodococci and related genera *Gordona* and *Tsukamurella* rarely are involved in human disease, but seem to be increasing in their role as opportunistic pathogens in the immunocompromised host.

Primarily a pathogen of livestock, *R. equi* is the species most commonly associated with human infections in this group. The organism has been implicated with pulmonary disease, including lung abscess and cavitation, often concomitantly with bacteremia in patients with neoplastic disease or on immunosuppression (5, 51).

Separate reports have implicated *R. equi* in localized infection of a cervical lymph node as well as in endophthalmitis (52). The organism also may cause disease in patients with implanted Hickman catheters or in those with AIDS (53).

Infections caused by *Gordona* species are very rare, and colonization or contamination of specimens may occur. These species have been implicated in primary cutaneous infection (*G. terrae*), chronic pulmonary infection (*G. aichiensis*, *G. bronchialis*, and *G. rubropertincta*), catheter-associated bacteremia, and a brain abscess (1).

The genus *Tsukamurella* also is a rare opportunistic pathogen, having been associated with immunocompromised patients and those undergoing dialysis and with indwelling catheters; bacteremias have been documented on several occasions (1).

Epidemiology

The rhodococci may be found in a number of environmental locations, especially soil, but also fresh water, marine habitats, the gut of some arthropods, and feces of herbivores, swine, and fish. The presence of fecal contamination significantly increases the multiplication of the organisms in the soil. Isolation of the organism from sputum and central lines suggests its acquisition from air or dust, which may be inhaled or may contaminate medical devices. Often however, patients do not have a history of exposure to animals (53). Corticosteroids and immunocompromising conditions enhance patient susceptibility to clinically manifested infection (1, 4, 49, 53).

Gordona also are well distributed in nature, with various species being isolated from soil, aeration tank foam in sewage treatment plants, biofilters from waste gas treatment, and sputum. *Tsukamurella* species, which prefer cooler growth temperatures, have been isolated from soil, sludge and arthropods (1, 4).

Diagnostic Procedures

Specimen Collection

Follow the same guidelines given for the nocardia. Bacteremia may be present during rhodococcal and related organism infection of immunocompromised patients and of long-term indwelling lines.

Primary Isolation

On culture plates, colonies of *Rhodococcus equi* initially are clear to white, nonhemolytic, rounded, and frequently mucoid and often turn salmon-pink to red within 2 to 5 days (Fig. 54.5.). Pigmentation may vary from red to cream. Growth is good between 15°C and 40°C, but is slow, often requiring several days for visual observation. Microscopically, the organisms are Gram-positive and coccobacillary or bacillary, resembling other coryneform bacteria. A small number of cells within a population may be stained by the modified Kinyoun weak acid-fast stain, contributing to their being confused with fragmented nocardial cells. Unlike the nocardia, however, rhodococci are susceptible to vancomycin and erythromycin *in vitro*.

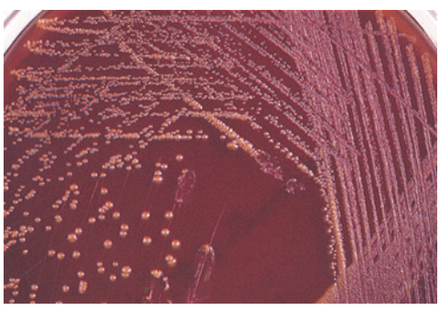


FIGURE 54.5. Colonies of *Rhodococcus* species with typical salmon-pink to red pigmentation (shown on sheep blood agar).

In addition to colonial and microscopic recognition, the rhodococci may be preliminarily recognized and identified by biochemical characteristics using API Coryne system (bioMerieux, Maarcy l'Etoile, France), API ZYM, and the VITEK GPI card (bioMeriux Vitek, Hazelwood, MO) (53). Supplemental tests may include the presence of *equi* factor (CAMP-like test), catalase, urease, lipase, oxidase, orthonitrophenyl-ß-D-galactopyranoside (ONPG), growth on MacConkey agar, and hydrolysis of xanthine, tyrosine, hypoxanthine and gelatin (53). Definitive identification is difficult and may require cell analysis for meso-DAP, arabinose, and galactose, as well as more extensive physiological and molecular testing (1, 4, 5, 7, 51, 52 and 53).

Direct Detection

The rhodococci are seen typically as Gram-positive short rods or coccobacilli and are confused frequently with the diphtheroids. Although having partially acid-fast properties, these organisms are seen less frequently as being acid-fast than the nocardia. The organisms do not have characteristic morphologies that would lead one to suspect them as being rhodococci on direct Gram stain. However, the histopathology in bronchial secretions may show characteristic pleomorphic Gram positive bacteria engulfed by polymorphonuclear leukocytes. Microabscesses, pseudotumors and malakoplakia (granulomatous inflammation) also may be evident by histopathology (1).

Therapeutic Considerations

The rhodococci normally are susceptible to erythromycin, rifampin, vancomycin, the aminoglycosides, penicillin G, and doxycycline. Because of the organisms' intracellular nature, the combinations of erythromycin or imipenem with rifampin seem to act synergistically. Fluoroquinolones have been suggested as alternative therapy, but resistant strains have been noted in Asia (53). Duration of therapy is uncertain, although many weeks may be required because relapses may occur after shorter periods. Surgical intervention also may be necessary in instances. Standardized susceptibility tests are not available, although NCCLS guidelines for the coryneform bacteria may be applied (1, 3, 7, 49, 53, 54).

ACTINOMADURA, *DERMATOPHILUS*, *NOCARDIOPSIS*, AND *STREPTOMYCES*

Microbiology and Classification

Species within the genus *Actinomadura* share a chemotype III cell wall and contain the sugar madurose. These organisms are Gram-positive rods that do not fragment and produce both substrate and aerial branched hyphae. They can be distinguished easily from *Dermatophilus* morphologically. Originally classified in the genus *Nocardia*, the *Actinomadura* were later given individual genus status. The genus contains 26 species, of which three (*A. madurae*, *A. pelletieri*, and *A. latina*) are clinically significant. Previous members such as the species *dassonvillei* were found to lack madurose and were transferred to the genus *Nocardiopsis* (1, 4).

The genus *Dermatophilus* contains but a single species, *Dermatophilus congolensis*, an obligate animal pathogen with meso-DAP and madurose as major cell-wall components. The organism characteristically divides by transverse, longitudinal, and horizontal septation of the mycelium, leading to the formation of motile zoospores. A tough, gelatinous capsule is formed surrounding the branching vegetative mycelium at points of division where zoospores mature, representing a multilocular sporangium (1, 55).

Nocardiopsis contains seven species, only one of which really may be considered pathogenic (*N. dassonvillei*). The organisms produce characteristic chains of arthrospores which fragment from aerial as well as substrate hyphae, contain meso-DAP and a chemotype III cell wall, but lack mycolates (1, 3).

The genus *Streptomyces* contains a large number of frequently poorly defined species of sporoactinomycetes with cell-wall type I (L-DAP, glycine but no sugar). Of all the species present in this genus, only *Streptomyces somaliensis* and *Streptomyces anulatus* (*griseus*) are commonly associated with human disease (1, 3).

Spectrum of Disease

Infectious manifestations of members of the genera *Actinomadura*, *Nocardiopsis*, and *Streptomyces* are primarily limited to the formation of localized, suppurative, mostly chronic lesions. These mycetomas occur commonly in parts of the body coming into frequent contact with soil or that may be abraded or traumatized. The most common human pathogens are *A. madurae*, *A. pelletieri*, and *S. somaliensis*, and less frequently, *S. griseus* and *N. dassonvillei*. The organisms gain entry into subcutaneous tissues through slight or imperceptible breaks in the skin. A painless nodule forms, becomes fluctuant, and usually forms a sinus tract through which pus is discharged. Draining pus often contains pigmented granules made up of filamentous, compacted microcolonies of the organism and ranging in color from white to yellow (*A. madurae* and *S. somaliensis*) or red to pink (*A. pelletieri*). Additional nodules and sinus tracts form over time, with fibrous tissue eventually deforming the areas involved. Unchecked, progression of the lesions may involve connective tissue, muscle, and bone. The immune response to actinomycotic infections characteristically is pyogenic. The disease remains localized. Hematogenous spread is encountered rarely (except with nocardia), although spread via the lymphatics to involve lymph nodes is possible (26, 42, 43).

D. congolensis, the only species in the genus, causes streptrotrichosis in animals; this dermatophilosis is a pustular, exudative dermatitis affecting many wild and domestic animals, especially sheep, cattle, and horses. The organism is found obligately on animals, never having been isolated from the environment. It may be transmitted to humans, either directly through contact with infected animals or possibly through bites by arthropods that may carry infective zoospores. Infection, however, probably requires some form of dermal abrasion because *Dermatophilus* is not invasive. Presenting either as multiple pustules or as a desquamative keratitis of the hands and forearms, the infection usually is self-limiting in humans (1, 4, 55).

Epidemiology

The clinically significant species of *Actinomadura*, *Nocardiopsis*, and *Streptomyces*, primarily soil inhabitants, have diverse geographic prevalences, although none is common in the United States. Mycetomas caused by *A. madurae* are cosmopolitan, although tending to be more frequent in tropical or subtropical regions. The organism is isolated frequently in North and South America, Africa, and Asia. *A. pelletieri* has a predilection for areas with heavy rainfall and is more restricted to areas of Africa and South America (26, 42). *S. somaliensis* prefers arid regions, sandy soil, with vegetation consisting of thorny plants. It therefore causes mycetomas most frequently in Africa, Mexico, and portions of South America, although it may be found in other continents as well.

Dermatophilus is an obligate parasite associated with livestock and other feral animals. Direct contact of abraded skin with material such as skin or hair of diseased animals carrying infective zoospores is the normal route of transmission. The disease is uncommon in the United States (1, 55).

Diagnostic Procedures

Specimen Collection

Biopsy of deeper tissue is preferred for the evaluation of actinomycotic mycetomas, because surface material may be contaminated or colonized. Other specimens may include pus, drainage fluid, or sinus scrapings either from open lesions or from aspirates.

Specimens for diagnosis of *Dermatophilus* infections include material from unopened pustules, exudates, biopsies, scrapings, and scabs. Specimens should be submitted in sterile tubes or in adequate volume on premoistened sterile swabs and should be kept at room temperature (1, 55).

Primary Isolation

Culture methods for the actinomadurae, the nocardiopsis, and the streptomycetes are similar to those for the nocardia. The actinomadurae may at times be difficult to isolate from clinical specimens. The former, together with the streptomycetes, may not grow on all of the media described for the nocardia, but may produce small colonies on nocardia selective agars. Thus, such media should be supplemented by other routine media as outlined

for the nocardia, as well as by Sabouraud dextrose agar, both with and without chloramphenicol (0.05 mg/mL). Incubation should be at 26°C as well as at 35° for up to 3 weeks (1, 3).

Isolation of *Dermatophilus* is difficult, but may be accomplished by inoculating the specimens onto beef infusion-blood agar and incubating for up to a week aerobically at 35°. Plates also may be incubated in the presence of 5% to 10% CO_2. Specimens with heavy contamination may require selective processing, employing either animal passage or the predilection of zoospores for CO_2 (55).

Identification

Members of the genus *Actinomadura* normally are slow-growing, with colonies that are glabrous or waxy, heaped, and either white to cream (*A. madurae*) or pink to coral red (*A. pelletieri*) in color. Microscopically, they are Gram-positive, non–acid-fast, bacillary organisms with branching vegetative mycelia. The mycelia of the two clinically important species only rarely fragment and usually do not produce aerial forms. Streptomycetes, on the other hand, are fast-growing and produce colonies similar to those of the actinomadurae.

Simplified differential identification schemes may include cell-wall analysis for the presence of meso-DAP and madurose in the actinomadurae, of meso-DAP without characteristic sugars in nocardiopsis, and of L-DAP in the streptomycetes. *A. madurae* and *A. pelletieri* can be differentiated by the red coloration of the latter's colony (1, 26).

Dermatophilus may be recognized as tiny, pitting colonies after incubation for 24 hours, turning orange within 2 to 5 days with ß-hemolysis evident in areas of heavy growth. The single species in this genus can be identified by the characteristic microscopic morphology, which is similar to that seen in clinical specimens. Colonial growth is optimal at 37°C when enriched media are used, and production of aerial mycelium is promoted by 10% CO_2. Whole-cell hydrolysate analysis for DL-DAP and madurose, together with casein and starch hydrolysis may help in furthering identification of a few strains, but such studies are rarely needed (55).

Direct Detection

In cases of mycetoma, biopsy or drainage material should be examined for granules, which usually are hard and of variable sizes (1, 26). The granules should be washed with sterile saline and then examined microscopically after emulsification in a drop of 10% potassium hydroxide or after crushing between two slides. They also can be stained by the Gram method to visualize the tangled filamentous forms of the actinomycetes, thereby differentiating them from the broader, cross-walled mycelia of the fungi. A positive modified acid-fast stain also can differentiate between the nocardia and the actinomadura and streptomyces; however, negative stains also may occur with the nocardia. As with the nocardia, the Brown and Brenn's modification of the Gram stain and the Gram-Weigert stain will detect the aerobic actinomycetes in tissue (1, 26, 42, 43).

Direct detection of *Dermatophilus* is best accomplished using the Giemsa stain of submitted smears or tissue. The Gram-stain and methylene blue wet mount also may be useful. The organisms typically show 2 to 5 µm wide, branched filaments with transverse and longitudinal division. Packets containing coccoid cells may be present (55).

Therapeutic Considerations

Actinomycotic mycetoma responds to antimicrobial therapy, although regimens of several months' duration are necessary to curb relapse. Surgical debridement may enhance healing. Streptomycin sulfate, usually combined with another agent, is commonly suggested. The second agents recommended may be dapsone for *A. madurae* and *S. somaliensis*, and SXT for *A. pelletieri*. SXT also has been suggested for streptomycetes not initially responding to dapsone. Susceptibility studies are not recommended routinely (42, 43).

Infection with *Dermatophilus* normally is self-limited, and systemic therapy usually is not required.

OTHER ACTINOMYCETES

Part of "54 - Aerobic Actinomycetes"

Other actinomycetes of some medical importance include the genera *Oerskovia*, *Micromonospora*, and the three thermophilic genera *Saccharomonospora* spp., *Saccharopolyspora* spp., and *Thermoactinomyces* spp. These genera, however, are encountered rarely in clinical laboratories and are well described elsewhere (1).

Interesting characteristics of the oerskoviae include their being facultative anaerobes, having the ability to fragment into motile bacillary elements, and having a yellow pigment. Additionally, morphologic characteristics, absence of DAP and mycolates, and presence of type VI cell wall constituents, with lysine and galactose as major components, separate the oerskoviae from other aerobic actinomycetes (5). Found primarily in soil and plant debris, these organisms occasionally have been reported to cause infections in humans, usually in association with traumatic inoculation or contamination of indwelling foreign objects (56).

SUMMARY AND CONCLUSIONS

Part of "54 - Aerobic Actinomycetes"

Many of the aerobic actinomycetes have gained increased prominence as pathogens at a time when the number of patients with immunocompromising conditions also is on the rise. Sophisticated medical technology together with newer diseases that induce immunodeficiency are at least partially responsible. Added recognition of the infective capability of the aerobic actinomycetes and better understanding of their epidemiology, clinical course, and susceptibilities to antimicrobial agents are paramount to rapid diagnosis and better patient care. Molecular identification and typing techniques will bring a better understanding of the interrelatedness of the various genera and species in these groups of organisms. Clinical laboratories must become aware of the role played by the aerobic actinomycetes in disease and must recognize when identification, susceptibility testing, and therapeutic or surgical interventions are necessary.

REFERENCES

1. Brown JM, McNeil MM, Desmond EP. *Nocardia, Rhodococcus, Gordona, Actinomadura, Streptomyces*, and other actinomycetes of medical importance. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of Clinical Microbiology*. 7th ed. Washington, D.C.: American Society for Microbiology, 1999:370–398.

2. Metchock BG, Nolte FS, Wallace Jr RJ. *Mycobacterium*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of Clinical Microbiology*. 7th ed. Washington, D.C.: American Society for Microbiology, 1999:399–437.

3. McNeil MM, Brown JM. The medically important aerobic actinomycetes: epidemiology and microbiology. *Clin Microbiol Rev* 1994;7:357–417.

4. Goodfellow M, Cross T. Classification. In: Goodfellow M, Mordarski M, Williams ST, eds. *The Biology of the Actinomycetes*. London: Academic Press, 1984:7–164.

5. Lechevalier HA, Goodfellow MA, Lechevalier MP, et al. Nocardioform actinomycetes. In: Williams ST, Sharpe ME, Holt JG, eds. *Bergey's Manual of Systematic Bacteriology*. Baltimore: Williams & Wilkins, 1989;4:2348–2382.

6. Wayne LG, Kubica GP. Genus *Mycobacterium*. In: Holt JG, ed. *Bergey's Manual of Systematic Bacteriology*. Baltimore: Williams & Wilkins, 1986;2:1436–1457.

7. Coyle MB, Lipsky BA. Coryneform bacteria in infectious diseases: clinical and laboratory aspects. *Clin Microbiol Rev* 1990;3:227–246.

8. Lerner PI. Nocardiosis. *Clinical Infectious Diseases* 1996;22:891–905.

9. Ellner PD, Kiehn TE, Cammarata R, et al. Rapid detection and identification of pathogenic mycobacteria by combining radiometric and nucleic acid probe methods. *J Clin Microbiol* 1988;26:1349–1352.

10. Butler WR, Kilburn JO. High-performance liquid chromatography patterns of mycolic acids as criteria for identification of *Mycobacterium chelonae*, *Mycobacterium fortuitum*, and *Mycobacterium smegmatis*. *J Clin Microbiol* 1990;28:2094–2098.

11. Steingrube VA, Wilson RW, Brown BA, et al. Rapid identification of clinically significant species and taxa of aerobic actinomycetes, including *Actinomadura*, *Gordona*, *Nocardia*, *Rhodococcus*, *Streptomyces*, and *Tsukamurella* isolates, by DNA amplification and restriction endonuclease analysis. *J Clin Microbiol* 1997;35:817–822.

12. Wilson RW, Steingrube VA, Brown BA, et al. Clinical application of PCR-restriction enzyme analysis for rapid identification of aerobic actinomycete isolates. *J Clin Microbiol* 1998;36:148–152.

13. Devallois A, Goh KS, Rastogi N. Rapid identification of mycobacteria to species level by PCR-restriction fragment length polymorphism analysis of the *hsp65* gene and proposition of an algorithm to differentiate 34 mycobacterial species. *J Clin Microbiol* 1997;35:2969–2973.

14. Kim BJ, Lee SH, Lyu MA, et al. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (rpo B). *J Clin Mirobiol* 1999;37:1714–1720.

15. Cumitech. Laboratory diagnosis of the mycobacterioses. In: Weissfeld AS, ed. *Cumulative Techniques and Procedures in Clinical Microbiology, Number 16A*. Washington DC: American Society for Microbiology, 1994.

16. Association of State and Territorial Public Health Laboratory Directors and U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention. *Mycobacterium Tuberculosis: Assessing your Laboratory*. 1995. Association of State and Territorial Public Health Laboratory Directors and Centers for Disease Control and Prevention, Atlanta, GA.

17. Tenover FC, Crawford JT, Huebner RE, et al. The resurgence of tuberculosis: is your laboratory ready? *J Clin Microbiol* 1993;31:767–770.

18. Centers for Disease Control-National Institutes of Health. *Biosafety in Microbiological and Biomedical Laboratories*, 3rd ed. HHS Publication (CDC) 93-8395, Washington D.C.: U.S. Government Printing Office, 1993.

19. Jost Jr KC, Dunbar DF, Barth SS, et al. Identification of *Mycobacterium tuberculosis* and *M. avium* complex directly from smear-positive sputum specimens and BACTEC 12B cultures by high performance-liquid chromatography with fluorescence detection and computer-driven pattern recognition models. *J Clin Microbiol* 1995;33:1270–1277.

20. Lang AM, Feris-Iglesias J, Pena C, et al. Clinical evaluation of the Gen-Probe amplified direct test for detection of *Mycobacterium tuberculosis* complex organisms in cerebrospinal fluid. *J Clin Microbiol* 1998;36:2191–2194.

21. Wang SX, Tay L. Evaluation of three nucleic acid amplification methods for direct detection of *Mycobacterium tuberculosis* complex in respiratory specimens. *J Clin Microbiol* 1999;37:1932–1934.

22. Stender H, Lund K, Petersen KH, et al. Fluorescence in situ hybridization assay using peptide nucleic acid probes for differentiation between tuberculous and nontuberculous mycobacterium species in smears of mycobacterium cultures. *J Clin Microbiol* 1999;37:2760–2765.

23. Ichiyama S, Ito Y, Sugiura F, et al. Diagnostic value of the strand displacement amplification method compared to those of Roche Amplicor PCR and culture for detecting mycobacteria in sputum samples. *J Clin Microbiol* 1997;35:3082–3085.

24. Ichiyama S, Iinuma Y, Yamori S, et al. *Mycobacterium* growth indicator tube testing in conjunction with either the AccuProbe or the AMPLICOR-PCR assay for detecting and identifying mycobacteria from sputum samples. *J Clin Microbiol* 1997;35:2022–2025.

25. Cantanzaro A, Davidson BL, Fujiwara PI, et al. Rapid diagnostic tests for tuberculosis: what is the appropriate use? In: *Proceedings of the American Thoracic Society Workshop*. *Am J Respir Crit Care Med* 1997;155:1804–1814.

26. Venugopal TV, Vengopal PV. Mycetoma. In: Braude AI, ed. *Medical Microbiology and Infectious Disease*. Philadelphia: WB Saunders, 1981:1762–1775.

27. Bothamley GH. Immunological procedures for diagnosis of tuberculosis and leprosy. In: Rose NR, Conway de Macario E, Folds JD, et al., eds. *Manual of Clinical Laboratory Immunology 5th ed*. Washington, DC: American Society for Microbiology, 1997:534–541.

28. Beaman BL. Nocardiae and agents of actinomycosis. In: Rose NR, Conway de Macario E, Folds JD, et al., eds. *Manual of Clinical Laboratory Immunology 5th ed*. Washington, DC: American Society for Microbiology, 1997:543–546.

29. Woods GL, Washington II JA. Mycobacteria other than *Mycobacterium tuberculosis*: review of microbiologic and clinical aspects. *Rev Infect Dis* 1987;9:275–294.

30. American Thoracic Society. Diagnosis and treatment of disease caused by nontuberculous mycobacteria; official statement. *Am J Respir Crit Care Med* 1997;156:S1–S25.

31. Cage GD, Saubolle MA. Infrequently isolated and unusual mycobacterial species as agents of human disease. *Rev Med Microbiol* 1997;8:125–135.

32. Falkinham III JO. Epidemiology of infection by nontuberculous mycobacteria. *Clin Microbiol Rev* 1996;9:177–215.

33. Hastings RC, Gillis TP, Krahenbuhl JL, et al. *Clin Microbiol Rev* 1988;1:330–348.

34. Piersimoni C, Scarparo C, Cicher P, et al. Multicenter evaluation of the MB-redox medium compared with radiometric BACTEC System, *Mycobacteria* Growth Indicator Tube (MGIT), and Lowenstein-Jensen medium for detection and recovery of acid-fast bacilli. *Diagn Microbiol Infect Dis* 1999;34:293–299.

35. Cambau E, Wichlacz C, Truffot-Pernot C, et al. Evaluation of the new MB Redox System for detection of growth of mycobacteria. *J Clin Microbiol* 1999;37:2013–2015.

36. Somoskovi A, Magyar P. Comparison of the *Mycobacteria* Growth Indicator Tube with MB Redox, Lowenstein-Jensen, and Middlebrook 7H11 Media for recovery of *Mycobacteria* in clinical specimens. *J Clin Microbiol* 1999;37:1366–1369.

37. Zanetti S, Ardito F, Sechi L, et al. Evaluation of nonradiometric system (BACTEC 9000 MB) for detection of mycobacteria in human clinical samples. *J Clin Microbiol* 1997;35:2072–2075.

38. Forbes BA. Critical assessment of gene amplification approaches on the diagnosis of tuberculosis. *Immunol Invest* 1997;26:105–116.

39. American Thoracic Society. Treatment of tuberculosis and tuberculosis infection in adults and children; official statement. *Am J Respir Crit Care Med* 1994;149:1359–1374.

40. Nolte FS. Antimicrobial susceptibility test procedures for mycobacteria. In: Wentworth BB, ed. *Diagnostic Procedures for Bacterial Infections*. 7th ed. Washington, DC: American Public Health Assocation, 1987;669–678.

41. Beaman BL, Beaman L. Nocardia species: host-parasite relationships. *Clin Microbiol Rev* 1994;7:213–264.

42. Mahgoub ES. Agents of mycetoma. In: Mandell GL, Bennett JE, Dolin R, eds. *Principles and Practice of Infectious Diseases*. 4th ed. New York: Churchill Livingstone, 1995:2327–2330.

43. Tight RR, Barlett MS. Actinomycetoma in the United States. *Rev Infect Dis* 1981;3:1139–1150.

44. Houang ET, Lovett IS, Thompson FD, et al. *Nocardia asteroides* infection—a transmissible disease. *J Hosp Infect* 1980;1:31–40.

45. Blumel J, Blumel E, Yassin AF, et al. Typing of *Nocardia farcinica* by pulse-field gel electrophoresis reveals an endemic strain as source of hospital infections. *J Clin Microbiol* 1998;36:118–22.

46. Wallace RJ, Steele LC. Susceptibility testing of *Nocardia* species for the clinical laboratory. *Diagn Microbiol Infect Dis* 1988;9:155–166.

47. Ambaye A, Kohner PC, Wollan PC, et al. Comparison of agar dilution, broth microdilution, disk diffusion, E-test, and BACTEC radiometric methods for antimicrobial susceptibility testing of clinical isolates of the *Nocardia asteroides* complex. *J Clin Microbiol* 1997;35:847–852.

48. Biehle JR, Cavalieri SJ, Saubolle MA, et al. Comparative evaluation of the E test for susceptibility testing of Nocardia species. *Diagn Microbiol Infect Dis* 1994;19:101–110.

49. Prescott JF. *Rhodococcus equi*: an animal and human pathogen. *Clin Microbiol Rev* 1991;4:20–34.

50. Lasker BA, Brown JM, McNeil MM. Identification and epidemiological typing of clinical and environmental isolates of the genus *Rhodococcus* with use of a digoxigenin-labeled rDNA gene probe. *Clin Infect Dis* 1992;15:223–233.

51. Van Etta LL, Felice GA, Ferguson RM, et al. *Corynebacterium equi*: a review of 12 cases of human infection. *Rev Infect Dis* 1983;5:1012–1018.

52. Ebersole LL, Paturzo JL. Endophthalmitis caused by *Rhodococcus equi* Prescott serotype 4. *J Clin Microbiol* 1988;26:1221–1222.

53. Hsueh PR, Hung CC, Teng LJ, et al. Report of invasive *Rhodococcus equi* infections in Taiwan, with an emphasis on the emergence of multidrug-resistant strains. *CID* 1998;27:370–375.

54. McNeil MM, Brown JM. Distribution and antimicrobial susceptibility of Rhodococcus equi from clinical specimens. *Eur J Epidemiol* 1992;8:437–443.

55. Schaal KP. Laboratory diagnosis of actinomycete diseases. In: Goodfellow M, Mordarski M, Williams ST, eds. *The Biology of the Actinomycetes*. London: Academic Press, 1984:441–456.

56. Rihs JD, McNeil MM, Brown JM, et al. *Oerskovia xanthineolytica* implicated in peritonitis associated with peritoneal dialysis: case report and review of *Oerskovia* infections in humans. *J Clin Microbiol* 1990;28:1934–1937.

55
Antimicrobial Susceptibility Testing

Carl L. Pierson

Antimicrobial susceptibility test results are regarded by most clinicians as vital information for the rational selection of antimicrobics to treat an established infection or to prevent an infection. Unfortunately, few clinicians who use susceptibility test results understand the methods used to generate them and, more importantly, the limitations of using such *in vitro* procedures to predict efficacy in their patients. With new antimicrobics being added almost yearly to the physician's armamentarium with their varied antimicrobial spectra, pharmacokinetics, and pharmacodynamics, the complexity of choosing and administering an appropriate therapeutic agent has increased dramatically. In addition, our expanding knowledge regarding the number and complexity of antimicrobic resistance mechanisms is requiring the testing laboratories to add or modify testing procedures to detect such mechanisms. Therefore, clinical pathologists and microbiologists must work closely with the infectious disease specialists and pharmacists within a medical care setting to provide meaningful reports with interpretations to assist in this selective process. Quality assurance indicators designed to monitor how clinicians use laboratory information often reveal where false assumptions are being made from a laboratory report, e.g., “the antibiotic with the lowest minimum inhibitory concentration (MIC) is the one to use.” Clinical pathologists and microbiologists must continue to stay abreast of the laboratory testing methods being used and their clinical relevance, and must interact with the clinicians using this information to produce useful reports leading to appropriate drug utilization.

The clinical relevance of *in vitro* susceptibility results is not altogether clear in many instances. Bacterial pathogens declared “susceptible” to an antimicrobic selected for therapy may be treated effectively in an immunocompetent patient, but the drug may be ineffective in an immunocompromised patient or if the infection occurs in a poorly vascularized anatomic site. Most studies that attempt to correlate treatment success with *in vitro* susceptibility fail to show a positive correlation; however, there frequently is a positive correlation between organisms that test resistant and treatment failure (1, 2, 3 and 4). Therefore, the apparent value of the susceptibility test is to monitor for organism resistance or declining susceptibility to selected antimicrobics.

The developmental history of susceptibility test methods has been reviewed (5). Much effort has gone into developing testing standards to provide consistent interlaboratory results. Drs. Hans Ericsson and John Sherris worked with an international team to publish the International Collaborative Study report in 1971 (6). These remarkable studies laid the groundwork upon which most subsequent work on test standardization for disk diffusion and broth dilution techniques was done. In the United States, this continuing task has been taken up by the National Committee for Clinical Laboratory Standards (NCCLS) subcommittee on antimicrobial susceptibility tests, which publishes testing standards and supplemental updates for these standards on a periodic basis. NCCLS Subcommittees have been formed to develop methods for performing other special tests such as the minimum bactericidal concentration (MBC) test and the serum bactericidal test (SBT), which, by consensus approval, could become approved standards.

Microbiology specimens submitted for routine bacteriologic culture usually arrive in the laboratory accompanied by a requisition for “culture and susceptibility.” This implies that a susceptibility test is to be done if a possible pathogen is isolated that is known to vary in its susceptibility to antimicrobics commonly used for treatment, and methods exist to perform and interpret the test. Organisms that typically show no variation in susceptibility to the recommended drugs of choice need not be tested, e.g., to date *Streptococcus pyogenes* has always tested susceptible to penicillin when testing is done appropriately. Other pathogens may need to be treated empirically without the benefit of susceptibility testing either because the *in vitro* results are known not to reflect the clinical response to therapy with a drug or because the laboratory has no way of interpreting the results obtained from such a test. An example of the former would be the activity of cefotetan against *Stenotrophomonas maltophilia*; an example of the latter would be the activity of ampicillin against *Bacillus cereus*.

TABLE 55.1. LABORATORY RESPONSIBILITIES FOR PERFORMING ANTIMICROBIAL SUSCEPTIBILITY TESTS

Select isolates to be tested
Select antimicrobics to test
Perform tests using standardized procedures whenever possible
Assist with interpretation of nonstandardized tests
Report results
Accurately
Timely
Selectively

As antimicrobics became available for clinical use, a method of determining which agents needed to be tested and/or reported against appropriate bacteria had to be developed. The current NCCLS documents M2-A7 (standards for the disk diffusion method) and M7-A5 (standards for dilution methods) offer suggested groupings that can be used to limit the number of drugs requiring testing based on drug class and probability of obtaining like results (7, 8). Suggestions also are offered for selective reporting of results to encourage the use of less expensive yet effective antimicrobics. If an organism shows resistance to a

first-tier antimicrobic, another one of a similar class that shows good activity can be reported from a secondary tier. Antimicrobics approved for use by the hospital formulary committee should be used as a guide for drug selection. Consultation with local infectious disease specialists who utilize the laboratory also can offer valuable assistance on drug selection and drug concentrations to test, but, to promote interlaboratory uniformity in testing, NCCLS standards should be followed unless there are overriding local circumstances to the contrary. Whenever an alternative method is used, it should be justified, and the clinician should be notified that a nonstandardized method was used. Under such circumstances, NCCLS interpretations should not be used, and it may be necessary to report the result without an interpretation. The areas of responsibility for the performance of susceptibility testing that rest with the microbiology laboratory are listed in Table 55.1.

COMMON IN VITRO ANTIMICROBIAL SUSCEPTIBILITY TESTS

Part of "55 - Antimicrobial Susceptibility Testing"

There are at least a dozen different laboratory test methods in use to determine organism susceptibility to various antimicrobics (Table 55.2 and Table 55.3) but only a few are used with any frequency in most clinical microbiology laboratories doing such testing on a routine basis. In addition, reference laboratories also may perform susceptibility tests for mycobacteria, fungi, and certain viruses.

TABLE 55.2. FREQUENTLY USED *IN VITRO* SUSCEPTIBILITY TESTS

Dilution tests
Broth microdilution
Agar
Semiautomated
Diffusion tests
Disk (aka Kirby Bauer)
Fixed gradient (Etest or epsilometer)
Spot tests (e.g., ß-lactamase)

To date, there are four susceptibility testing standards issued by the NCCLS:

M7-A5 *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically*, 4th ed. (2000);

M2-A7 *Performance Standards for Antimicrobial Disk Susceptibility Tests*, 6th ed. (2000);

M11-A4 *Methods for Antimicrobial Susceptibility Testing of Anaerobic Bacteria*, 4th ed. (1997); and

M27-A *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts* (1997).

In addition, a supplement — M100-S10 *Performance Standard for Antimicrobial Susceptibility Testing*; Ninth Informational Supplement (2000), provides updated tables for standards M7-A5 and M2-A7. These tables are updated annually.

NCCLS publications for other susceptibility testing procedures are available but have not been given standard status to date. Still, they contain valuable information regarding the state of the art. These publications include:

M21-A *Methodology for the Serum Bactericidal Test* (1999); Approved Guideline

M23-T3 *Development of In Vitro Susceptibility Testing Criteria and Quality Control Parameters* (1998); Tentative Guideline

M24-T *Antimycobacterial Susceptibility Testing* for Mycobacterium tuberculosis (1995); Tentative Guideline

M26-A *Methods for Determining Bactericidal Activity of Antimicrobial Agents* (1999); Approved Guideline, and

M38-P *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Conidium-Forming Filamentous Fungi* (1998); Proposed Standard.

Broth Microdilution Minimum Inhibitory Concentration Test

The broth microdilution minimum inhibitory concentration (MIC) test provides a quantitative measurement (in µg/mL) of the lowest concentration of an antimicrobial agent that inhibits the growth of a bacterium. The reader is advised to consult the NCCLS standard M7-A5 for a detailed description of procedures for this test method (8).

Selected antimicrobics can be obtained either from the pharmaceutical manufacturer, the U.S. Pharmacopeia, Rockville, MD, or a commercial distributor. Only standardized powders or liquids specifying specific activity and expiration date are to be used. The agents are dissolved in appropriate solvents, finally diluted to appropriate concentrations and distributed into the wells of molded microtiter plates (Fig. 55.1). The filled plates can be either frozen at -70°C or desiccated and stored at room temperature. The concentrations selected for each agent are determined by the achievable blood (or urine) concentrations obtained with recommended dosage schedules and routes of administration. Traditional MIC panels contain a two-fold dilutional series of concentrations for

each agent, ranging from easily achievable concentrations in blood (or urine)(susceptible) to those that usually are unachievable (resistant). Because of the large number of antimicrobics available for testing, some panels are designed as “breakpoint panels” that limit the number of dilutions used to the minimum necessary to provide interpretive information. This allows for more drugs to be tested per panel. Other drugs are present in single concentrations and may be used primarily for screening for one-point levels of resistance.

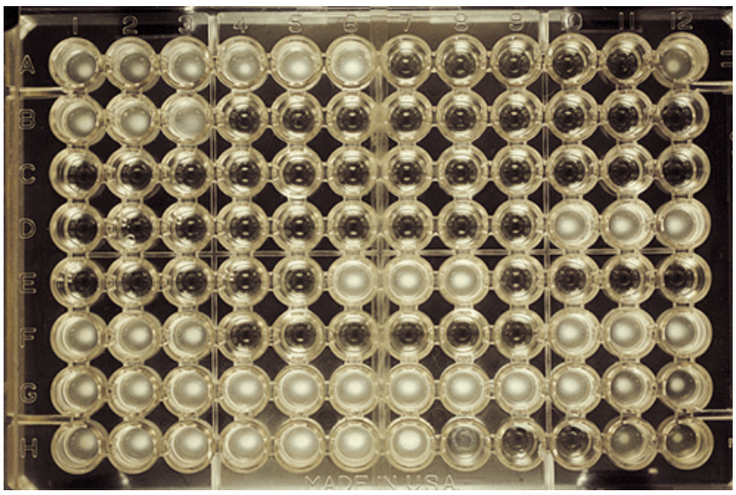


FIGURE 55.1. Microdilution minimum inhibitory concentration (MIC) tray showing growth endpoints using indirect lighting.

Frozen panels must be defrosted and subsequently inoculated with the test organism, making sure that the inoculum volume does not exceed more than 10% of the final volume in each well. Usually 1 to 5 µL of inoculum is added to 100 µL of diluted antimicrobic solution in each well. Alternatively, the drug concentration in each well can be prepared and stored at twice the final desired concentration and diluted two-fold with the diluted inoculum. Dessicated panels usually are rehydrated and inoculated simultaneously by suspending the appropriate number of organisms directly in the rehydrating growth medium.

TABLE 55.3. INFREQUENTLY USED *IN VITRO* SUSCEPTIBILITY TESTS

Minimum bactericidal concentration (MBC)
Serum bactericidal test (SBT, “Schlichter test”)
Time-kill kinetic assay
Drug synergy tests
Disk approximation
“Checkerboard”
Time-kill kinetic assay

Several different growth media have been used to assess antimicrobic activity. Some nutritionally fastidious bacteria (e.g., *Haemophilus influenzae*) require the addition of supplementary growth factors (9). Others may require higher concentrations of cations to express resistance characteristics not seen otherwise (10, 11 and 12). Serum or serum components may be added to show the effects of protein-binding on antimicrobic activity (13, 14 and 15). For routine testing, however, the growth medium selected for use in MIC testing in the United States is Mueller-Hinton broth (MHB) that contains a final concentration of 20 to 25 mg/L of Ca^{2+} and 10 to 12.5 mg/L of Mg^{2+} described as cation-adjusted Mueller-Hinton broth (CAMHB). These cation concentrations are especially important when testing the activity of certain aminoglycosides against *Pseudomonas aeruginosa* (11). To optimize the expression of resistance to oxacillin in staphylococci, the CAMHB also must contain 2% NaCl (12, 16). The pH of the medium must fall within the range of 7.2 to 7.4 (17).

Bacteria to be used for inoculation can be harvested from either solid media or broth culture. The organism's phase of growth appears to have little effect on the results, but it is important to have a relatively fresh culture to assure that a high percentage of the bacterial cells are viable (e.g., ≤24-hour culture). Initially, a portion of four to five similar colonies from a pure culture or from colonies well isolated on a plate are to be harvested and suspended in water and the turbidity adjusted to match that of a 0.5 McFarland turbidity standard or an optical density (O.D.) of approximately 0.08 @625 nm. This provides an approximate starting concentration of 10^8 viable cells/mL. This suspension must be further diluted in CAMHB to achieve a concentration such that the final concentration in the microdilution well will be about 5×10^5 CFU/mL. Although this concentration is somewhat arbitrary, it has been determined that when inoculum concentrations substantially higher than this are used, e.g., 10^7 CFU/mL, false-resistant results can be obtained; lower inoculum concentrations can yield false-susceptible results (18). A sample from the growth control well of an inoculated MIC panel should be removed and a colony count performed to document that the viable inoculum concentration is within the acceptable range. Inoculated panels are to be incubated 18 to 24 hours at 35°C in a room-air incubator. CO_2 incubators are not to be used unless necessary to achieve growth because the CO_2 increases the acidity of the growth medium, which can significantly alter the activity of certain antimicrobics, e.g., erythromycin. A full 24-hour incubation time is recommended when testing *Staphylococcus aureus* for oxacillin resistance. These resistant organisms are frequently slow-growing, and “heteroresistant” strains, i.e., when only a small number of cells express resistance, they may be difficult to detect with shorter incubation times (12). It may be

necessary to incubate coagulase-negative staphylococci up to 48 hours to detect oxacillin resistance. A full 24-hour incubation time also is necessary to optimally detect vancomycin resistance in enterococci. Rapid-growing mycobacteria and *Norcardia* spp. also may be tested for susceptibility using this method, but they usually require 3 to 4 days of incubation prior to reading (19, 20).

Panels are read by holding them over a light source that provides indirect background lighting. This type of illumination is optimal for detecting faint buttons of growth at the base of the wells. All panels are to have a growth control well that contains no antimicrobic, and an uninoculated sterility well that contains no antimicrobic. Both control wells should be examined prior to assessing the rest of the panel. The sterility well should show no evidence of growth, and the growth control well should show obvious growth that is consistent with the pattern of growth expected for the organism type inoculated. Gram-positive cocci usually form a discrete button at the bottom of the well, whereas Gram-negative bacilli frequently grow in a more diffuse pattern, although neither of these characteristics is absolute. If the control wells do not appear as expected, the reading of the panel should be aborted and the isolate retested.

If the control wells are acceptable, each series of dilutions of an antimicrobic should be examined for an endpoint where the lowest concentration in the antimicrobic series has completely inhibited the growth of the inoculated organism. This concentration is recorded for each antimicrobic tested. Most endpoints are quite easy to determine; some, especially certain bacteriostatic drugs, may not produce a sharp endpoint, and exhibit a "trailing effect." With experience, an endpoint usually can be determined where there is a sharp (≥80%) decrease in the amount of growth seen.

Specific control strains must be tested at least weekly to monitor drug concentrations in the panels and to assure that the technique being used to set up the panels is adequate to obtain accurate results. Quality control (QC) guidelines are available from the NCCLS (M23-T3). Currently, there are four strains that are recommended for routine quality control of MIC panels:

S. aureus, American Type Culture Collection (ATCC) 29213 (ß-lactamase positive);

Enterococcus faecalis, ATCC 29212;

Escherichia coli, ATCC 25922; and

P. aeruginosa, ATCC 27853.

Additional strains should be tested when certain drugs or specific bacteria are to be tested:

E. coli, ATCC 35218 when ß-lactam/ß-lactam inhibitor combinations are being tested (ß-lactamase positive);

H. influenzae, ATCC 49247 or 49766 when testing an *H. influenzae* isolate; and

Neisseria gonorrhoeae, ATCC 49226 when testing an isolate of *N. gonorrhoeae*.

All of the above strains are available from the ATCC, Rockville, MD. The ideal control strain would yield an MIC that is in the intermediately susceptible range for the drug being monitored. Unfortunately, many of the new broad-spectrum antimicrobics are highly active against the older control strains and produce MICs that are much lower than the lowest concentration being tested. New, stable, resistant control strains are needed for these drugs to provide adequate QC.

The MIC results are to be reported in µg/mL and an interpretation given for each drug as to whether the MIC falls within the range considered by the NCCLS standard to be susceptible, intermediately susceptible, or resistant. Infections caused by organisms that yield intermediately susceptible results still may be treated effectively by the drug if higher dosages can be administered or if the drug is known to concentrate within the infected locus such as in urine or bile. When organism-drug combinations are tested and no interpretations are available, no interpretations should be given until there is enough clinical evidence to show that an interpretation is justified. As mentioned previously, attempts should be made to limit the number of drugs reported to those that offer a more limited yet good spectrum of activity and are less expensive to purchase and administer.

Agar Dilution Method

The reader is referred to the section on the testing of anaerobes for a description of this method.

Disk Diffusion Method

This method, commonly referred to as the Kirby-Bauer test, provides a qualitative measure of the ability of an antimicrobic to inhibit the growth of a rapidly growing bacterium. The disk diffusion method was the first standardized susceptibility method available and remains a very useful method despite the continuing shift of laboratories to either the microdilution MIC method or a semiautomated procedure.

Disks containing a given concentration of an antimicrobic are placed on a confluently inoculated agar plate and incubated for 16 to 24 hours. At the end of the incubation period, zones of growth inhibition are measured across the disk diameter and recorded to the nearest millimeter (Fig. 55.2). The results are interpreted as susceptible, intermediately susceptible, or resistant by using the interpretive tables provided in the NCCLS standard M2-A5 (7). No quantitative results can be given with this method.

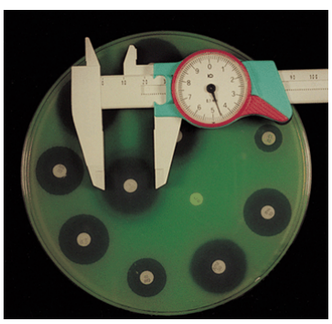


FIGURE 55.2. Measuring zones of inhibition around antimicrobic disks using a micrometer.

For routine disk diffusion testing, most clinical laboratories use plastic culture dishes with diameters of 150 mm that have been filled with Mueller-Hinton agar to a depth of 4 mm. Cation supplementation usually is unnecessary because the agar itself contributes adequate amounts of Ca^{2+} and Mg^{2+}, and the companies providing the media now monitor their products and supplement them when necessary. The media can be supplemented with other products such as horse blood, which supplies thymidine phosphorylase and allows for more accurate measurement of trimethoprim and sulfamethoxazole activity, which are inhibited by the presence of thymidine (21). Blood-supplemented plates often are necessary to obtain adequate growth of some fastidious organisms such as some streptococci and diphtheroids (9). To test *N. gonorrhoeae*, the NCCLS standards recommend the use of GC agar instead of Mueller-Hinton agar. The 150 mm plates can accommodate about 12 individual susceptibility disks; the actual number depends on the anticipated

size of the zones of inhibition, which should not overlap. This size is partly dependent on the concentration of the drug in the disk and the diffusion characteristics of the drug through the agar. Drugs such as penicillin can produce very large zones of inhibition when testing very susceptible organisms, whereas others such as most aminoglycosides produce relatively small zones but can still be considered clinically effective. The agar plates can be maintained at 4°C for up to the given expiration date. Regardless of the date, however, the agar surface should be smooth, level, and of the correct depth (4 mm) prior to placing a plate into service. No condensation should be apparent on the plate surface after the plate has been allowed to reach room temperature.

Inoculum preparation is similar to that previously described for MIC preparation. The tops of four to five well-isolated uniform colonies are touched with an inoculating loop and used to inoculate 4 to 5 mL of a nutrient broth (e.g., TSB or BHI). The broth is incubated at 35°C to 37°C for sufficient time to allow the growth to reach a tubidity equal to or greater than a 0.5 McFarland turbidity standard (ca. 1-2 × 10^8 cfu/mL) and subsequently diluted if necessary with broth to be equivalent to this standard. An acceptable alternative method is to remove sufficient fresh growth from a nonselective agar medium to obtain adequate tubidity. A nontoxic swab is immersed into the broth, swirled to eliminate air bubbles, and the excess fluid pushed out of the swab by rotating it against the wall of the tube above the fluid level. The swab is used to streak the agar plate for confluent growth. Care must be taken not to over- or underinoculate the plate because this can alter the resulting zone of inhibition (22). A correctly inoculated plate will reveal individual but closely packed colonies with careful inspection of the plate after incubation.

Filter paper disks containing designated amounts of an antimicrobic can be obtained from either the manufacturer or several commercial supply firms. They usually are supplied in sealed cartridges with a desiccating agent to assure dryness. Moisture can decrease potency. It is necessary to put the disks in a non–self-defrosting freezer for long-term storage and keep a working supply in a refrigerator under desiccation. The disks should be allowed to warm to room temperature prior to opening them for use.

The disks must be placed onto the surface of the inoculated agar plate no later than 15 minutes after inoculation. Once applied, they cannot be repositioned because the antimicrobic begins to diffuse immediately. The plates are to be incubated in a non-CO_2 incubator at 35°C for 16 to 18 hours unless otherwise specified. After this time, the zone of inhibition can be measured to the nearest mm and the results interpreted following the appropriate NCCLS Standard table. Generally, supplements should be added only when necessary to obtain adequate growth. Certain fastidious organisms such as *H. influenzae, N. gonorrhoeae*, and *S. pneumoniae* will require blood supplements and a 5% to 7% CO_2 atmosphere to grow. Whenever media supplements are used, appropriate control strains must be tested to assure proper test function and interpretation of results.

The determination of disk diffusion interpretation is based on correlating MIC results with corresponding zone diameters. These are inversely proportional; i.e., as the MIC increases, the zone diameter of inhibition decreases. When a variety of organisms with varying susceptibilities to an antimicrobic are tested by both methods and the results plotted with the increasing MIC results on the *y* axis and the increasing zone sizes in millimeters plotted on the x axis, a regression curve with a negative slope is obtained (Fig. 55.3). Depending in large part on the pharmacokinetics of the drug, interpretive zone sizes can be determined that correspond to those used for interpreting MIC results. For various reasons, this correlation does not always exist and discrepant results can occur. Nevertheless, commercial products have been developed that use such regression analysis to calculate an MIC from zone size data. An example is the BioMIC (Giles Scientific, Inc., New York, NY), which uses an imaging system and computer to display the plate image on a computer monitor and marks the zone of inhibition around each disk (Fig. 55.4). This zone can be adjusted, if necessary, by the technologist prior to entering the data into the system. The computer then records the zone size and calculates the MIC based on stored regression data. Studies have been published correlating the calculated MIC with the standardized microdilution methods (23).

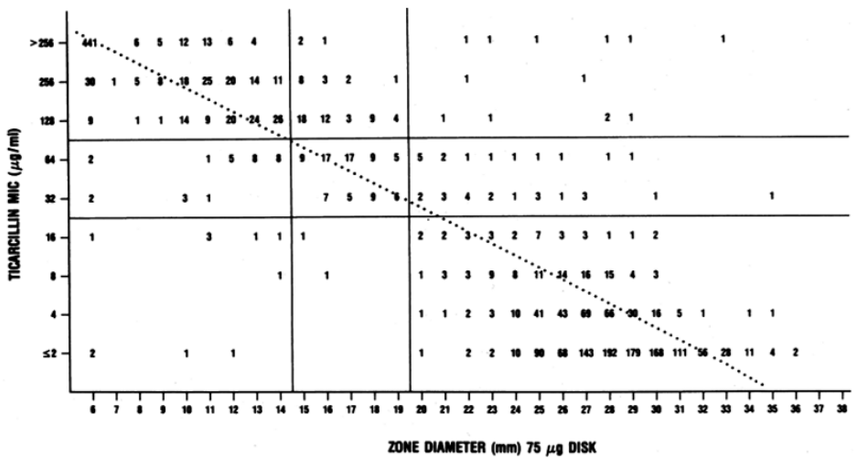


FIGURE 55.3. Regression analysis of *Enterobacteriaceae* susceptibility to ticarcillin. Broth microdilution minimum inhibitory concentrations (MICs) *(y-axis)* are plotted against zones of inhibition *(x-axis)* obtained by disk diffusion. Horizontal and vertical lines represent current breakpoints for interpretation of test results. (Reprinted with permission from Barry A, Fuchs P, Gerlach E, et al. Antimicrobial. *Agents Chemother*. 1992;36:139.)

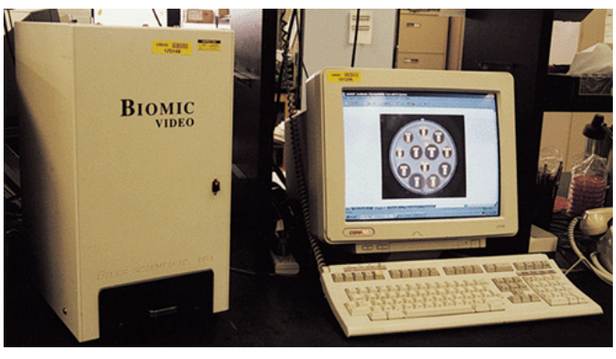


FIGURE 55.4. The BioMIC System showing on-screen zone measurement.

A new test method has been developed recently called the E Test (AB Biodisk, North America, Inc., Piscataway, NJ), which is a modification of the disk diffusion test but provides an MIC result (24, 25 and 26). This test is set up in the same way as the disk diffusion test except that, in place of the disks, one uses 5-mm × 50-mm plastic strips having a continuous, fixed gradient of antimicrobic immobilized on one side and an MIC interpretive scale corresponding to 15 two-fold MIC dilutions on the other side. After incubation, the zone of inhibition takes the shape of an ellipse. The MIC is read at the point where the zone intersects the MIC scale on the strip (Fig. 55.5). Studies show this method to give greater than 95% agreement with the standardized broth microdilution method (27).

Again, most testing laboratories are opting to use methods that provide an MIC result rather than the disk-diffusion test. It should be noted that both methods provide clinically useful information. Under most clinical situations, the interpretive result is all that is needed for the choice of antimicrobic (28, 29). There also are advantages to the use of each method. With the disk diffusion method, it is easy to substitute one disk for another. Unless you prepare your own MIC trays, you are dependent on a commercial provider for the drug profiles available. It also is easier to spot contamination and low-level resistance than it is with the broth microdilution method. However, the disk-diffusion method is approved for use only with rapidly growing organisms, whereas the broth-microdilution method can be incubated for longer periods of time and, therefore, can be used to test slower growing nonfermenters, anaerobes, and even the rapid-growing mycobacteria (27, 30, 31). It also lends itself more readily to automation with prepared trays having long outdates. MBCs cannot be done using agar diffusion techniques.

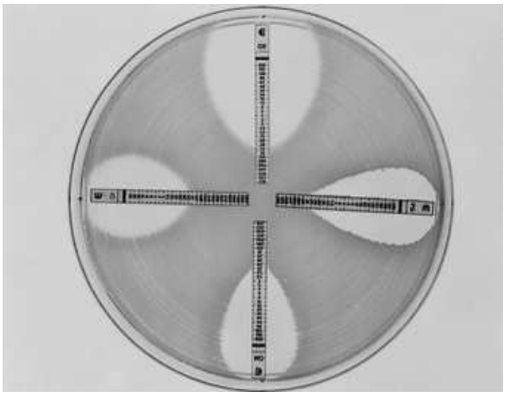


FIGURE 55.5. The E-test. Elliptical zones of inhibition around antimicrobic-impregnated strips. The minimum inhibitory concentration (MIC) is read from the on-strip scale where the zone touches the side of the strip.

SUSCEPTIBILITY TESTING OF ANAEROBIC BACTERIA

Part of "55 - Antimicrobial Susceptibility Testing"

Although testing standards exist for anaerobes (NCCLS M11-A4), it is now recommended that most clinical laboratories consider limiting the frequency of such testing (32). Empiric choice of agents shown by clinical trials to be effective against pathogenic anaerobes is usually sufficient. Susceptibility testing of anaerobes is indicated primarily in situations where the infection did not respond to recommended therapy or for periodic screening of local isolates to assess the relative activity of recommended antimicrobics (33, 34).

Various methods have been used, but none appear totally adequate for the testing of all pathogenic anaerobes. An agar dilution method has been chosen as the reference method, but it is not a convenient method for most clinical laboratories to use. The reader is referred to the M11-A4 standard for details on this method. In brief, antimicrobics are made up from standard powders, as mentioned earlier, and added to melted Wilkins-Chalgren agar supplemented with hemin and vitamin K_1 to achieve the desired concentration. The molten agar is immediately poured into Petri plates and allowed to solidify and the surface to dry after reaching room temperature before inoculation. Wadsworth Brucella blood agar using 5% laked sheep blood, has been suggested as an alternative medium and appears to support the growth of a broader range of anaerobes. MICs may average about one dilution higher with this medium. Thioglycollate broth enriched with hemin and vitamin K_1 is inoculated with a fresh isolate taken from an anaerobic blood agar plate and incubated to obtain turbid growth. The growth is adjusted to obtain a concentration such that 10^5 CFU are deposited onto the surface of the agar in the form of a spot.

Several organisms can be tested on each plate containing a specific antimicrobic concentration. The concentrations recommended for testing for each agent are given in the standard but should span the physiologically achievable range with maximum dosage. The aforementioned procedures can be done under aerobic conditions, but following inoculation, the plates must be incubated at 35°C to 37°C under anaerobic conditions for 48 hours before determining endpoints. The endpoint is the lowest concentration tested that inhibits growth of the isolate. Some inoculum spots may show a haze or one to two colonies of reduced size compared to the growth control, but should be ignored. If trailing occurs, the endpoint is read at the concentration where a major decrease in colony count or size occurs. The result is interpreted as either susceptible or resistant. No intermediate or moderately susceptible catagories exist, because the maximum dosage is recommended for the treatment of anaerobic infections.

Most laboratories testing anaerobes use a broth microdilution method. The methodology is quite similar to that described for testing organisms that grow aerobically; however, some significant differences exist. Various test media have been used including Schaedler's, West-Wilkins, brain heart infusion, and Wilkins-Chalgren broth. None are capable of supporting the growth of all anaerobes that one may want to test. Any of these can be used so long as the control strains respond appropriately. The recommended control strains are:

Bacteroides fragilis, ATCC 25285;

Bacteroides thetaiotaomicron, ATCC 29741;

Clostridium perfringens, ATCC 13124; and

Eubacterium lentum, ATCC 430555.

It is recommended that at least two of these control strains be run with each batch of tests or at least weekly. The inoculum is prepared as described for the standard agar dilution, but the final concentration should be about 10^6 CFU/mL in each dilution well. The inoculum volume should not exceed 0.01 mL, and the final volume in each well should not be less than 0.1 mL. The incubation time should be 48 hours under anaerobic conditions. The endpoint should be the lowest concentration of each drug dilutional series that shows either no growth or significantly reduced growth compared to the growth control well. It has been noted that the MICs obtained by broth microdilution usually run about one dilution lower than those obtained by agar dilution.

Older studies refer to a disk elution technique where disks containing a given concentration of an antimicrobic were added to a given volume of test medium and time was allowed for the antimicrobic to elute into the medium (35). Usually, only one concentration was tested, and the isolate was considered either sensitive or resistant depending on whether or not it was inhibited. This technique is no longer recommended by the NCCLS, especially when testing cephalosporins.

The ß-Lactamase Test

The ß-lactamase test is a rapid, sensitive procedure for screening isolates for resistance to ß-lactamase-susceptible antimicrobics. The genetic ability of bacteria to produce small amounts of ß-lactamase is universal as it is needed for normal cell wall development. However, certain bacteria are capable of producing sufficient quantities of these enzymes to effectively neutralize the killing capacity of certain ß-lactam antimicrobics before they can

reach vulnerable “penicillin-binding proteins” that serve as their primary targets. The genetic information coding for these enzymes can be carried on plasmids or integrated into the chromosome. The reader is referred to Bush et al. (36) for a comprehensive review of the ß-lactamase and their current classification. Some ß-lactamases are produced constitutively and others must be induced to produce detectable amounts (37, 38). Irrespective of the method of production, if a ß-lactamase is detected, all susceptible ß-lactam antimicrobics are considered to be ineffective against the organism. This type of resistance spawned the development of many new ß-lactam drugs that either are inherently resistant to ß-lactamase inactivation or are coupled with specific enzyme inhibitors that irreversibly bind with the enzyme to render them inactive.

Three methods have been used in the clinical laboratory for the detection of ß-lactamases: acidometric, chromogenic cephalosporin, and iodometric.

The acidometric method takes advantage of the enzyme's hydrolytic action on the ß-lactam ring, producing penicilloic acid, which has an additional carboxyl group. This causes the pH of the medium to decrease sufficiently to cause the pH indicator to change color. A convenient technique is to draw a solution containing penicillin and a pH indicator into a capillary tube and subsequently to stab one or more colonies with the end of the tube creating an organism plug that makes contact with the solution. If ß-lactamase is produced, the solution adjacent to the plug will begin to become acidic. This method is easy and quite inexpensive but may not react with certain enzymes that are more specifically active against cephalosporins than against penicillins.

The chromogenic cephalosporin method currently is considered the optimal method because it appears to react with all classes of ß-lactamases known to date. This method takes advantage of the chromogenic properties of specific cephalosporins that change color when hydrolyzed. A disk containing a chromogenic cephalosporin, e.g., nitrocefin or cefesone, is wetted with water or saline and the test organism is removed from an agar plate and streaked onto the disk surface. If an isolate is a ß-lactamase producer, it will cause a color change, usually within minutes but it may take as long as 1 hour for some staphylocci. Impregnated disks or cartridges are available from commercial sources (39, 40).

The iodometric method is seldom used in clinical laboratories but can be a very sensitive method for detecting penicillinases. The method relies on the ability of penicilloic acid to function as a reducing agent, converting iodine to iodide. When a solution containing penicillin, iodine and starch is inoculated with a penicillinase-producing bacterium, the iodine is reduced to iodide and the starch solution changes from a blue-black color to colorless.

Organisms that are tested routinely for ß-lactamase production are *H. influenzae* and anaerobes such as nonfragilis *Bacteroides* spp., *Fusobacterium*, and *Clostridium* spp. The vast majority of staphylococci produce a penicillinase that is readily detectable following induction, but the incidence is so high that few laboratories bother testing for the few isolates that are nonproducers. *Moraxella catarrhalis*, *N. gonorrhoeae*, and *Enterococcus* spp. are tested infrequently by most laboratories. *M. catarrhalis* usually is positive and both *N. gonorrhoeae* and *Enterococcus* spp. are rarely positive. Certain Gram-negative facultative anaerobes also produce ß-lactamases that render specific antimicrobials inactive. Such activity results in high MICs, and it usually is not necessary to specifically test such isolates for ß-lactamase production.

Certain Gram-negative bacilli such as *Enterobacter cloacae* may only produce a ß-lactamase following induction with a specific ß-lactam antimicrobic such as cefoxitin or imipenem. The clinical significance of this is uncertain at present; however, certain strains may convert to constitutive enzyme production, which is clinically significant (37, 41, 42).

Inducible strains can be detected using a modification of the disk diffusion test. At the time of inoculation, a disk containing a low concentration of an inducing antimicrobic is placed just outside the zone of inhibition normally obtained with an enzyme-susceptible antimicrobic. After incubation, the shape of the zone of inhibition is noted. If no enzyme is induced, the shape of the zone will remain circular; however, if induced ß-lactamase is produced in the presence of the inducing agent, the zone will become flattened at the side adjacent to the inducing disk (Fig. 55.6). Such a test usually is necessary only when an appropriate organism appears to be susceptible *in vitro* but resistant *in vivo* and a potentially inducing antimicrobic is being administered in addition to the ß-lactamase-susceptible drug.

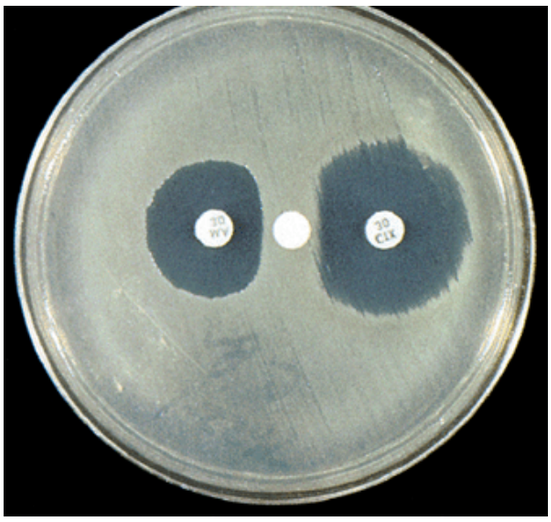


FIGURE 55.6. Evidence of the induction of ß-lactamase by the disk-diffusion technique. The zones of inhibition are flattened adjacent to the central disk containing an inducing agent.

RESISTANCE SCREENING PLATES

Part of "55 - Antimicrobial Susceptibility Testing"

Agar plates containing a specified concentration of a single antimicrobial agent are frequently used to screen selected organisms for resistance to the agent.

S. aureus isolates can be screened for resistance to oxacillin using a CAMHA plate containing 4% NaCl and 6 µg/mL of oxacillin. Each isolate's concentration is adjusted to match a 0.5 McFarland standard and spotted onto the agar plate and incubated at 35°C for a full 24 hours prior to examination for growth. All isolates testing resistant are to be reported as resistant

to all other ß-lactam antimicrobics. This method is not approved for screening coagulase-negative staphylococci (43).

E. faecium and *E. fecalis* can be screened for vancomycin resistance in a similar way by spotting the prepared inoculum onto a brain-heart infusion agar plate containing 6 µg/mL of vancomycin (44, 45). Isolates that screen resistant should be confirmed by performing a standard microdilution MIC test. This same plate can be used to screen *S. aureus* isolates for increased resistance to vancomycin (46, 47). *S. pneumoniae* isolates can be screened for resistance to penicillin by confluent inoculation of a standard 5% sheep blood MHA plate and application of a 1 µg oxacillin disk followed by 20- to 24-hour incubation in a 5% CO_2 atmosphere (48). Isolates yielding a inhibitory zone size of >20 mm are considered susceptible to penicillin; those yielding zones of <19 mm must be further tested by microdilution MIC or Etest to confirm resistance.

DETECTION OF EXTENDED-SPECTRUM ß-LACTAMASE-PRODUCING ISOLATES

Part of "55 - Antimicrobial Susceptibility Testing"

In the mid-1980s, reports began appearing in the literature of certain enteric Gram-negative bacilli, primarily *E. coli* and *Klebsiella* spp., demonstrating resistance to various extended-spectrum cephalosporins (49). Further investigations revealed that these organisms carried mutated TEM and SHV plasmids that coded for ß-lactamases capable of hydrolyzing extended-spectrum cephalosporins but were inhibited by the ß-lactamase inhibitors such as clavulanic acid and remained susceptible to cephamycins, carbapenems, and cephem class drugs. These enzymes are called extended-spectrum ß-lactamases or ESBLs. Unfortunately, organisms carryings ESBLs may test susceptible *in vitro* to one or more of the third generation cephalosporins but usually the zone of inhibition is smaller and the MIC is somewhat higher than is typically seen for truly susceptible organisms (50, 51 and 52). ESBLs also have been found in other members of the *Enterobacteriaceae* (53, 54). Various methods have been used to detect isolates expressing ESBL resistance but recently the NCCLS has established a standardized method for this purpose (7). Most testing laboratories monitor one or more of three ß-lactam antimicrobics for reduced activity against isolates of *E. coli* and *Klebsiella*: ceftazidime, aztreonam, and/or cefpodoxime. If reduced activity is detected by either disk-diffusion or MIC, a confirmation test is to be performed that involves testing the isolate's relative susceptibility to both ceftazidime and cefotaxime with and without the addition of clavulanic acid. If either the zone of inhibition increases by >5 mm or the MIC decreases <3 two-fold dilutions when clavulanic acid is added, the isolate is declared an ESBL-producing strain and it is reported as resistant to all cephalosporins and monobactams (Fig. 55.7). Isolates testing susceptible to cephamycins, carbapenems, cephems, and ß-lactam/inhibitor combinations should be reported as susceptible. AB Biodisk (Solna, Sweden) has produced a dual gradient strip with either cefotaxime or ceftazidime at one end of the strip and the same drug with a fixed concentration of clavulanic acid at the opposite end (26). When placed on an inoculated agar, an elliptical zone of inhibition could be produced at each end. If an MIC ratio of drug:drug + clavulanic acid is >16:1, the isolate is declared and ESBL-producer (Fig. 55.8).

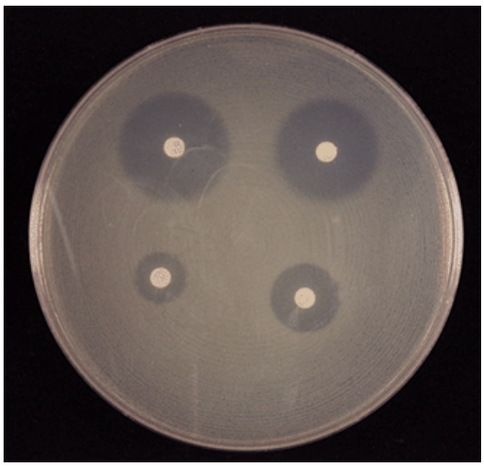


FIGURE 55.7. Extended-spectrum ß-lactamase disk test. Zones of inhibition with and without the addition of clavulanic acid to standard cefotaxime and ceftazidime disks.

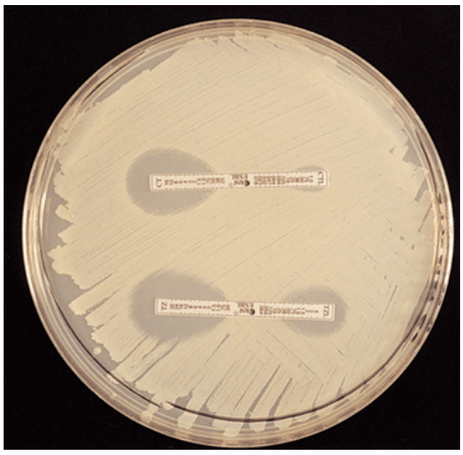


FIGURE 55.8. Etest ESBL test strips. A concentration gradient of either cefotaxime or ceftazidime is on the left side of each strip and the same drug plus a fixed concentration of clavulanic acid is on the right side of each strip.

OTHER TESTS TO DETECT DRUG-INACTIVATING ENZYMES

Part of "55 - Antimicrobial Susceptibility Testing"

Relatively rare isolates of *H. influenzae* produce choramphenicol acetyltransferase, which renders them resistant to chloramphenicol (55). A 10-minute disk or tube test is available to detect this enzyme. Most laboratories do not perform this test unless a strain

of *H. influenzae* capable of producing this enzyme has been detected in the community or there is a treatment failure.

Aminoglycoside-modifying enzymes can be detected by testing for enzyme-substrate activity or, more recently, by specific DNA probes but, to date, these tests are performed only in research settings (56, 57).

Semiautomated Susceptibility Testing

Several instruments have been introduced that facilitate susceptibility testing; a few approach full automation and usually perform organism identification procedures as well (58, 59). Two systems have been used extensively in clinical microbiology laboratories: the Vitek System (AMS) (bioMerieux Vitek, Inc., Hazelwood, MO)(60, 61) and the MicroScan System (Dade MicroScan, Inc., West Sacramento, CA)(62). Briefly, each system is composed of three major components: a module that serves as an incubator/reader; a computer with a keyboard that allows data/command entry and controls all system functions; and a printer to provide hard copy. Prepared susceptibility test panels, usually provided in the dried form for longer shelf life, are packaged to be run only on the designated system. Most companies provide several different formats to choose from, or they can customize panels at extra cost. The test isolate is prepared by diluting the suspension in an inoculating medium to obtain a designated turbidity. This inoculum is used to simultaneously hydrate and inoculate the panels. This is done either by hand-held disposable transfer devices, mechanical inoculators or a vacuum unit. Once inoculated, the panels are appropriately labeled and accessioned in the computer for subsequent tracking. The panels are then loaded into the incubator/reader. The panels are “read” by the instrument at periodic intervals using light emitting diodes or fiberoptics as light sources, coupled with detectors to monitor changes in turbidity in each well. Incubation times vary, depending on the system, from 3.5 hours using fluorogenic substrates to 18 hours with conventional substrates. As each test is finalized, the position becomes available for another panel. Systems in clinical use have been evaluated for accuracy by comparing results with the standardized reference methods (63, 64).

Testing errors have been categorized as very major, major, or minor. A very major error occurs when the test system calls an isolate susceptible to an antimicrobic, but the reference system calls it resistant (false-susceptible). A major error occurs when the test system calls an isolate resistant, and the reference method calls it susceptible (false-resistant). Minor errors result when the interpretation varies between susceptible and intermediately susceptible or between resistant and intermediately susceptible relative to the reference method. The FDA requires the manufacturer to recommend to the user of a system that an alternative method should be used whenever a drug/bug combination produces a very major error more than 1.5% of the time or a major error more than 3% of the time. Through software modifications, most approved systems now produce acceptable and reproducible results, but some problems still occur, especially with slow-growing and fastidious bacteria, oxacillin-heteroresistant staphylococci, and gentamicin-resistant enterococci (65). Systems that report results within 5 to 6 hours may have difficulty detecting resistance in isolates that produce inducible ß-lactamase. Such systems also have difficulty detecting low-level vancomycin resistance in enterococci mediated by the *vanB* gene (66).

Sites using such systems report improved reproducibility of results and a modest savings in personnel time, although usually not enough to free up a technologist for other duties. Two-way interfaces help to limit the amount of data input necessary during accessioning. The system's computer is able to generate useful summary reports and susceptibility trend analyses. Programs can be added allowing for input of pharmacy data and compares the susceptibility of the patient isolate to the antimicrobics being administered and flags instances when the organism appears resistant to the antimicrobic(s) being given.

To date, these systems are not fully automated. The inoculum must still be prepared by hand and the panels inserted into the system. No system has been free of major error; therefore, each testing site must maintain a backup system or use a reference laboratory to obtain results for specified bug/drug combinations. Some panels cannot be read manually in the event of a system malfunction. The primary motive for going to an automated system is to increase accuracy, efficiency, and to realize a cost savings. To date, these systems have had only a modest favorable impact. The manufacturers of these systems are actively addressing these concerns, and substantial improvements in performance have been made since the first introduction of these systems into clinical laboratories.

INFREQUENTLY USED ANTIMICROBIC SUSCEPTIBILITY TEST METHODS

Part of "55 - Antimicrobial Susceptibility Testing"

Because of the complexity of performing and interpreting the following special susceptibility tests, it is advisable that such requests be routed through a laboratory director accompanied by a pathology consultation. The director should ascertain why the test is being ordered and determine if the test being requested is appropriate or if the information needed can be better obtained using a different, less expensive methodology.

The Minimum Bactericidal Concentration Test

Under special clinical circumstances, it may be warranted to determine the minimum antimicrobic concentration necessary to kill the organism. This information usually is unnecessary if the infected patient is immunocompetent. In that case, both bacteriostatic and bactericidal drugs appear to be effective therapeutic agents. However, patients who are leukopenic, have defects in cell-killing capacity or have endocarditis may benefit from receiving treatment with a bactericidal agent (67, 68 and 69). Certain organisms may be inhibited by low concentrations of an antimicrobic but require much higher concentrations to be killed. This can be determined by the minimum bactericidal concentration (MBC). Bactericidal tests should only be done using drugs that are known to have bactericidal activity against the isolate being tested. At present, there is no standardized test procedure, but

the NCCLS subcommittee has released an approved testing guideline (M26-A).

For efficiency, the MBC should be requested at the same time that the routine susceptibility test is requested, especially if the routine method is a broth MIC test. The endpoint being sought is the lowest concentration of antimicrobic that kills at least 99.9% of the viable inoculum after 18 to 24 hours of incubation. This endpoint is used because it is difficult to kill 100% of the organisms. A small number of "persisters" survive, probably from low metabolic activity. To obtain this information, the exact number of viable organisms used for inoculation must be known. Also, because the cidal activity of many antimicrobics is dependent on active cellular metabolic activity, the test must be performed using actively growing bacterial cells.

The isolate to be tested is inoculated into a nutrient broth and allowed to grow at 35°C to 37°C until it reaches late log phase. The growth must be adjusted to obtain an estimated final concentration in each well of about 5×10^5 CFU/mL. A sample is further diluted and plated for viable colony count. The rest of the setup is the same as with the routine microdilution MIC.

Following the reading of the MIC, all wells of drug(s) being tested that show no visual growth must be sampled by removing an aliquot and using it to inoculate a nutrient agar plate. The inoculum should be spread out over an area of at least 5 cm^2 to provide isolated colonies. The inoculated plate is then incubated long enough to obtain visible colonies and the number of colonies counted. Care must be taken to sample an adequate volume from each well. The endpoint is the drug concentration that is able to kill at least 99.9% of the original inoculum (5×10^5 CFU/mL). Accordingly, the residual viable organism concentration must be 500 CFU/mL or less. To detect 500 CFU/mL, one would have to plate an undiluted volume of 0.01 mL to obtain five colonies.

The 99.9% kill endpoint remains controversial because there are no convincing clinical studies showing that concentrations that kill 99% or even 90% of inoculated organisms result in more frequent treatment failure. Another area of controversy is the meaning of in vitro tolerance. The working definition of tolerance is when the MBC:MIC is greater than or equal to 32:1. For example, enterococci frequently show tolerance to the cidal activity of vancomycin. Therefore, vancomycin is not effective when used alone for the treatment of endocarditis from *Enterococcus* species (70). The results obtained are very technique-dependent, and laboratories not using consistent methods will get variable results with the same isolate. Another phenomenon seen with this technique is the "paradoxical" or Eagle effect, where killing is seen with lower antimicrobic concentrations but less killing with higher concentrations (71). The higher drug concentrations are thought to inhibit various enzymatic steps required to kill the organism resulting in a bacteriostatic condition. With rapid dilution of the drug, the organisms remain viable.

The MBC has been performed with both micro- and macrodilution techniques using volumes of 0.01 mL and up to 5 mL/assay tube, respectively. Discrepant results are frequently obtained between these two methods depending on the technique used. The current NCCLS approved guideline for bactericidal tests advocates using the microdilution technique to enhance reproducibility and convenience, but there are no clinical data to support the use of one method over the other.

The Serum Bactericidal Test

The serum bactericidal test (SBT) measures the level of cidal antimicrobial activity in patient serum directed against the patient's isolate. The result is expressed as the highest serum dilution that retains the ability to kill 99.9% of the inoculum (72). The NCCLS has published an Approved Guideline for this test, M 21A. Sera are collected when the concentration of the antimicrobic(s) being administered is either at its peak, at its nadir, or at both times. For test interpretation, both the time and route of drug administration and the time that the blood was drawn must be accurately documented. Serial doubling dilutions of the serum are prepared using Mueller-Hinton broth containing 50% normal human serum as the diluent yielding dilutions from 1:2 to 1:256. The inoculum is prepared in the same way as described for the MBC test. A quantitative viability count is necessary to document the starting inoculum concentration. After 18 to 24 hours of incubation, each tube or well is examined for visible growth. All tubes showing no growth are sampled by removing and plating an aliquot and counting the number or colonies after overnight incubation and the 99.9% killing endpoint determined. If both peak and trough sera were assayed, there is usually a significant drop in titer with normal drug clearance.

Test interpretation remains a problem (73). If the patient is being treated for endocarditis, clinicians may want the nadir (trough) titer to stay above 1:8 (74). For the treatment of staphylococcal osteomyelitis using oral therapy, they may want to see high peak titers indicating good absorption of the drug from the gastrointestinal tract and patient compliance with dosage orders. The number of requests for SBTs remains low in most tertiary care centers and is used primarily for cases of endocarditis and osteomyelitis being treated with drug combinations or for patients being switched from parenteral to oral therapy (75).

Selected tests methods have been used to test for antimicrobical interaction but none have proven to be strongly predictive of clinical outcome and, therefore, are infrequently ordered despite the frequency with which drug combinations are used clinically (76, 77 and 78).

Disk Approximation Synergy Test

Disk Approximation is the least complex method to be used as a screening method to assess how two drugs interact when combined *in vitro*. The test is performed in two stages. The first stage is to perform a routine disk diffusion test to obtain the zone of inhibition radius for the drugs to be tested. The second stage is to repeat the setup but place the disks such that one would expect to obtain a small zone of growth (e.g., 4 mm to 5 mm) between the two zones of inhibition produced around each disk if there was no drug interaction. After an adequate incubation time, the area between the two disks is examined for any distortion in the zone of inhibition. If no growth is seen between the two disks, this is evidence

for possible drug synergy. If a zone is blunted, this may be evidence of drug antagonism or enzyme induction (Fig. 55.7).

Time-Kill Kinetics Test

As performed, the MBC and the SBT measure the number of surviving organisms at one time point, usually after 24 hours of exposure to varying concentrations of the drug. The time-kill kinetic assay measures the rate of killing while being exposed to a single drug concentration, usually one that is physiologically achievable at standard dosage (79). Tubes containing Mueller-Hinton broth with and without the test antimicrobic(s) are inoculated with the test organism at an estimated viable concentration (documented by colony counts at the time of inoculation) and maintained at 35°C during the testing period. Aliquots are removed at periodic intervals, usually at 2, 4, 8, or 12 and 24 hours, and plated to determine the number of viable organisms at each time point. The results are plotted using semilog paper or standard graph paper after calculating the $\log_{10}$ of each data point. The viable concentration (in CFU/mL or $\log_{10}$ CFU/mL) is plotted along the y-axis as a function of time on the x-axis. Drugs showing more complete and/or rapid rates of killing may be more efficacious in treating infections occurring in individuals with low immunocompetence.

The time-kill kinetic assay is currently thought to be the best method to study in vitro drug synergy. As the term implies, drugs that act in synergy with one another are more efficient in killing than either one alone acting independently (80). The cidal activity of each drug must be tested as well as the combination. If more than one concentration of either drug is tested, the testing complexity becomes geometric. A combination is considered to be acting in a synergistic manner when the viable count is reduced at least 100-fold compared with the most actively cidal drug of the pair (Fig. 55.9).

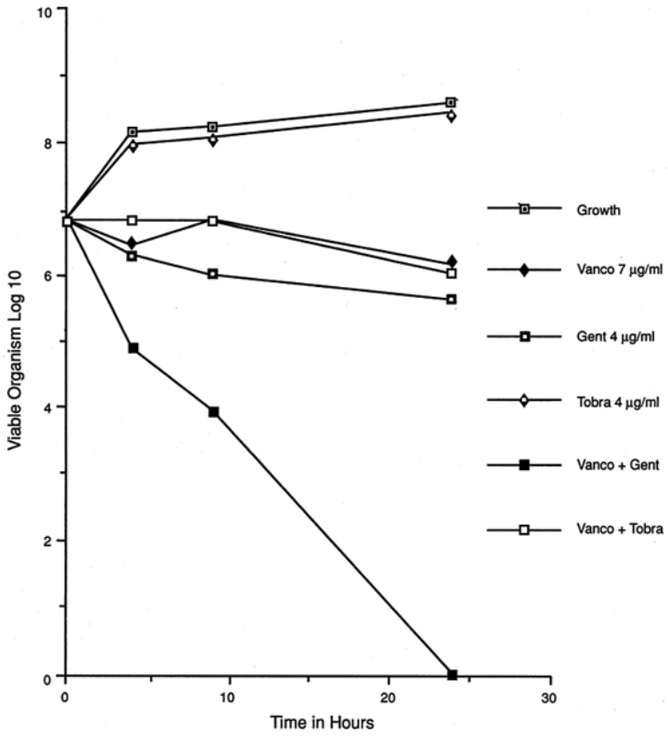


FIGURE 55.9. Time-kill kinetics. Rate of bactericidal activity of single and combined agents on *Enterococcus faecium*. Synergistic killing is occurring with the combination of vancomycin plus gentamicin.

Checkerboard Microdilution Test

An alternative method of performing drug synergy studies is the "checkerboard" microdilution technique (81). A standard 96-well microtiter tray is loaded horizontally with increasing concentrations of drug A, and vertically with increasing concentrations of drug B so that as one scans the plate from lower left to upper right, both drugs in each well are present in increasing concentrations that span the therapeutic range from easily achievable (susceptible) to unachievable (resistant) (Fig. 55.10). Each drug also is present in either a row or column as a single agent to determine the MIC and MBC of each agent without the other being present. The wells are inoculated with the test organism prepared as just described and the plate incubated for 18 to 24 hours.

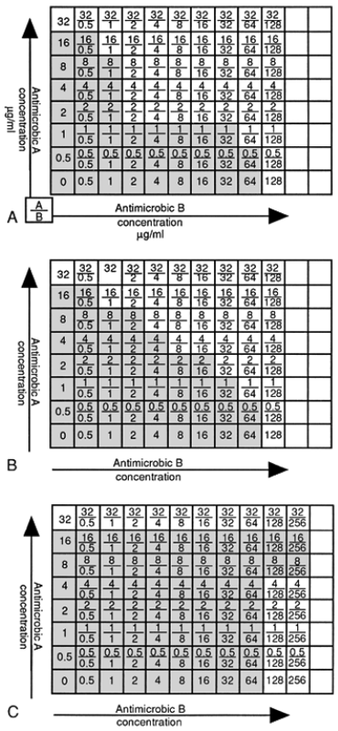


FIGURE 55.10. "Checkerboard" technique for determining antimicrobic interaction. (**A**) Drug synergy. (**B**) Additive effect. (**C**) Drug antagonism.

Because both drugs are present at their lowest concentrations at the lower left, one would expect growth to occur in this region of the plate. As one progresses toward the upper right, however, a point is reached for each dilutional series where no visible growth occurs. If one marks the first well of the lowest concentrations in each row and column where no growth is seen and connects these marks with a continuous line, a crude isobologram is produced, the shape of which connotes the test results. If the line bows substantially toward the lower left, the two drugs are expressing synergy; if it bows substantially toward the upper right, the two drugs are antagonistic; if the line is essentially straight, the effects are additive. A more practical method of interpretation is to express the results as a fractional index. The concentration of drug A necessary to inhibit or kill the organism in any row or column is divided by the concentration needed to do the same without drug B present. This is the fractional inhibitory or cidal concentration (FIC) of drug A (FIC_A). The same is done for drug B to obtain the FIC_B. The two factional concentrations are added together to obtain the fractional index.

$$\frac{A}{\mathrm{MIC_A}} + \frac{B}{\mathrm{MC_B}} = \mathrm{FC_A} + \mathrm{FC_B} = \mathrm{FI}$$

There are several possible choices of combinations to choose for calculating the index. Thus, the logical choice is to select those concentrations that are within the lower physiological range for each drug. If the index is 0.5 or less, the drugs are acting with synergy. As the index increases, the drug interactions progress from additive to indifferent to antagonistic.

REFERENCES

1. Washington JA II. Discrepancies between in vitro activity of and in vivo response to antimicrobial agents. *Diagn Microbiol Infect Dis* 1983;1:25–31.

2. Williams JD. The correlation of in vitro susceptibility tests with in vivo results of antibiotic treatment. *Scand J Infect Dis* 1978;13(suppl):64–66.

3. Greenwood D. In vitro *veritas*? Antimicrobial susceptibility tests and their clinical relevance. *J Infect Dis* 1981;144:380–385.

4. Snydman DR, Cuchural GJ Jr, McDermont L, et al. Correlation of various in vitro testing methods with clinical outcomes in patients with *Bacteroides fragilis* group infections treated with cefoxitin: a retrospective analysis. *Antimicrob Agents Chemother* 1992;36:540–544.

5. Sherris JC. Antimicrobic susceptibility testing: a personal perspective. *Clin Lab Med* 1989;9:191–202.

6. Ericsson HM, Sherris JC. Antibiotic sensitivity testing-report of an International Collaborative Study. *Acta Pathol Microbiol Scand* [B] 1971;217(suppl):1–90.

7. National Committee for Clinical Laboratory Standards. *Performance Standards for Antimicrobial Disk Susceptibility Tests*. 4th ed. Standard M2-A4. Villanova, PA: National Committee for Clinical Laboratory Standards, 1990.

8. National Committee for Clinical Laboratory Standards. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically*. 2nd ed. Standard M7-A2. Villanova, PA: National Committee for Clinical Laboratory Standards, 1990.

9. Hindler JA, Swenson JM. Susceptibility testing of fastidious bacteria. In: Murray PR, Baron EJ, Pfaller MA, Tenover FC, and Yolken RH. eds. *Manual of Clinical Microbiology*. 7th ed. Washington D.C.: American Society for Microbiology, 1999:1544–1554.

10. D'Amato RF, Thornsberry C. Calcium and magnesium in Mueller-Hinton agar and their influence on disk diffusion susceptibility results. *Curr Microbiol* 1979;2:135–138.

11. Barry AL, Reller LB, Miller GH, et al. Revision of standards for adjusting the cation content of Mueller-Hinton broth for testing susceptibility of *Pseudomonas aeruginosa* to aminoglycosides. *J Clin Microbiol* 1992;30:585–589.

12. Pfaller MA, Wakefield DS, Stewart B, et al. Evaluation of laboratory methods for the classification of oxacillin-resistant and oxacillin-susceptible *Staphylococcus aureus*. *Am J Clin Pathol* 1988;89:120–125.

13. Reimer LG, Stratton CW, Reller, LB. Minimum inhibitory and bactericidal concentrations of 44 antimicrobial agents against three standard control strains in broth with and without human serum. *Antimicrob Agents Chemother* 1981;19:1050–1055.

14. Traub WH, Spohr M, Bauer D. *Pseudomonas aeruginosa*: In vitro susceptibility to antimicrobial drugs, single and combined, with and without defibrinated human blood. *Chemother* 1988;34:284–297.

15. Pruul H, McDonald PJ. Potentiation of antibacterial activity of azithromycin and other macrolides by normal human serum. *Antimicrob Agents Chemother* 1992;36:10–16.

16. Kohner P, Uhl J, Kolbert C, et al. Comparison of susceptibility tersting methods with *mecA* gene analysis for determining oxacillin (methicillin) resistance in clinical isolates of *Staphylococcus aureus* and coagulase-negative *Staphylococcus* spp. *J Clin Microbiol* 1999;37:2952–2961.

17. Hwang JMD, Piccinini TE, Lammel CL, et al. Effect of storage temperature and pH on the stability of antimicrobial agents in MIC trays. *J Clin Microbiol* 1986;23:959–961.

18. Bulger RR, Washington JA II. Effect of inoculum size and ß-lactamase production on in vitro activity of new cephalosporins against *Haemophilus* species. *Antimicrob Agents Chemother* 1980;17:393–396.

19. Gay JD, DeYoung DR, Roberts GD. In vitro activities of norfloxacin and ciprofloxacin against *Mycobacterium tuberculosis*, *M. avium*, *M. chelonei*, *M. fortuitum* and *M. kansasii*. *Antimicrob Agents Chemother* 1984;26:94–96.

20. Swenson JM, Wallace Jr, RJ, Silcox VA, and Thornsberry C. Antimicrobial susceptibility of *Mycobacteria fortuitum* and *Mycobacterium chelonae*. *Antimicrob Agents Chemother* 1985;28:807–811.

21. Ferguson RW, Weissfield AS. Comparison of the suitability of three common bacterial media for susceptibility testing of trimethoprim-sulfamethoxazole. *J Clin Microbiol* 1984;19:85–86.

22. Flournoy DJ, Shirley RA. Influence of inoculum concentration on zone size in disk diffusion testing. *Lab Med* 1985;16:616–618.

23. Berke I, Tierno PM. Comparison of efficacy and cost-effectiveness of BIOMIC VIDEO and vitek antimicrobial susceptibility test system for use in the clinical microbiology laboratory. *J Clin Microbiol* 1996;34:1980–1984.

24. Baker CN, Stocker SA, Culver DH, et al. Comparison of the E-Test to agar dilution, broth microdilution, and agar diffusion susceptibility testing techniques by using a special challenge set of bacteria. *J Clin Microbiol* 1991;29:533–538.

25. Jorgensen JH, Ferraro MJ, McElmeel ML, et al. Detection of penicillin and extended-spectrum cephalosporin resistance among Streptococcus pneumoniae clinical isolates by use of the E test. *J Clin Microbiol* 1994;32:159–163.

26. Cormican MG, Marshall SA, Jones RN. Detection of extended-spectrum ß-lactamase (ESBL)-producing strains by the Etest ESBL screen. *J Clin Mictobiol* 1996;34:1880–1884.

27. Huang M, Baker PN, Banerjee S, et al. Accuracy of the E test for determining antimicrobial susceptibilities of staphylococci, enterococci, Campylobacter jejuni, and gram-negative bacteria resistant to antimicrobial agents. *J Clin Microbiol* 1992;30:3243-3248.

28. Craig WA. Qualitative susceptibility tests versus quantitative MIC tests. *Diagn Microbiol Infect Dis* 1993;16:321–236.

29. Fuchs PC, Barry AL, Brown SD, et al. Reproducibility of broth microdilution and disk diffusion susceptibility tests of nine antimicrobial agents against *Streptococcus pneumoniae* ATCC 49619. *Diagn Microbiol Infect Dis* 1997;28:27–29.

30. Citron DM, Ostoravi MI, Karlsson A, Goldstein EJC. Evaluation of the E test for susceptibility testing of anaerobeic bacteria. *J Clin Microbiol* 1991;29:2197–2203.

31. Rosenblatt J, Gustafson, DR. Evaluation of the Etest for susceptibility testing of anaerobic bacteria. *Diagn Microbiol Infect Dis* 1995;22;279–284.

32. Finegold SM. Susceptibility testing of anaerobic bacteria. *J Clin Microbiol* 1988;26:1253–1256.

33. Aldridge KE, Gelfand M, Reller LD, et al. A five-year multicenter study of the susceptibility of the *Bacteroides fragilis* group isolates to cephalosporins, cephamycins, penicillins, clindamycin, and metronidazole in the United States. *Diagn Microbiol Infect Dis* 1994;18:235–241.

34. Snydman DR, Jacobus NV, McDermott LA, et al. Multicenter study of in vitro susceptibility of the *Bacteroides fragilis* group, 1995-1996, with comparison of resistance trends from 1990 to 1996. *Antimicrob Agents Chemother*. 1999;43:2417–2422.

35. Wilkins TD, Thiel T. Modified broth-disk method for testing the antibiotic susceptibility of anaerobic bacteria. *Antimicrob Agents Chemother* 1973;3:350–356.

36. Bush K, Jacoby GA, and Medeiros, AA. A functional classification scheme for ß-lactamases and its correlation with molecular structure. *Antimicrob Agents Chemother* 1995;39:1211–1233.

37. Sanders WE, Sanders CC. Inducible ß-lactamases: clinical and epidemiologic implications for use of newer cephalosporins. *Rev Infect Dis* 1988;10:830–838.

38. Bush K. Beta-lactamase inhibitors from laboratory to clinic. *Clin Microbiol Rev* 1988;1:109–123.

39. Doern BV, Jones RN, Gerlach EH, et al. Multicenter clinical laboratory evaluation of a ß-lactamase disk assay employing a novel chromogenic cephalosporin, S1. *J Clin Microbiol* 1995;33:1665–1667.

40. Dyke JW, Brown WJ, Morris JM, et al. Evaluation of the new Dry-Slide beta-lactamase test. *Am JH Clin Pathol* 1994;101:726–728.

41. Livermore DM. ß-lactamases in laboratory and clinical resistance. *Clin Microbiol Rev* 1995;8:557–584.

42. Knapp CC, Washington JA. Conparison of the E test and microdilution for detection of ß-lactam-resistant mutants that are stably derepressed for type 1 ß-lactamase. *J Clin Microbiol* 1992;30:214–215.

43. Tenover FC, Jones RN, Swenson JM. et al. Methods for improved detection of oxacillin resistance in coagulase-negative staphylococci: results of a multicenter study. *J Clin Microbiol* 1999;37:4051–4058.

44. Van Horn KG, Gedris CA, Rodney KM, Mitchell JB. Evaluation of commercial vancomycin agar screening plates for detection of vancomycin-resistant enterococci. *J Clin Microbiol* 1996;34:2042–2044.

45. Swenson JM, Clark N, Ferraro MJ et al. Development of a standardized screening method for detection of vancomycin-resistant enterococci. *J Clin Microbiol* 1994;32:1700–1704.

46. Tenover FC, Lancaster MV, Hill BC, et al. Characterization of staphylococci with reduced susceptibility to vancomycin and other glycopeptides. *J Clin Microbiol* 1998;36:1020–1027.

47. Hubert SK, Mohammed JM, Fidkin FC, et al. Glycopeptide-intermediate *Staphylococcus aureus*: evaluation of a novel screening method and results of a survey of selected U.S. hospitals. *J Clin Microbiol* 1999;37:3583–3585.

48. Doern GV, Brueggemann AB, Pierce G. Assessment of the oxacillin disk screening test for determining penicillin resistance in *Streptococcus pneumoniae*. *Eur J Clin Microbiol Infect Dis* 1997;16:311–331.

49. Sanders CC, Sanders WE. Microbial resistance to newer generation of ß-lactam antibiotics: clinical and laboratory implications. *J Infect Dis* 1985;151:399–406.

50. Vercauteren E, Descheemaecker P, Ieven M, et al. Comparison of screening methods for detection of extended-spectrum ß-lactamases and their prevalence among blood isolates of *Escherichia coli* and *Klebsiella* spp. in a Belgian teaching hospital. *J Clin Microbiol* 1997;35:2191–2197.

51. Katsanis GP, Spargo J, Ferraro MJ, et al. Detection of Klebsiella pneumoniae and *Escherichia coli* strains producing extended-spectrum ß-lactamases. *J Clin Microbiol* 1994;32:691–696.

52. Jacoby GA, Han P. Detection of extended-spectrum ß-lactamases in clinical isolates of *Klebsiella pneumoniae* and *Escherichia coli*. *J Clin Microbiol* 1996;34:908–911.

53. Thomson KS, Sanders CC, and Moland ES. Use of microdilution panels with and without ß-lactamase inhibitors as a phenotypic test for ß-lactamase production among *Escherichia coli*, *Klebsiella* spp., *Enterobacter* spp., *Citrobacter freundii*, and *Serratia marcescens*. *Antimicrob Agents Chemother* 1999;43:1393–1400.

54. Pitout JD, Moland ES, Sanders CC, et al. ß-lactamases and detection of ß-lactam resistance in *Enterobacter* spp. *Antimicrob Agents Chemother* 1997;41:35–39.

55. Walker CW, Brown DFJ. The reliability of methods for detecting chloramphenicol resistance in *Haemophilus influenzae*. *J Antimicrob Chemother* 1988;22:905–910.

56. Courvalin P. Genotypic approach to the study of bacterial resistance to antibiotics. *Antimicrob Agents Chemother* 1991;35:1019–1023.

57. Tenover FC. Studies of antimicrobial resistance genes using DNA probes. *Antimicrob Agents Chemother* 1986;29:721–725.

58. Pierson CL, McClatchey KD. Automation in clinical microbiology. In: Lederberg J, ed. *Encyclopedia of Microbiology*. New York: Academic Press 1992;1:171–179.

59. Jorgenson J. *Automation in Clinical Microbiology*. Boca Raton, FL: CRC Press, 1987.

60. Knapp C, Ludwig MD, Washington JA, et al. Evaluation of the Vitek GPS-SA card for testing of oxacillin against borderline-susceptible staphylococci that lack mecA. *J Clin Microbiol* 1996;34:1603–1605.

61. Jett B, Free L, Sahm DF. Factors influencing the Vitek gram-positive susceptibility system's detection of *vanB*-encoded vancomycin resistance among enteroccci. *J Clin Microbiol* 1996;34:701–706.

62. Jorgensen J, McElmeel M, Crawford S. Evaluation of the Dade MicroScan MICroSTREP antimicrobial susceptibility testing panel with selected *Streptococcus pneumoniae* challenge strains and recent clinical isolates. *J Clin Microbiol* 1998;36:788–791.

63. Cherubin CD, Eng R, Appleman M. A critique of semiautomated susceptibility systems. *Rev Infect Dis* 1987;9:655–659.

64. Kiehlbauch J, Kendle JM, Carlson LG, et al. Automated antibiotic susceptibility testing: comparative evaluation of four commercial systems and present state. *Clin Lab Med* 1989;9:319–340.

65. Tenover FC, Baker CN, Swenson JM. Evaluation of commercial methods for determining antimicrobial susceptibility of *Streptococcus pneumoniae*. *J Clin Microbiol* 1996;34:10–14.

66. Tenover FC, Swenson JM, O'Hara CM, et al. Ability of commercial and reference antimicrobial susceptibility testing methods to detect vancomycin resistance in enterococci. *J Clin Microbiol* 1995;33:1524–1527.

67. Sherris JC. Problems in in vitro determination of antibiotic tolerance in clinical isolates. *Antimicrob Agents Chemother* 1986;30:633–637.

68. James PA. Comparison of four methods for the determination of MIC and MBC of penicillin for viridans sptreptococci and the implications for penicillin tolerance. *J Antimicrob Chemother* 1990;25:209–216.

69. Betriu C, Gomez M, Sanchez A, et al. Antibiotic resistance and penicillin tolerance in clinical isolates of group B streptococci. *Antimicrob Agents Chemother* 1994;38:2183–2186.

70. Murray BE. The life and times of the enterococcus. *Clin Microbiol Rev* 1990;3:46–65.

71. Eagle H, Musselman AD. The rate of bactericidal action of penicillin in vitro as a function of its concentration, and its paradoxically reduced activity at high concentrations against certain organisms. *J Exp Med* 1948;88:99–131.

72. Schlichter JG, MacLean H. A method of determining the effective therapeutic level in treatment of subacute bacterial endocarditis with penicillin. *Am Heart J* 1947;34:209–211.

73. MacGowan A, McMullin C, James P, et al. External quality assessment of the serum bactericidal testL results of a methodology/interpretation questionnaire. *J Antimicrob Chemother* 1997;39:277–284.

74. Reller LB. The serum bactericidal test. *Rev Infect Dis* 1986;8:803–808.

75. Standiford HC, Tatem BA. Technical aspects and clinical correlations of the serum bactericidal test. *Eur J Clin Microbiol Infect Dis* 1986;5:79–87.

76. Visallli M, Jacobs MR, and Appelbaum PC. Determination of activities of levofloxacin, alone and combined with gentamicin, ceftazidime, cefpirome and meropenem against 124 strains of *Pseudomonas aeruginosa* by checkerboard and time-kill methodology. *Antimicrob Agents Chemother* 1998;42:953–955.

77. Cappelletty DM, Rybak MJ. Comparison of methodologies for synergism testing of drug combinations against resistant strains of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 1996;40:677–683.

78. D'Alessandri RM, McNeely DJ, Kluge RM. Antibiotic synergy and antagonism against clinical isolates of *Klebsiella* species. *Antimicrob Agents Chemother* 1998;10:889–892.

79. Eliopoulos GM, Moellering Jr, RC. Antimicrobial combinations. In: Lorian V, ed. *Antibiotics in Laboratory Medicine*. 3rd ed. Baltimore: Williams & Wilkins, 1991;432–492.

80. Blaser J. Interactions of antimicrobial combinations in vitro: the relativity of synergism. *Scand J Infect Dis* 1991;74(suppl):71–79.

81. Van der Auwera P. Synergistic, additive, and antagonistic activity of antimicrobial agents: how to measure? What are the clinical consequences? Applications to the new fluoroquinolones. *Infect Dis Newsl* 1992;11:49–54.

56

Molecular Techniques for Diagnosis of Infectious Diseases

Frederick S. Nolte

Since the publication of the first edition of this book significant changes have occurred in the practice of diagnostic molecular microbiology. Chief among them is the widespread application of nucleic acid amplification techniques to the diagnosis and management of patients with infectious diseases. The advent of commercially available test kits has facilitated the use of this technology in the clinical laboratory. Molecular microbiology has emerged as the leading area in molecular pathology in terms of both numbers of tests performed and clinical relevance. This technology has reduced the dependency of the clinical microbiology laboratory on culture-based methods and created new opportunities for the clinical laboratory to impact patient care. This chapter will cover amplified and nonamplified probe techniques, postamplification detection and analysis, clinical applications of these techniques, and the special challenges and opportunities that these techniques provide for the clinical laboratory.

NUCLEIC ACID PROBES

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

Nucleic acid probes are pieces of DNA or RNA labeled with radioisotopes, enzymes, or chemiluminescent reporter molecules that can bind to complementary nucleic acid sequences with high specificity. Probes range from 15 to thousands of nucleotides in size, however synthetic oligonucleotides (<50 nucleotides) are the most commonly incorporated into commercial kits. The probes can be designed to identify microorganisms at any taxonomic level (e.g., genus, species, or subspecies). A number of commercially available DNA probes have been developed for direct detection of pathogens in clinical specimens and identification of pathogens after isolation by culture.

The commonly used formats for probe hybridization includes liquid-phase, solid phase, and in situ hybridization. The leading method used in clinical microbiology laboratories is a liquid-phase hybridization protection assay (Gen-Probe). In this method a single-stranded DNA probe labeled with an acridinium ester is incubated with the target nucleic acid. Alkaline hydolysis follows the hybridization step and probe binding is measured in a luminometer after addition of peroxides. In a positive sample, the acridinium ester on the bound probe is protected from hydrolysis and upon addition of peroxides emits light. The hybridization protection assay can be completed in several hours and does not require removal of unbound single-stranded probe or isolation of probe-bound double stranded sequences (1).

In solid-phase hybridization, target nucleic acids are bound to nylon or nitrocellulose and hybridized with a probe in solution (2). The unbound probe is washed away and the bound probe is detected by means of fluorescence, luminescence, radioactivity, or color development. Although solid-phase hybridization is a powerful research tool, the length of time required and the complexity of the procedure limit its application in clinical practice.

In situ hybridization is another type of solid-phase hybridization where the nucleic acid is contained in tissues or cells affixed to microscope slides and is governed by the same basic principles (3). In most clinical applications, formalin-fixed, paraffin embedded tissue sections are used. The sensitivity of *in situ* hybridization is limited by the accessibility of target nucleic acid in the cells.

In general, the application of nonamplified probe techniques to direct detection of pathogens in clinical specimens is limited to those situations where the number of pathogens is great because of the poor analytical sensitivity of these techniques. These include Group A streptococcal pharyngitis and genital tract infections with *Neisseria gonorrhoeae* and *Chlamydia trachomatis*. These techniques are used to their best advantage in culture confirmation assays for mycobacteria and systemic dimorphic fungi. These culture confirmation tests have a positive impact on patient management by providing rapid and accurate diagnoses for these slowly growing, often difficult to identify pathogens.

NUCLEIC ACID AMPLIFICATION TECHNIQUES

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

The development of the polymerase chain reaction (PCR) by Mullis and colleagues (4) was a milestone in biotechnology and heralded the beginning of molecular diagnostics. Although PCR is the best developed and most widely used nucleic amplification strategy, other strategies have been developed and several have commercial potential. These strategies are based on target, probe, or signal amplification. Examples of each category will be discussed in the sections that follow. These techniques have sensitivity unparalleled in laboratory medicine, have created new opportunities for the clinical laboratory to impact patient care,

and have become the new gold standards for laboratory diagnosis of several infectious diseases. Table 56.1 compares the features of the nucleic acid amplification methods.

TABLE 56.1. NUCLEIC ACID AMPLIFICATION METHODS[a]

Method	Category	Manufacturer	Amplification System	Temperature Requirement	Nucleic Acid Target	Key Reference
PCR	Target	Roche	DNA polymerase	Thermal cycle	DNA (RNA)	(4)
TMA	Target	Gen-Probe	RT, RNA polymerase, RNase H	Isothermal	RNA (DNA)	(11)
NASBA	Target	Organon-Teknika	RT, RNA polymerase, RNase H	Isothermal	RNA (DNA)	(10)
SDA	Target	Becton-Dickinson	Restriction endonuclease, DNA polymerase	Isothermal	DNA (RNA)	(13)
LCR	Probe	Abbott	DNA ligase	Thermal cycle	DNA	(17)
bDNA	Signal	Bayer	Branched DNA probes	Isothermal	DNA, RNA	(24)
Hybrid capture	Signal	Digene	anti-DNA-RNA hybrid antibody	Isothermal	DNA, RNA	(28)

[a] Adapted from reference (66.)

Target Amplification Techniques

All of the target amplification systems share certain fundamental characteristics. They are enzyme-mediated processes, in which a single enzyme or multiple enzymes synthesize copies of target nucleic acid. The amplification products in all of the techniques are defined by two oligonucleotide primers that bind to complementary sequences on opposite strands of double-stranded targets. All result in the production of millions to billions of copies of the targeted sequence in a matter of hours, and, in each case, the amplification products can serve as templates for subsequent rounds of amplification. Because of this, all of the techniques are sensitive to contamination with product molecules that can lead to false-positive reactions. False-positive reactions because of product cross-contamination has impeded the routine use of techniques in the clinical laboratory.

Polymerase Chain Reaction

PCR is a simple, *in vitro*, chemical reaction that permits the synthesis of essentially limitless quantities of a targeted nucleic acid sequence. This is accomplished through the action of a DNA polymerase, that under the right conditions, can copy a strand of DNA. At its simplest, a PCR consists of target DNA, a molar excess of two oligonucleotide primers, a heat-stable DNA polymerase, an equimolar mixture of deoxyribonucleotide triphosphates (dATP, dCTP, dGTP, and dTTP), $MgCl_2$, KCl, and a Tris-HCl buffer. The two primers flank the sequence to be amplified, typically <100 to several hundred bases, and are complementary to opposite strands of the target.

To initiate a PCR, the reaction mixture is heated to separate the two strands of target DNA, and then cooled to permit the primers to anneal to the target DNA in a sequence specific manner. The DNA polymerase then initiates extension of the primers at their 3′ ends toward one another. The primer extension products are dissociated from the target DNA by heating. Each extension product, as well as the original target, can serve as a template for subsequent rounds of primer annealing and extension.

A PCR cycle consists of denaturation and annealing. At the end of each cycle, the PCR products are theoretically doubled. Thus, after n PCR cycles the target sequence can be amplified 2^n-fold. The whole procedure is carried out in a programmable thermal cycler that precisely controls the temperature at which the steps occur, length of time that the reaction is held at the different temperatures and the number cycles. Ideally, after 20 cycles of PCR, a million-fold amplification is achieved and after 30 cycles, a billion-fold. In practice, the amplification may not be completely efficient because of failure to optimize the reaction conditions or the presence of inhibitors of the DNA polymerase. In such cases, the total amplification is best described by the expression $(1+e)^n$ where e is the amplification efficiency ($0 \geq e \leq 1$) and n is the total number of cycles.

Reverse transcriptase-PCR

PCR as it was originally described was a technique for DNA amplification. Reverse transcriptase-PCR (RT-PCR) was developed to amplify RNA targets. In this process complementary DNA (cDNA) is first produced from RNA targets by reverse transcription, and then the cDNA is amplified by PCR. As originally described RT-PCR employed two enzymes, a heat-labile RT such as avian myeloblastosis virus reverse transcriptase (AMV-RT), and a thermostable DNA polymerase. Because of the temperature requirements of the heat-labile enzyme, cDNA synthesis had to occur at lower temperatures. This presented problems both in terms of the nonspecific primer annealing and inefficient primer extension because of formation of RNA secondary structures. These problems have been largely overcome by the development of a thermostable DNA polymerase derived from *Thermus thermophilus* that under the proper conditions can function efficiently as both a RT and a DNA polymerase (5). RT-PCRs using this enzyme are more specific and efficient than previous protocols using conventional, heat-labile RT enzymes. Commercially available kits (Roche) employing this single enzyme technology are available for detection of hepatitis C virus (HCV) RNA and for quantitation of human immunodeficiency virus type 1 (HIV-1) and HCV RNA in clinical specimens.

Nested PCR

Nested PCR was developed to increase both the sensitivity and specificity of PCR (6). It employs two pairs of amplification

primers and two rounds of PCR. Typically, one primer pair is used in the first round of PCR of 15 to 30 cycles. The products of the first round of amplification are then subjected to a second round of amplification using the second set of primers that anneal to a sequence internal to sequence amplified by the first primer set. The increased sensitivity arises from the high total cycle number and the increased specificity arises from the annealing of the second primer set to sequences found only in the first round products, thus verifying the identity of the first round product. The major disadvantage of nested PCR is the high rates of contamination that can occur during the transfer of first round products to the second tube for the second round of amplification. This can be avoided either by physically separating the first- and second-round amplification mixtures with a layer of wax or oil, or by designing single-tube amplification protocols. In practice, the enhanced sensitivity afforded by nested PCR protocols is required rarely in diagnostic applications, and the identity of an amplification product usually is confirmed by hybridization with a nucleic acid probe.

Multiplex PCR

In multiplex PCR, two or more primer sets, designed for amplification of different targets, are included in the same reaction mixture (7). With this technique, more than one target sequence in a clinical specimen can be coamplified in a single tube. The primers used in multiplexed reactions must be selected carefully to have similar annealing temperatures and to lack complementarity. Multiplex PCRs are proven more complicated to develop and are less sensitive than single primer set PCR reactions.

Quantitative PCR and Reverse Transcriptase PCR

A linear relationship may exist between the quantity of input template and the amount of amplification product. However, because the final amount of PCR product depends on exponential amplification of the initial quantity of template, minor differences in amplification efficiency may lead to very large and unpredictable differences in the final product yield (8). The tube-to-tube differences may depend on sample preparation and nucleic acid purification procedures, presence of inhibitors, and thermal cycler performance. For these reasons, simple quantitation of amplified product and the use of external standard reference curves do not provide reliable quantitation of the template initially present in the sample.

A variety of PCR-based strategies have been developed to accurately quantitate DNA and RNA targets in clinical specimens. It generally is accepted that a competitive PCR (cPCR) approach is the most reliable and robust.

The basic concept behind cPCR is the coamplification in the same reaction tube of two different templates of equal or similar lengths and with the same primer binding sequences. Because both templates are amplified with the same primer pair, identical thermodynamics and amplification efficiency are ensured. The amount of one of the templates must be known and, after amplification, products from both templates must be distinguishable. Different types of competitors have been used in cPCR, but in general, those competitors similar in size and base composition to the target work most effectively. RNA competitors should be used in quantitative RT-PCRs to address the problem of variable RT efficiency.

The yield of PCR product is described by the equation, $Y=I(1+e)^n$, where Y is the quantity of PCR product, I is the quantity of template at the beginning of the reaction, *e* is the efficiency of the reaction, and *n* is the number of cycles. In cPCR, this equation is written for both templates, as follows: competitor, $Y_c=I_c(1+e)^n$; and target, $Y_t=I_t(1+e)^n$. Because *e* and *n* are the same for both the competitor and target, the relative product ratio Y_c/Y_t directly depends on their initial concentration ratio I_c/I_t and the function, $Y_c/Y_t=I_c/I_t$, is linear.

A single concentration of competitor is sufficient, in theory, to quantitate an unknown amount target without the use of a standard curve. However, because analysis of two template species present in a sample at widely different amounts may be difficult and imprecise in practice, cPCR using several concentrations of competitor within the expected concentration range of the target were generally performed. However, this approach provided no more accurate results than the use of a single concentration of competitor in a recent study of different approaches to standardization of cPCR (9). The commercially available quantitative PCR and RT-PCR assays or cytomegalovirus (CMV), HIV-1, and HCV (Roche) all use a single concentration of a competitor to determine the initial concentration of the target.

Transcription Amplification Methods

Nucleic acid sequence-based amplification (NASBA) and transcription mediated amplification (TMA) both are isothermal RNA amplification methods modeled after retroviral replication (10,11 and 12). The methods are similar in that the RNA target is reversed transcribed into cDNA and then RNA copies are synthesized using an RNA polymerase. NASBA uses AMV-RT, RNase H, and T7-RNA polymerase while TMA uses a RT enzyme with endogenous RNase H activity and T7-RNA polymerase.

Amplification involves the synthesis of cDNA from the RNA target with a primer containing the RNA polymerse promoter. RNase H then degrades the initial strand of target RNA in the RNA-cDNA hybrid. The second primer then binds to the cDNA and is extended, resulting in the formation of double-stranded DNA containing the T7 RNA polymerase-binding sequence. Both strands can serve as transcription templates for the RNA polymerase that generates multiple copies of single-stranded RNA (both sense and anti-sense). These RNA product molecules reenter the cycle with subsequent formation of more double-stranded cDNAs that can serve as templates for more RNA synthesis. A billion-fold amplification of the target RNA can be achieved in less than 2 hours with this method.

Transcription-based amplification systems have several strengths including no requirement for a thermal cycler, rapid kinetics, and the single-stranded RNA product does require denaturation prior to detection. Also, single-tube clinical assays and an RNA product that is more easily degraded than DNA, may help minimize contamination risks. Limitations include the poor performance with DNA targets and concerns about the stability of complex, multi-enzyme systems. TMA-based assay for

detection of *Mycobacterium tuberculosis* (Gen-Probe), *C. trachomatis* (Gen-Probe), HCV (Gen-Probe, Bayer), and HIV-1 (Gen-Probe), are commercially available. A NASBA is commercially available (Organon-Teknika) for the detection and quantitation of HIV-1 RNA.

Strand Displacement Amplification

Strand displacement amplification (SDA) is an isothermal DNA amplification technique that consists of a target generation step followed by an exponential amplification phase that replicates the target sequence through a series of primer extension, nicking, and strand displacement steps (13, 14). This conceptually complicated amplification method was initially plagued by nonspecific amplification products. However, the use of organic solvents to increase the stringency of primer binding and heat stable enzymes have effectively addressed this problem (15). Clinical applications of SDA are being developed by Becton Dickinson and include assays for *M. tuberculosis* and *C. trachomatis*. A quantitative RT-SDA assay for HIV-1 RNA also has been described (16).

Probe Amplification Methods

Probe amplification methods differ from target amplification in that the amplification products contain only a sequence present in the initial probes. Ligase chain reaction (LCR)(17), Qbeta replicase (18) and Cleavase/Invader technology (19) are all examples of probe amplification methods with commercial potential. To date, only the ligase chain reaction has had a significant impact on the clinical microbiology laboratory.

Ligase Chain Reaction

In a standard ligase chain reaction (LCR), two oligonucleotide probes hybridize adjacent to one another on each of the denatured target DNA strands such that a “nick” is formed. A thermostable DNA ligase then seals the “nick” by joining the 3′ end of one probe and the 5′ end of the other. Each ligated product, as well as the original target, can serve as a template in subsequent rounds of denaturation, annealing, and ligation resulting in an exponential accumulation of products (17, 20).

A modification of this technique called gapped LCR (G-LCR) differs from standard LCR in that a short gap is formed after annealing of the probes to the template. The gap is filled by a thermostable DNA polymerase and the resulting nick then is ligated by the DNA ligase (21). Although LCR is convenient and easily automated, control of contamination with ligation products may prove difficult. A combination G-LCR kit for detection of both *C. trachomatis* and *N. gonorrhoeae* is commercially available (Abbott Laboratories, Chicago, IL).

Signal Amplification Methods

In signal amplification methods, the concentration of probe or target does not increase. The increased analytical sensitivity comes from increasing the concentration of label molecules attached to the target nucleic acid. Multiple enzymes, multiple probes, multiple layers of probes, and reduction of background noise have been used to enhance target detection (22). Target amplification systems generally have greater analytical sensitivity than signal amplification methods, but technological developments, particularly in branched DNA assays, have lowered the limits of detection to levels that may rival target amplification assays in some applications (23).

Signal amplification assays have several advantages over target amplification assays. In signal amplification systems, the number of target molecules is not altered and, as a result, the signal is directly proportional to the amount of target sequence present in the clinical specimen. This reduces concerns about false-positives because of cross-contamination and simplifies the development of quantitative assays. Because signal amplification systems are not dependent on enzymatic processes to amplify target sequences, they are not affected by the presence of enzyme inhibitors in clinical specimens. Consequently, less cumbersome nucleic acid extraction methods may be used. Typically, signal amplification systems employ either larger probes or more probes than target amplification systems and, consequently, are less susceptible to errors resulting from target sequence heterogeneity. Finally, RNA can be measured directly without the synthesis of a cDNA intermediate.

Branched DNA

The branched DNA (bDNA) signal amplification system is a solid phase, sandwich hybridization assay incorporating multiple sets of synthetic oligonucleotide probes (24). The key to this technology is the amplifier molecule; a bDNA molecule with 15 identical branches, each of which can bind three labeled probes.

Multiple target-specific probes are used to capture the target nucleic acid onto the surface of a microtiter well. A second set of target-specific probes also bind to the target. Preamplifier molecules bind to the second set of target probes and up to eight bDNA amplifiers. Three alkaline phosphatase-labeled probes hybridize to each branch of the amplifier. Detection of the bound labeled probes is achieved by incubating the complex with an enzyme-triggerable substrate, dioxetane, and measuring the light emission in a luminometer. The resulting signal is directly proportional to the quantity of target in the sample. The quantity of target in the sample is determined from an external standard curve.

Nonspecific hybridization of any of the amplification probes and nontarget nucleic acids leads to amplification of the background signal (noise). To reduce the hybridization potential to all nontarget the non-natural bases, isocytidine (isoC) and isoguanosine (isoG), were incorporated into the amplification probes of the third generation bDNA assays (25). IsoC and isoG bases pair with each other but not with any of the four naturally occurring bases (26). The use of isoC and isoG probes in bDNA assays increases target-specific amplification without a concomitant increase in background, thereby, greatly enhancing the detection limits. The detection limit of the third generation bDNA assay for HIV-1 RNA is 50 copies/mL. bDNA assays for the quantitation of hepatitis B virus (HBV) DNA, HCV RNA, and HIV-1 RNA are commercially available (Bayer). The System 340 platform for bDNA assays automates the incubation, washing, reading, and data processing.

Hybrid Capture

The hybrid capture system is a solution hybridization antibody capture assay that uses chemiluminescent detection. The target DNA in the specimen is denatured and hybridized with a specific RNA probe. The DNA-RNA hybrids are captured by antibodies specific for DNA-RNA hybrids that are coated onto the surface of a tube. Alkaline phosphatase-conjugated antihybrid antibodies bind to the immobilized hybrids. The bound antibody conjugate is detected with a chemiluminescent substrate and the light emitted is measured in a luminometer. The intensity of the emitted light is proportional to the amount of target DNA in the specimen. Hybrid capture assay for detection of human papillomavirus (HPV) (27) and CMV (28) in clinical specimens are commercially available (Digene).

POSTAMPLIFICATION DETECTION AND ANALYSIS

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

Gel Analysis

Visualization of amplification products in agarose gels after electrophoresis and ethidium bromide staining was the earliest detection method. After gel electrophoresis DNA may be transferred to a nitrocellulose or nylon membrane and hybridized to a specific probe to increase both the specificity and sensitivity of detection. Membranes with bound radiolabeled probes are placed in proximity to radiograph film and the hybrids are visualized as dark bands. Enzyme-labeled probes can be visualized through either light or color production after the addition of the appropriate chemiluminescent or chromogenic substrates. Many of these nonisotopic approaches are at least as sensitive as isotopic methods and are faster and the labeled probes more stable. Although gel electrophoresis and blotting remain important research tools, these techniques are being replaced by faster and simpler methods in the clinical laboratory.

Single-strand conformation polymorphism (SSCP) and restriction fragment length polymorphism (RFLP) analysis have been used to ascertain information about the base composition of the amplification products visualized in the gel. In SSCP analysis PCR product is denatured and then subjected to electrophoresis in a nondenaturing gel (29). Variations in the physical conformation of the PCR products are related to the base composition and are detected by differential gel migration. This technique has been used successfully to detect mutations causing rifampin resistance in *M. tuberculosis* (30).

RFLP analysis employs restriction endonucleases to cleave amplification products at specific recognition sites. The fragments are separated by electrophesis and the resulting banding pattern provides information about the nucleic acid sequence. When coupled with a hybridization reaction, RFLP analysis can also provide information about the location and number of loci homologous to the probe. Both SSCP and RFLP analysis of short products may soon be replaced by direct DNA sequencing as this technology improves and the costs drop.

Colorimetric Microtiter Plate Systems

Colorimetric microtiter plate (CMP) systems are convenient alternatives to traditional blotting techniques for detection of amplified products. In these systems amplified product is captured in microtiter plate wells by specific oligonucleotide probes coating the plastic surface. Bound product is detected by a color change that takes place after addition of an enzyme conjugate and the appropriate substrate. These systems resemble enzyme immunoassays and use microtiter plate washers and readers commonly found in clinical laboratories. CMP systems are more practical and faster than traditional membrane hybridization techniques described above.

Several variations are commercially available. In one popular approach, biotinylated primers are used to amplify the target and the biotin containing PCR product is denatured and added to the microtiter well. After hybridization with a capture probe, bound product is detected with a streptavidin-enzyme conjugate and a chromogenic substate (31). Enzyme conjugated-antibodies directed against double-stranded DNA also have been used detect PCR product in another CMP system (32). Another approach uses digoxigenin-dUTP to label PCR product and enzyme-conjugated antidigoxigenin antibodies to detect the captured product (33).

Automated Detection Systems

Target and probe amplification assays consist of three major steps: specimen processing, nucleic acid amplification, and product detection. The amplification and detection steps have proven the easiest to automate. Unfortunately, sample processing usually is the most labor-intensive step and represents the biggest challenge for manufacturers of automated test systems.

The COBAS system (Roche) automates the amplification and detection steps. The Abbott LCx systems automates the detection of G-LCR products. Organon Teknika manufactures automated sample preparation and detection systems for NASBA assays. All of the major manufacturers are developing high-throughput, fully automated, nucleic-acid analyzers. One example, the TIGRIS system (Gen-Probe), can reportedly process 500 nucleic acid detection tests in 8 hours.

Real Time PCR

Real-time PCR describes methods by which the target amplification and detection steps occur simultaneously in the same tube. These methods require special thermal cyclers with precision optics that can monitor the fluorescence emission from the sample wells. The computer software supporting the thermal cycler monitors the data throughout the PCR at every cycle and generates an amplification plot for each reaction.

In its simplest format, PCR product is detected as it is produced using fluorescent dyes that preferentially bind to double-stranded DNA. SYBR Green I is one such dye that has been used in this application (34). In the unbound state, the fluorescence is relatively low, but when bound to double-stranded DNA the fluorescence is greatly enhanced. The dye will bind to both the specific and nonspecific PCR products. The specificity of the detection can be improved through melting curve analysis. The specific amplified product will have a characteristic melting peak at its predicted melting temperature (T_m) whereas the primer dimers and other nonspecific products should have different T_m or give broader peaks (35).

The specificity of real-time PCR also can be increased by including hybridization probes in the reactions mixture. These probes are labeled with fluorescent dyes or with combinations of fluorescent and a quencher dyes. In the 5′ nuclease PCR assay (Taqman), the 5′ to 3′ exonuclease activity of Taq DNA polymerase is used to cleave a nonextendable hybridization probe during the primer extension phase of PCR (36). This approach uses dual-labeled fluorogenic hybridization probes. One fluorescent dye serves as reporter and its emission spectra is quenched by the second fluorescent dye. The nuclease degradation of the hybridization probe releases the reporter dye resulting in an increase in its peak fluorescent emission. The increase in fluorescent emission indicates that specific PCR product has been made and the intensity of fluorescence is related to the amount of product (37).

Fluorescence resonance energy transfer (FRET) is the basis of another approach to real-time PCR (38). This method requires two specially designed sequence-specific oligonucleotide probes. These hybridization probes are designed to hybridize next to each other on the product molecule. The 3′ end of one probe is labeled with a donor dye and the 5′ end of the other probe is labeled with an acceptor dye. The donor dye is excited by an external light source and instead of emitting light, transfers its energy to the acceptor dye by a process called FRET. The excited acceptor dye emits light at a longer wavelength than the unbound donor dye, and the intensity of the acceptor dye light emission is proportional to the amount of PCR product.

Real-time detection and quantitation of PCR product also can be accomplished using molecular beacons (39). Molecular beacons are hairpin-shaped oligonucleotide probes with an internally quenched fluorophore whose flurorescence is restored when they bind to a target nucleic acid. They are designed in such a way that the loop portion of the probe molecule is complementary to the target sequence. The stem is formed by the annealing of complementary arm sequences on the ends of the probe. A fluorescent dye is attached to one end of one arm and a quenching molecule is attached to the end of the other arm. The stem keeps the fluorophore and quencher in close proximity such that no light emission occurs. When the probe encounters a target molecule, it forms a hybrid that is longer and more stable than the stem, and undergoes a conformational change that forces the stem apart and causes the fluorophore and quencher to move away from each other, restoring the fluorescence.

Real-time PCR methods decrease the time required to perform nucleic acid assays because there are no post-PCR processing steps. Also, because amplification and detection occur in the same closed tube, these methods eliminate the post-amplification manipulations that can lead to laboratory contamination with amplicon. In addition, real-time PCR methods lend themselves well to quantitative applications because analysis is performed early in the log phase of product accumulation.

Direct Sequencing

The combination PCR and dideoxynucleotide chain termination methods can be used to determine DNA sequence in clinical samples (40). There are two platforms currently used for sequencing PCR products: electrophoretic separation on polyacrylamide slab gels or glass capillaries, and DNA microarrays. Fluorescent dye terminator chemistry and laser scanning in a polyacrylamide gel electrophoresis format has been the standard in electrophoretic separation technology. However, the recent application of capillary electrophoresis techniques to the separation of PCR and dideoxy chain termination products has streamlined the sequencing process by eliminating some of the labor intensive steps (41).

Although direct sequencing of PCR products using electrophoretic technology is a powerful research tool, its routine use in the clinical laboratory depends upon development of high throughput systems with integrated data bases and data analysis software. One such system for HIV-1 drug resistance genotyping is currently in clinical trials (Visible Genetics).

DNA MICROARRAYS

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

DNA microarrays are produced by attaching or synthesizing hundreds or thousands of oligonucleotides to a solid support. A labeled amplification product is hybridized to the probes and hybridization signals are mapped to various positions within the array. If number of probes is sufficiently large, the sequence of the PCR product can be deduced from the pattern of hybridization with the aid of computer software. A number of manufacturers are developing DNA microarrays and the instrumentation required to acquire and analyze the data.

One of the most developed approaches marries advances in synthetic nucleic acid chemistry with photolithography, a process used in the manufacture of semiconductors for the computer industry. This approach uses light to direct the synthesis of short oligonucleotides on a silica wafer (42). On a 15-mm square chip, thousands of individual sites or features can be established. At each feature, specific oligonucleotides are assembled, one nucleotide at a time, by light-activated chemistry.

The DNA chip is incubated in a flow cell with DNA product that has been fragmented and labeled with a fluorophore. After hybridization, a scanning laser confocal microscope evaluates the surface fluorescence intensity of the chip. Automated scanning by the microscope requires only a few minutes to acquire an image of the entire surface of the chip, and the computer software analyzes the fluorescent image and determines the nucleic acid sequence of the PCR product. A DNA chip based on this technology for the detection of HIV-1 drug resistance mutations is commercially available (Affymetrix).

DNA microarrays hold much promise for molecular diagnostics, however, the current technology has several limitations including the complexity of fabricating the microarrays, limited availability, and high test cost.

CLINICAL APPLICATIONS

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

Molecular methods have created new opportunities to the clinical microbiology laboratory to impact patient care in the areas of initial diagnosis, disease prognosis, and monitoring response to therapy. However, in most cases the new molecular tests supplement rather than replace the conventional diagnostic tests and add significant expense to laboratory budget. As a consequence, only those clinical applications that add true value to the overall diagnostic strategy will stand the test of time.

Detection of Slowly Growing and Unculturable Pathogens

With the development of molecular methods, the clinical microbiology laboratory is no longer solely reliant on the traditional culture methods for detection of pathogens in clinical specimens. Culture-based methods have long been the “gold standard” for infectious disease diagnosis but, in several diseases, nucleic acid based tests have replaced culture as the gold standard. Hepatitis C, pertussis, and genital infections with *C. trachomatis* are some examples of diseases in which nucleic acid-based tests are the new gold standards for diagnosis. This technology has been used to its best advantage in situations where the traditional methods are slow, insensitive, expensive, or not available. These techniques work particularly well with fragile or fastidious microorganisms that may die in transit or be overgrown by contaminating flora when cultured. It is beyond the scope of this chapter to review all of the possible applications or to provide a compendium of methods for detection of various pathogens. The reader is directed to other excellent resources for this information (43, 44).

Opportunities to actually replace culture for bacterial pathogens in routine practice are limited by the need to isolate the organisms for antibiotic susceptibility testing. In those applications where culture actually has been replaced by nucleic acid testing the pathogens are of predictable susceptibility and consequently, routine susceptibility testing is not performed e.g., *N. gonorrhoeae* and *C. trachomatis*.

Molecular methods have had the biggest impact in clinical virology where the molecular approaches are often faster, more sensitive, and more cost-effective than the traditional approaches. The diagnosis of enteroviral meningitis, herpes simplex virus encephalitis, and diagnosis of CMV infections in immunocompromised patients are examples of clinical relevant and cost effective applications of nucleic acid based tests. There are greater opportunities to replace the conventional methods in virology than in bacteriology because the culture-based methods are costly and antiviral susceptibility testing is not routinely performed. In those situations where antiviral susceptibility testing is required, it is amenable to molecular approaches.

Perhaps the greatest impact of molecular methods has been in the discovery of previously unrecognized or unculturable pathogens. During the past 15 years, a number of infectious agents were first identified directly from clinical material using molecular methods (Table 56.2). HCV, the principal etiologic agent of what was once known and non-A, non-B hepatitis, was discovered in 1989 thorough the application of molecular cloning techniques by investigators from the Centers for Disease Control and Prevention and the Chiron corporation (45). Cloning and analysis of the HCV genome led to production of viral antigens that now serve as the basis of the specific serological tests used to screen the blood supply and for diagnosis of hepatitis C. Currently, RT-PCR is used to detect, quantitate, and genotype HCV in infected individuals. To date, HCV has resisted all attempts at *in vitro* cultivation.

TABLE 56.2. EXAMPLES OF HUMAN PATHOGENS FIRST IDENTIFIED FROM CLINICAL SPECIMENS USING MOLECULAR APPROACHES[a]

Disease	Pathogen	Reference
Non-A, non-B hepatitis	Hepatitis C virus	(67)
Bacillary angiomatosis	*Bartonella henselae*	(68)
Whipple's disease	*Tropheryma whippelii*	(46)
Hantavirus pulmonary syndrome	Sin nombre virus	(69)
Kaposi's sarcoma	Human herpes virus 8	(70)

[a] Adapted form reference (47)

Tropheryma whippelii, the causitive agent of Whipple's disease, is another example of an unculturable microorganism which was initially identified by molecular methods (46). It was discovered by the use of broad-range PCR, in which primers directed against conserved sequences in the bacterial 16S ribosomal RNA gene. Sequence analysis of the PCR product and comparison with known 16S ribosomal RNA gene sequences were used to characterize the organism and establish its disease association. This approach provides a new paradigm for discovery of unrecognized pathogens and will be of value in other diseases with features that suggest an infectious etiology (47).

HCV and *T. whippelii* are two of the best examples of an important new principle in infectious diseases. A etiologic agent can be detected and characterized, and its disease association established long before its existence can be substantiated by traditional culture and serologic methods.

Disease Prognosis

Molecular techniques have created opportunities for the laboratory to provide important information that may predict disease progression. Probably the best example is HIV-1 viral load as a predictor of progression to acquired immune deficiency syndrome (AIDS) and death in infected individuals. This was first demonstrated in 1996 as part of a multicenter AIDS cohort study (48). These investigators showed that the risk of progression to AIDS and death was directly related to the magnitude of the plasma viral load at study entry. Plasma viral load was a better predictor of disease progression than the number of CD4+ lymphocytes. Subsequent studies have confirmed that baseline viral load critically influences disease progression.

Subtyping of certain viruses by molecular methods also may have prognostic value. Subtyping respiratory syncytial viruses may provide information about the severity of infection in hospitalized infants, with those infected with group A viruses having poorer outcomes (49). HPV causes dysplasia, intraepithelial neoplasia, and carcinoma of the cervix in women. HPV type 16 and 18 are associated with high risk of progression to neoplasia while types 6 and 11 are associated with a low risk of progression (50). Different genotypes of HCV have their own distinct global distributions and associations with progression to chronic infections and hepatocellular carcinoma (51).

Response to Therapy

Molecular methods have been developed to detect the genes responsible for resistance to single antibiotics or classes of antibiotics in bacteria and in many cases are superior to the phenotypic, growth-based methods. The detection of methicillin resistance in staphylococci, vancomycin resistance in enterococci, and rifampin resistance in *M. tuberculosis* are examples

where molecular methods are used to supplement the growth based methods (52). However, it is difficult to imagine, given our current state of knowledge of molecular genetics of antimicrobial resistance and the technological limitations, a genotypic approach to routine antimicrobial susceptibility testing of bacteria that would rival the phenotypic methods in terms of information content and cost.

Molecular techniques are playing an increasing role in predicting and monitoring patient response to antiviral therapy. The laboratory may have a role in predicting response to therapy by detection of specific drug resistance mutations, determining viral load, and by genotyping.

Both viral load and genotype are independent predictors of response to combination therapy with interferon and ribavirin in chronic HCV infections (53). Those patients with pretreatment viral loads of >2 million copies/mL or with genotype 1 infections have poor sustained response rates. Patients with type 1 infections require a full 48 weeks of combination therapy for maximum response whereas those patients infected with other genotypes do not benefit from therapy beyond 24 weeks duration. Therefore a simple genotype determination could spare those infected with genotypes other than 1 the discomfort and expense associated with an additional 24 weeks of therapy.

Drug resistance mutations in HIV-1 RT and protease genes lead to lower sensitivity to antiretroviral agents and are an important cause of drug failure (54). In a recent randomized clinical trial, genotypic resistance testing was found to have a significant benefit on virological response when choosing therapeutic alternatives (55). Although more data are need to document the clinical benefit of assays for detection of drug resistance mutations, these assays are likely to be used routinely to predict response to antiretroviral drugs.

Quantitative tests for HIV-1 RNA are the standard of practice for guiding clinicians in initiating, monitoring, and changing antiretroviral therapy. Several commercially available HIV-1 viral load assays have been developed and guidelines for their use in clinical practice have been published (56). Viral load assays have also been used in monitoring response to therapy in patients chronically infected with HBV and HCV (57, 58).

SPECIAL CONCERNS

Part of "56 - Molecular Techniques for Diagnosis of Infectious Diseases"

The unparalleled analytical sensitivity of nucleic acid amplification techniques coupled with their susceptibility to cross contamination present unique challenges to the routine application of these techniques in the clinical laboratory. There are special concerns in the areas of specimen processing, workflow, quality assurance, and interpretation of test results. Additional information can be found in the National Committee for Clinical Laboratory Standards (NCCLS) document MM3-A, Molecular Diagnostic Methods for Infectious Diseases; Approved Guideline (59).

Specimens

Specimen collection and processing are as critical for molecular microbiology tests as they are for culture-based tests. However, techniques designed to preserve and recover viable organisms may not be optimal for nucleic acids. Specimen processing for nucleic acid amplification tests should efficiently recover target nucleic acid from the specimen, ensure its integrity, and remove amplification inhibitors. Specimen processing also should concentrate the nucleic acid into a small volume of fluid that is compatible with the amplification reaction and address any biosafety concerns. Most specimen processing protocols for nucleic acid amplification tests require that the target DNA or RNA be extracted and partially purified. A wide variety of conventional methods and commercial kits are employed. Specific recommendations for processing vary with the target nucleic acid, the specimen, and the test method.

Prevention of Cross-Contamination

All of the target and probe amplification methods are subject to false-positives because of cross contamination with amplicon and target nucleic acid. Special laboratory design, practices, and workflow are necessary to reduce the possibility of cross contamination. The molecular diagnostic laboratory must be able to accommodate four basic functions: reagent preparation, sample processing, reaction set-up, and product detection. Ideally, these functions should be done in four physically separated areas. However, because of space constraints, separate workstations for these functions within the same room are often used. Regardless of the degree of physical separation, each workstation should have dedicated equipment and supplies. Additional containment measures include controlled airflow, vestibule entrances to the rooms, and the use of dead-air boxes at the workstations. Nucleic acid amplification tests should not be performed in areas of the laboratory used for cultivation of the target organisms because of concern about cross contamination with the target. The need for strict physical isolation of the workstations is less for those applications that include amplicon inactivation methods or use closed systems.

Unidirectional workflow is another means to reduce the possibility of cross-contamination. The workflow is should always be from those areas free of contaminating nucleic acids to those areas that are likely to be contaminated, proceeding from the reagent preparation area to the amplification area, and then finally to product detection area. Flow of personnel in the opposite direction should be discouraged. In addition to unidirectional workflow, a number of other routine preventive measures have been described including the use of aerosol barrier pipette tips, single use reagents, diluted bleach to wipe down counters, and low copy number positive controls (60, 61).

One of the most effective measures for controlling false-positive PCRs from amplicon cross-contamination is use of dUTP and uracil *N*-glycosylase (UNG) to modify the PCR product and selectively destroy any carried-over product from a previous reaction, respectively (62). With this amplicon inactivation method, dUTP replace dTTP in the reaction mixture and results in amplicon with deoxyuracil residues in thymidine resides. This unnatural DNA is chemically distinct from the target DNA and the UNG present in the reaction mixture cleaves any carried over product DNA at the uracil residues destroying its template activity. Target DNA is unaffected by UNG and new product can be synthesized because the UNG is thermally denatured at the elevated temperature used to denature the target and initiate the

PCR (63). Other methods include of amplicon inactivation include, ultraviolet (UV) light irradiation, use of photochemical cross-linkers, and hydroxlamine treatment (64, 65).

Quality Assurance

The number of nucleic acid amplification tests cleared by the FDA for diagnostic use is relatively small and limited to a few pathogens (Table 56.3). Tests for herpes simplex virus (HSV) DNA in cerebrospinal fluid (CSF) of patients suspected of having HSV encephalitis or enteroviral RNA in CSF of children with aseptic meningitis, and for HCV RNA in serum of patients with hepatitis are some examples of tests that are used routinely for diagnostic purposes, but for which there are no kits cleared by the Food and Drug Administration (FDA). Laboratories that develop assays "in house" must take responsibility for the appropriate validation and verification of the test performance. These activities are more demanding than described in CLIA'88 for FDA-cleared kits and are outlined in the NCCLS document 3A (59).

TABLE 56.3. NUCLEIC ACID AMPLIFICATION TESTS CLEARED FOR *IN VITRO* DIAGNOSTIC USE BY THE FDA

Test	Method (Manufacturer)
Chlamydia trachomatis	PCR (Roche), LCR (Abbott), TMA (Gen-Probe)
Neisseria gonorrhoeae	PCR (Roche), LCR (Abbott)
Mycobacterium tuberculosis complex	TMA (Gen-Probe), (PCR Roche)
Human immunodeficiency virus type 1	Quantitative RT-PCR (Roche)
Cytomegalovirus	HCA (Digene)

FDA, Food and Drug Administration

Proficiency testing is extremely important in molecular diagnostic laboratories because of the number of nonstandardized tests that are in routine use. Unfortunately, organized proficiency testing programs in this area are not well developed. The College of American Pathologists has offered an ungraded molecular microbiology survey over the past several years. The Centers for Disease Control and Prevention has on-going proficiency testing programs in HIV-1 viral load and *M. tuberculosis* testing. These programs and published reports of investigator-initiated proficiency testing studies all demonstrate that increased standardization is needed and that the laboratory performing the test is relatively more important than the test methodology. In the absence of an organized proficiency-testing program, the laboratory can design its own internal program or arrange to send samples to laboratories using similar methods.

Interpretation of Results

Molecular diagnostic tests often are more sensitive than culture methods that traditionally have been considered the gold standards for diagnosis. In these situations, the sensitivity and specificity of the molecular tests is defined in clinical rather than laboratory terms. Also, because molecular methods bring new capabilities to the clinical microbiology laboratory, they can lead to the definition of new disease states. Interpretation is even more problematic in this situation because the molecular methods may have no acceptable standard to use in a comparative evaluation.

The simple detection of an organism's DNA in clinical specimen does not guarantee its viability or its association with a disease process. Dead organisms may contain amplifiable DNA and, as a result, nucleic-acid–based tests are not useful as tests of cure for bacterial infections. Detection of an organism's nucleic acid sequence in a normally nonsterile body site may indicate colonization, contamination, or infection.

Interpretation of molecular diagnostic test results must take into consideration the factors that can lead to both false-positive and false-negative results. Assays using carry-over prevention formats and multiple negative controls decrease the likelihood that false positive results will be reported. Assays with internal positive controls increase the likelihood of detecting false-negatives due to sample inhibition and lead to increased confidence in negative results.

Culture methods usually are designed to recover all of the common pathogens that are likely to be found in particular specimen, whereas, molecular methods usually are designed for a single agent. Panels of single-target molecular tests and multiplexed reactions often are required to provide the same broad range of pathogen detection offered by culture.

REFERENCES

1. Arnold Jr LJ, Hammond PW, Wiese WA, et al. Assay formats involving acridinium ester-labeled DNA probes. *Clin Chem* 1989;35:1588.

2. Tenover FC. Diagnostic deoxyribonucleic acid probes for infectious diseases. *Clin Microbiol Rev* 1988;1:82.

3. Hankin RC. In situ hybridization: principles and applications. *Lab Med* 1992;23:764.

4. Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988;239:487.

5. Myers TW, Gelfand DH. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* 1991;30:7661.

6. Haqqi TM, Sarkar G, David CS, et al. Specific amplification with PCR of a refractory segment of genomic DNA. *Nucleic Acids Res* 1988;16:11844.

7. Chamberlain JS, Gibbs RA, Rainer JE, et al. Delection screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res* 1988;16:1141.

8. Clementi M, Menzo S, Bagnarelli P, et al. Quantitative PCR and RT-PCR in Virology. *PCR Methods and Applic* 1993;2:191.

9. Haberhausen G, Pinsl J, Kuhn C-C, et al. Comparative study of different standardization concepts in quantitative competitive reverse transcription-PCR assays. *J Clin Microbiol* 1998;36:628.

10. Compton J. Nucleic acid sequence-based amplification. *Nature* 1991;350:91.

11. Kwoh DY, David GR, Whitfield KM, et al. Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. *Proc Natl Acad Sci USA* 1989;86:1173.

12. Guatelli JC, Whitfield KM, Kwoh DY, et al. Isothermal in vitro amplification of nucleic acids by multienzyme reaction modeled after retroviral replication. *Proc Natl Acad Sci USA* 1990;87:1874.

13. Walker GT, Fraiser MS, Schram JL, et al. Strand displacement amplification - an isothermal, in vitro DNA amplification technique. *Nucleic Acids Res* 1992;20:1691.

14. Walker GT, Little MC, Nadeau JG, et al. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc Natl Acad Sci USA* 1992;89:392.

15. Spargo CA, Fraiser MS, Van Cleve M, et al. Detection of M. tuberculosis DNA using thermophilic strand displacement amplification. *Molec Cellular Probes* 1996;10:247.

16. Nycz CM, Dean CH, Haaland PD, et al. Quantitative reverse transcription strand displacement amplification; quantitation of nucleic acids using an isothermal amplification technique. *Anal Biochem* 1998;259:226.

17. Wu DY, Wallace RB. The ligation amplification reaction(LAR) - amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* 1989;4:560.

18. Kramer FR, Lizardi PM. Replicatable RNA reports. *Nature* 1989;339:401.

19. Lyamichev V, Mast A, Hall JG, et al. Polymorphism identification and quatitative detection from genomic DNA by invasive cleavage of oligonnucleotide probes. *Nature Biotech* 1999;17:292–296.

20. Barany F. Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc Natl Acad Sci USA* 1991;88:189.

21. Birkenmeyer LG, Mushahwar IK. DNA probe amplification methods. *J Virol Meth* 1991;35:117.

22. Kricka LJ. Nucleic acid detection technologies — labels, strategies, and formats. *Clin Chem* 1999;45:453.

23. Kern D, Collins M, Fultz T, et al. An enhanced-sensitivity branched-DNA assay for quantification of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol* 1996;34:3196.

24. Nolte FS. Branched DNA signal amplification for direct quantitation of nucleic acid sequences in clinical specimens. *Advan Clin Chem* 1999;33:201.

25. Collins ML, Zayati C, Detmer JJ, et al. Preparation and characterization of RNA standards for use in quantitative branched-DNA hybridization assays. *Analy Biochem* 1995;226:120.

26. Piccirilli JA, Krauch T, Moroney SE, et al. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* 1990;343:33.

27. Cope JJ, Hildesheim A, Schiffman MH, et al. Comparison of the hybrid capture tube test and PCR for detection of human papillomavirus DNA in cervical specimens. *J Clin Microbiol* 1997;35:2262.

28. Mazzulli T, Drew LW, Yen-Lieberman B, et al. Multicenter comparison of the digene hybride capture CMV DNA assay (version 2.0), the pp65 antigenemia assay, and cell culture for detection of cytomeglovirus viremia. *J Clin Microbiol* 1999;37:958.

29. Orita M, Iwahana H, Kanazawa H, et al. Detection of polymorphism of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA* 1989;86:2766.

30. Telenti A, Imboden P, Marchesi F, et al. Direct, automated detection of rifampin-resistant *Mycobacterium tuberculosis* by polymerase chain reaction and single-strand conformation polymorphism analysis. *Antimicrob Agents Chemother* 1993;37:2054.

31. Loeffelholz MJ, Lewinski CA, Silver SR, et al. Detection of *Chlamydia trachomatis* in endocervical specimens by polymerase cain reaction. *J Clin Microbiol* 1992;30:2847.

32. Mantero G, Zonaro A, Albertini A, et al. DNA enzyme immunoassay: general method for detecting products of polymerase chain reaction. *Clin Chem* 1991;37:422.

33. Poljak M, Seme K. Rapid detection and typing of human papillomaviruses by consensus polymerase chain reaction and enzyme-lined immunosorbent assay. *J Virol Methods* 1996;56:231.

34. Morrison T, Weiss JJ, Wittwer CT. Quantification of low copy transcrips by continuous STBR green I dye monitoring during amplification. *Bio Techniques* 1998;24:954.

35. Ririe K, Rasmussen RP, Wittwer CT. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal Biochem* 1997;245:154.

36. Holland PM, Abramson RD, Watson R, et al. Detection of specific polymerase chain reaction product by utilizing the 5′ → 3′ exonuclease activity of *thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci USA* 1991;88:7276.

37. Heid C, Stevens J, Livak KJ, et al. Real time quantitative PCR. *Genome Res* 1996;6:986.

38. Lay MJ, Wittwer CT. Real-time fluorescence genotyping of factor V leiden during rapid cycle PCR. *Clin Chem* 1997;43:2262.

39. Tyagi S, Bratu DP, Kramer FR. Multicolor molecular beacons for allele discrimination. *Nature Biotech* 1998;16:49.

40. Innis MA, Myambo KB, Gelfand DH, et al. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc Natl Acad Sci USA* 1988;85:9436.

41. Felmlee TA, Oda RP, Persing DA, et al. Capillary electrophoresis of DNA potential utility for clinical diagnoses. *J Chromatogr* 1995;A717:127.

42. Pease AC, Solas D, Sullivan EJ, et al. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994;91:5022.

43. Persing DH, Smith TF, Tenover FC, et al. eds. *Diagnostic Molecular Microbiology: Principles and Applications*. Washington, D.C. ASM Press, 1993.

44. Persing DH. *PCR Protcols for Emerging Infectious Diseases*. Washington, D.C. ASM Press 1996.

45. Choo QL, Kuo G, Weiner AJ, et al. Isolation of a cDNA clone derived from a blood-borne non-A non-B viral hepatitis genome. *Science* 1989;224:359.

46. Relman DA, Schmidt TM, MacDermott RP, et al. Identification of the uncultured bacillus of Whipple's disease. *N Engl J Med* 1992;327:293.

47. Relman DA. The search of unrecognized pathogens. *Science* 1999;284:129.

48. Mellors JW, Rinaldo Jr. CR, Gupta P, et al. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* 1996;272:1167.

49. Walsh EE, McConnochie KM, Long CE, et al. Severity of respiratory syncytial virus infection is related to virus strain. *J Infect Dis* 1997;175:814.

50. Reid R, Greenberg M, Jensen AB, et al. Sexually transmitted papillomaviral infections. I. The anatomic distribution and pathologic grade of neoplastic lesions associated with different viral types. *Am J Obstet Gynecol* 1987;156:212.

51. Zein NN, Rakela J, Krawitt EL, et al. Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity and response to interferon therapy. *Ann Intern Med* 1996;125:634.

52. Tenover FC. Genotypic detection of antimicrobial resistance. In: Persing DH, ed. PCR Protocols for Emerging Infectious Diseases. Washington, D.C., ASM Press, 1996.

53. McHutchison JG, Gordon SC, Schiff ER, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. *N Engl J Med* 1998;339:1485.

54. Hirsch MS, Conway B, D'Aquila RT, et al. Antiretroviral drug resistance testing in adults with HIV infection: implications for clinical management. *JAMA* 1998;279:1984.

55. Durant J, Clevenberg P, Halfon P, et al. Drug-resistance genotypying in HIV-1 therapy: the VARADAPT randomized controlled trial. *Lancet* 1999;353:2195.

56. Saag MS, Holodniy M, Kuritzkes DR, et al. HIV viral load markers in clinical practice. *Nature Medicine* 1996;2:625.

57. Hendricks DA, Stowe BJ, Hoo BS, et al. Quantitation of HBV DNA in human serum using a branched DNA (bDNA) signal amplification assay. *Am J Clin Pathol* 1995;104:537.

58. Schiff ER, De Medina M, Kahn RS. New perspectives in the diagnosis of hepatitis C. *Seminars Liver Dis* 1999;19:3.

59. NCCLS. Molecular diagnostic methods for infectious diseases; approved guideline, NCCLS document MM3-A (SBN 1-56238-288-8). Wayne, PA, 1995.

60. Kwok S, Higuchi R. Avoiding false positives with PCR. *Nature* 1989;339:237.

61. Ehrlich HA. Caveats of polymerase chain reaction. *Clin Microbiol Newsl* 1991;13:149.

62. Rys PN, Persing DH. Preventing false positives: quantitative evaluation of three protocols for inactivation of polymerase chain reaction amplification products. *J Clin Microbiol* 1993;31:2356.

63. Longo MC, Berninger MS, Hartley JL. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. *Gene* 1990;93:125.

64. Aslanzadeh J. Application of hydroxylamine hydrochloride for post-PCR sterilization. *Mol Cell Probes* 1993;7:145.

65. Cimino GD, Gamper HB, Isaacs ST, et al. Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu Rev Biochem* 1989;54:1151.

66. Tang YW, Persing DH. Molecular detection and identification of microorganisms. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of Clinical Microbiology*. 7th ed. Washington, D.C., ASM Press, 1999.

67. Choo Q-L, Kuo G, Weiner AJ, et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989;244:359.

68. Relman DA, Louitt JS, Schmidt TM, et al. The agent of bacillary angiomatosis: an approach to the identificaiton of uncultured pathogens. *N Engl J Med* 1990;323:1573.

69. Nichols ST, Spiropoulou CF, Morzunov S, et al. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 1993;262:914.

70. Chang Y, Cesarman E, Pessin MS, et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 1994;266:1865.

57

Role of the Clinical Microbiology Laboratory in Hospital Epidemiology and Infection Control

Daniel J. Diekema

Michael A. Pfaller

The clinical microbiology laboratory is responsible for identifying pathogenic microorganisms in clinical specimens as accurately and rapidly as possible, thereby assisting clinicians in the timely diagnosis and treatment of infected patients. Hospital epidemiologists and infection control practitioners are charged with monitoring, preventing, and controlling the spread of infections in the hospital environment. Because preventing hospital-acquired (nosocomial) infections first requires the ability to detect infections when they occur, the clinical microbiology laboratory is inextricably linked to the infection control program at any hospital. In this chapter, we will discuss the impact of nosocomial infections, outline the organization of the hospital infection control program, and describe the important role of the clinical microbiology laboratory in the prevention and control of nosocomial infection.

NOSOCOMIAL INFECTION

Part of "57 - Role of the Clinical Microbiology Laboratory in Hospital Epidemiology and Infection Control"

Definition

A nosocomial infection is one that is acquired in a hospital or healthcare facility (i.e., the infection was not present or incubating at the time of admission). Although each individual infection must be considered in light of its usual incubation period, for most bacterial infections an onset of symptoms more than 48 hours after admission is evidence for nosocomial acquisition. An infection acquired in the hospital or healthcare facility may not be recognized until after discharge. These infections are still considered to be nosocomial. This is especially true for surgical wound infections. Because hospital stays are shorter now than ever before, infection control practitioners must devise strategies for postdischarge surveillance in order to accurately monitor nosocomial infection rates.

Infection Rates and Predominant Pathogens

Approximately 5% to 10% of inpatients acquire an infection during hospitalization. The urinary tract is the most commonly involved site, comprising 35% to 40% of all nosocomial infections. Surgical wound infections are the next most frequent nosocomial infection (20%), followed by lower respiratory tract (15%) and bloodstream (5% to 10%) infections (Fig. 57.1).

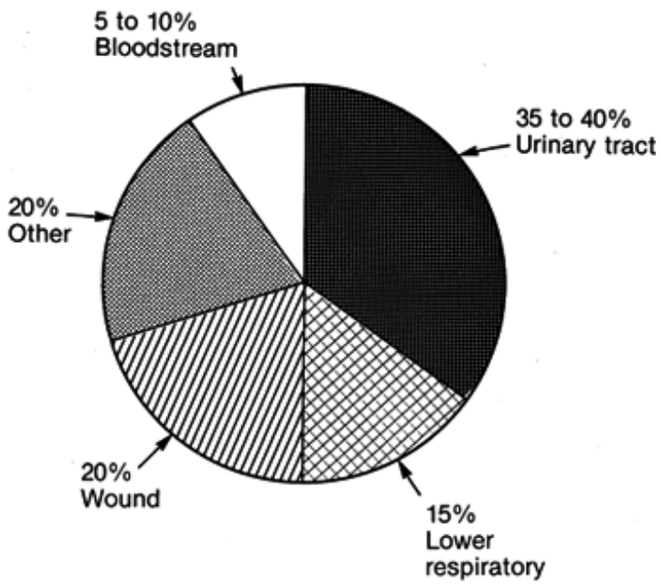


FIGURE 57.1. Nosocomial infections by site.

Table 57.1 lists the five most common bacterial pathogens causing various nosocomial infections in U.S. hospitals surveyed in the Centers for Disease Control and Prevention's (CDC) National Nosocomial Surveillance Survey (NNIS) (1, 2). The past two decades have seen a gradual shift in spectrum of nosocomial pathogens from Gram negative to Gram positive organisms, and the emergence of *Candida* spp. as a major problem (3). An important factor associated with the emergence of the staphylococci and enterococci as nosocomial pathogens has been their increasing resistance to most commonly used antimicrobial agents (4) (Table 57.2).

Morbidity, Mortality, and Cost

Nosocomial infections are estimated to be responsible for thousands of deaths annually, and likely contribute to thousands more (5, 6). Because patients with the most severe underlying illness also are those most vulnerable to nosocomial infection, it is very difficult to estimate the proportion of crude or overall mortality that is directly attributable to nosocomial infection. Studies have been performed that attempt to carefully control for as many confounding variables as possible, in order to estimate *attributable* mortality. These studies have demonstrated that the attributable mortality of nosocomial bloodstream infection was 14% for those because of coagulase-negative staphylococci (7), 31% and 37% for those because of vancomycin-susceptible and -resistant enterococci (8, 9), respectively, and 38% for infections from *Candida* spp. (10) (Table 57.3).

Nosocomial infections also cost the healthcare system billions of dollars annually, primarily through increasing hospital costs and length of stay (LOS). At the University of Iowa, we have demonstrated a median excess LOS of 8 and 30 days for nosocomial bloodstream infections resulting from coagulase-negative staphylococci and *Candida* spp., respectively (7, 10). Nosocomial

bloodstream infections in the intensive care unit have been associated with an excess LOS of 24 days and excess hospital costs of $40,000 per survivor (11). Surgical wound infections result in an excess LOS of over 7 days and an increase in hospital costs of over $3,000 per infection (12).

TABLE 57.2. EMERGING RESISTANCE AMONG *ENTEROCOCCUS* SPP. TO VANCOMYCIN AND *STAPHYLOCOCCUS AUREUS* TO METHICILLIN IN CLINICAL ISOLATES FROM U.S. HOSPITALS

Organism	Percent Resistance by Year				
	1990	1992	1994	1996	1997–1998
Enterococcus spp	<1	5.6	8	16–17	18
S. aureus	20–25	20–25	-	26	25–45

Adapted from reference 63.

TABLE 57.1. DISTRIBUTION OF THE FIVE MOST COMMON NOSOCOMIAL PATHOGENS ISOLATED FROM THE FOUR MAJOR INFECTION SITES IN THE INTENSIVE CARE UNIT, JANUARY 1986–APRIL 1997, NNIS[a]

Infection Site	Pathogen	Percent of Total at Each Infection Site
Bloodstream		
	CoNS	33.5
	S. aureus	13.4
	Enterococcus spp.	12.8
	C. albicans	5.8
	Enterobacter spp.	5.2
Pneumonia		
	P. aeruginosa	17.4
	S. aureus	17.4
	Enterobacter spp.	11.4
	K. pneumoniae	6.7
	H. influenza	4.9
Surgical Site		
	Enterococcus spp.	15.3
	CoNS	12.6
	S. aureus	11.2
	P. aeruginosa	10.3
	Enterobacter spp.	9.5
Urinary Tract		
	E. coli	19.2
	C. albicans	14.4
	Enterococcus spp.	14.1
	P. aeruginosa	11.2
	K. pneumoniae	5.8

[a] National Nosocomial Infections Surveillance System (1). CoNS, coagulase negative staphylococci

Although certain nosocomial infections are secondary to the use of devices or procedures that compromise normal host barriers to infection, many nosocomial infections are preventable. The Study of the Efficacy of Nosocomial Infection Control (SENIC) indicated that the presence of an active surveillance and infection control system was associated with a 32% decrease in nosocomial infection rates, while the absence of such a program was associated with an 18% increase in nosocomial infections (13). Clearly, the activities of an effective infection control program save hospitals money and, more importantly, improve patient care.

THE HOSPITAL INFECTION CONTROL PROGRAM

Part of "57 - Role of the Clinical Microbiology Laboratory in Hospital Epidemiology and Infection Control"

The hospital infection control program is generally directed by a physician-epidemiologist and enforced by the infection control committee. The program should include surveillance of nosocomial infections, continuing education of medical staff, control of infectious diseases outbreaks, protection of employees from infection, and advice on new products and procedures. The clinical microbiology laboratory has a role to play in each of these important activities.

Infection Control Committee

The infection control committee is responsible for reporting and evaluating nosocomial infection data and for drafting and implementing policies, procedures, and guidelines pertinent to the practice of infection control. The committee should be multidisciplinary, with representatives from all departments, including clinical microbiology. Every hospital also must have a working infection control staff, composed of one or more infection control practitioners supervised by the physician-hospital epidemiologist. These professionals are responsible for collecting data concerning nosocomial infections and providing it to the committee. The committee should meet every 1 to 2 months, depending upon the size of the hospital, to review hospital-specific nosocomial infection data and to formulate policy. Other responsibilities of the committee include review of technical information about new products, devices, or procedures pertinent to infection control, and institution of all necessary control measures in the event of an outbreak or other infection control emergency.

The clinical microbiologist has an important role on the infection control committee. Few, if any, of the other infection control committee members will have a background in diagnostic microbiology. The clinical microbiologist therefore provides expertise in the interpretation of culture results, advice about the appropriateness and feasibility of microbiological approaches to an infection control problem, and input regarding the laboratory resources necessary to accomplish the goals of the committee. Equally important is what the clinical microbiologist takes back to the laboratory: namely, the ability to better plan for epidemiologic investigations and surveillance efforts, and important insight into the practical problems confronting the hospital infection control program.

One of the most important contributions of the clinical microbiologist is to inform the infection control committee of the strengths and limitations of methods employed to detect and characterize nosocomial pathogens. He or she should describe the potential impact on infection control of any change in methods for detection, identification and susceptibility testing of nosocomial pathogens. The committee should also be made aware of the budgetary and personnel constraints under which the laboratory operates, to ensure that they do not expend valuable laboratory resources unless there is a clear epidemiologic indication to do so.

TABLE 57.3. ATTRIBUTABLE MORTALITY OF NOSOCOMIAL BLOODSTREAM INFECTION FROM SELECTED PATHOGENS

Organism	Mortality Among Cases (%)	Mortality Among Matched Controls (%)	Attributable Mortality (%)	Reference
CoNS[a]	31	17	14	7
Enterococcus spp.	43	12	31	8
VRE[b]	67	30	37	9
C. albicans	57	19	38	10

[a] Coagulase-negative *Staphylococcus* spp.

[b] Vancomycin-resistance *Enterococcus* spp.

Nosocomial Infection Surveillance

Although costly and time-consuming, systematic surveillance of nosocomial infections is important for many reasons. By monitoring the frequency and types of nosocomial infection, the infection control program may detect outbreaks, evaluate compliance with infection control guidelines, provide data for policy development, and monitor the impact of any infection control interventions on nosocomial infection rates. Additionally, surveillance is a requirement of national and state accrediting agencies. One of the most important aspects of surveillance is the timely feedback of infection rates to health care providers, along with suggestions for improvement and reemphasis of existing infection control practices. In any case, the existence of an active nosocomial infection surveillance program is associated with a reduction in infection rates (and their consequent morbidity and mortality) (13, 14).

Because surveillance consumes more time than any other single infection control activity (15), it is essential that the most efficient surveillance system be employed. The most complete and accurate surveillance program would employ daily chart reviews of all hospitalized patients, an approach that is obviously not practical in any but the smallest of hospitals. Limited resources should be focused in the highest risk areas (intensive care units, hematology-oncology, burn units, organ transplant wards), and various screens can be used to increase efficiency. Microbiology reports, nursing care plans (Kardex), antibiotic orders, radiology reports, temperature charts, and discharge diagnoses all can be used to determine which charts should be reviewed further.

Review of microbiology reports probably is the most common method for case finding used routinely in the performance of nosocomial infection surveillance. Review of microbiology reports alone compares favorably to more comprehensive ward-based surveillance (16, 17). Yakoe et al. (18) recently reported that review of microbiology data alone was both more resource efficient and as effective as applying the CDC/NNIS definition for detecting nosocomial bloodstream infection. Laboratory-based surveillance allows for the efficient review of large amounts of data, and medical information systems can enhance surveillance further by linking data from many sources, including pharmacy (antimicrobial use), laboratory, radiology, billing (diagnostic codes), and even nursing notes (temperature charts, care plans).

Although review of microbiology reports is an essential part of surveillance, these data alone will not detect all outbreaks or infections. The sensitivity and specificity of laboratory-based surveillance is heavily dependant upon both the frequency of culturing and the quality of the specimens received by the lab. Optimal surveillance will include a combination of the above data screens (for example, nursing care plan and microbiology reports) to help determine which charts deserve further review. The frequency of surveillance of specific hospital units should be determined in each hospital by infection control practitioners, based upon available resources, prevailing infection rates and other factors. The University of Iowa has validated a surveillance strategy using primarily microbiology reports and nursing care plans (Kardex), and found the sensitivity and specificity to be 81% and 98%, respectively (19).

Education

Communication is the cornerstone of effective infection control. Combining reminders about good infection control practice along with regular feedback on nosocomial infection rates is one effective way to educate health care providers. Infection control practitioners also require continuing education, and one area for special emphasis should be clinical microbiology. Regular laboratory rounds to review laboratory results, discuss ongoing infection control problems, and plan investigative efforts in outbreak settings serves to educate both microbiologists and infection control practitioners about each other's areas of expertise. The infectious diseases consultant(s) and any other interested clinicians also should be invited to participate in these rounds to add further clinical insight.

ROLE OF THE MICROBIOLOGY LABORATORY IN INFECTION CONTROL

Part of "57 - Role of the Clinical Microbiology Laboratory in Hospital Epidemiology and Infection Control"

Now that we have given a broad overview of the general structure and activities of the hospital infection control program, we

will focus on the most important specific roles played by the microbiology laboratory in the day to day practice of infection control. We will discuss the challenges of accurate reporting of identification and susceptibility results, the importance of timely reporting of data, and the specialized studies sometimes required in the setting of an outbreak of infection.

Accurate Identification and Susceptibility Testing of Nosocomial Pathogens

The expanding spectrum of organisms that colonize and infect seriously ill patients continues to challenge the ability of the clinical microbiology laboratory to accurately identify and characterize nosocomial pathogens. Many common nosocomial pathogens are easily detected and identified (e.g., *Staphylococcus* spp., *Enterococcus* spp., *Candida* spp., and the Enterobacteriaceae). However, many nonfermentative Gram-negative organisms (e.g., *Burkholderia* spp., *Stenotrophomonas* maltophilia, *Acinetobacter* spp.), fungi (e.g., *Aspergillus* spp., *Fusarium* spp., nonalbicans *Candida* spp.), and viruses also cause nosocomial infections and can be more difficult to identify.

Specimen Handling

The ability of the laboratory to accurately identify nosocomial pathogens requires that adequate, high-quality specimens be received from the clinical services (discussed in more detail in Chapter 50). Any inappropriate specimens (wrong transport media, leaking containers, undue delay from collection to transport) should not be processed. Specimens also should be examined microscopically whenever possible to help judge quality and assist in the interpretation of results. For example, a nasopharyngeal wash specimen for detection of respiratory syncytial virus (RSV) antigen that has no cellular material is likely to produce a falsely negative test result. Gram stain of sputum specimens also is extremely useful as a screen to determine which are contaminated with oropharyngeal secretions and therefore not likely to be of value in the diagnosis of respiratory tract infection (20). Careful monitoring of specimen quality and enforcement of strict criteria for acceptance of clinical specimens ensures that the most accurate, least misleading microbiological data are reported to the clinician and infection control practitioner. Because many nosocomial pathogens also are common colonizing organisms and culture contaminants (e.g., coagulase-negative staphylococci), specimen collection and handling can have an impact on nosocomial infection rates. In addition, improper handling and contamination of culture specimens in the laboratory can lead to pseudo-outbreaks of infection (21, 22), which expend valuable time and resources.

Rapid Diagnostic Testing

The past decade has seen the increasing development of rapid diagnostic testing using molecular or immunologic methods, independent of culture. The speed with which these methods produce accurate results can have important implications for infection control. For example, a variety of methods now are available for rapid detection of RSV in respiratory secretions (23), *Clostridium difficile* in stool (24), *Mycobacterium tuberculosis* in respiratory specimens that are smear-positive for acid-fast bacilli (AFB) (25), and *Legionella pneumophila* serogroup 1 in urine (26). A positive result from any of these tests may allow for more rapid institution of appropriate isolation precautions and/or early investigation of potential outbreaks. Of course, indiscriminant use or poor quality control of a rapid diagnostic test can lead to errors, including a series of falsely positive tests resulting in a pseudo-outbreak (27). Additionally, the infection control practitioner must be educated regarding the negative predictive value of some rapid tests (e.g., nucleic acid amplification for *M. tuberculosis* in AFB smear negative patients), which may not be sufficiently high to allow for discontinuation of previously instituted isolation precautions.

Another example of a rapid test with implications for infection control is a new, rapid ELISA for the detection of antibodies to the human immunodeficiency virus (HIV) (HIV, Murex, Single Use Diagnostic System (SUDS) HIV 1/2) (28). The high sensitivity and excellent negative predictive value of this test may assist the infection control practitioner in counseling a health care worker who sustains a blood or body fluid exposure to a source patient who tests negative by the rapid test. Again, however, the limitations of this rapid test should be communicated to infection control practitioners. If the source patient has recent (within the past 6 months) risk behaviors for HIV acquisition, any test for antibodies to HIV may be falsely negative because the patient may not have yet mounted a detectable immune response.

Finally, the most widely available rapid diagnostic test is the microscopic examination of Gram-stained material. This procedure provides information rapidly and may detect and presumptively identify organisms of epidemiologic importance that may not be apparent on routine culture. Thus, the microbiology laboratory should make sure that infection control personnel are aware of the results of the Gram stain, as well as culture, for epidemiologically important specimens.

Automated Identification and Susceptibility Testing

Most laboratories have the capability to identify microorganisms to species level and perform antimicrobial susceptibility testing (AST). Conventional macro- and microbroth dilution AST have been replaced in many laboratories by automated commercial systems, which are convenient, less labor intensive, more rapid, and more standardized than is currently the case with in-house methods (29, 30). In general, the performance of most of the automated identification and susceptibility testing systems is as good as the conventional manual methods. Accuracy of 90% to 95% relative to conventional methods are reported routinely for both identification and susceptibility testing (29,30,31,32 and 33). The widespread availability of these systems has provided an improved level of performance and standardization among laboratories and allows many laboratories to perform microbiologic testing that otherwise would have had to be sent out to reference laboratories.

The limitations of many of these automated AST systems, in particular those using short (3- to 5-hour) incubation periods, have become increasingly apparent in recent years. As new pathogens and new resistance mechanisms evolve, it has become necessary to supplement automated systems with additional

methodologies to guard against significant errors for some organism/antimicrobial combinations. In general terms, these are organisms that display heteroresistance to ß-lactam antibiotics, inducible resistance mechanisms, or high mutation rates to drug resistance. In many cases, these resistance mechanisms may become apparent only after an incubation period longer than the 3- to 5-hour interval used by many rapid automated systems. For example, many commercial systems are not satisfactory for detecting penicillin resistance in pneumococci (34), and may underestimate glycopeptide resistance among enterococci (35). In addition, some systems have had problems in accurately detecting oxacillin resistance among *Staphylococcus* species (36) and have not adequately detected extended-spectrum ß-lactamase production among certain Enterobacteriaceae (29, 37). Conversely, some systems have reported false resistance to imipenem and aztreonam among Gram-negative rods (38).

As these problems are brought to the attention of the manufacturers of these systems, improvements can be made in the instrumentation, panels, or software programs to improve accuracy. This process of ongoing independent evaluation of automated systems and feedback to responsive industry representatives is extremely important. Unfortunately, in the era of managed care and shrinking laboratory resources fewer laboratories can perform rigorous internal evaluations of new technology. Unrecognized problems in identification or susceptibility testing (e.g., inaccurate detection of glycopeptide-resistance among enterococci or oxacillin resistance among staphylococci) have obvious, major ramifications for infection control. Namely, serious problems and even outbreaks can go unrecognized, or conversely, infection control resources can be diverted toward spurious resistance problems.

Three of the most problematic resistances emerging as causes of nosocomial infection include extended-spectrum ß-lactamases (ESBLs) and stably derepressed Bush-Jacoby-Medeiros group 1 cephalosporinases among Enterobacteriaceae (39, 40), glycopeptide resistance among enterococci (41) and staphylococci (42, 43), and penicillin resistance among *S. pneumoniae* (44) and viridans-group streptococci (45). Effective infection control efforts obviously depend upon the ability of the laboratory to detect these epidemiologically important resistant pathogens. The laboratory director must keep up with current literature regarding the ability of automated systems to detect emerging resistances, and the need for any additional methods to detect or confirm resistance patterns. The Etest, a stable antimicrobial gradient method, has performed satisfactorily for susceptibility testing of fastidious species and for many other organism/drug combinations that have proven problematic for some automated systems (33, 46–48).

Reporting of Laboratory Data

Results of cultures and antimicrobial susceptibility testing are an important data source for infection control and usually are reviewed daily by infection control practitioners. Additional important information can be obtained via direct communication between laboratory and infection control personnel during regular laboratory rounds (e.g., issues of infection versus colonization, or the extent to which specimens should be worked up for epidemiologically important organisms).

Additionally, some culture results warrant an early phone call from the laboratory to the infection control practitioner to ensure that appropriate control measures are implemented. Examples include positive blood or normally sterile site cultures for *Neisseria meningitidis*, smears, or cultures positive for AFB, isolation of the enteric pathogens *Salmonella* or *Shigella*, and the isolation of certain antimicrobial resistant pathogens such as methicillin-resistant *S. aureus*. Additionally, the detection of new or unusual pathogens (e.g., *Legionella* spp., vancomycin-resistant Gram-positive organisms) also should be reported promptly to the infection-control practitioner.

Routine microbiology laboratory results should be readily accessible to laboratory personnel, clinicians, and infection-control practitioners. In most cases, results are stored in a computer database, facilitating retrieval and analysis. Information stored should include specimen type, date of collection, patient identification, hospital number, hospital service, ward location, organisms identified, antimicrobial susceptibility test results, and the results of any specialized testing performed (e.g., typing). This information can be of great value in establishing a baseline for nosocomial infection rates for the analysis of trends or outbreaks.

Both clinicians and infection-control workers may benefit from periodic summaries of selected microbiology results. These results should be organized into frequency of isolation of various nosocomial pathogens by anatomical site and hospital service. The tables generated should summarize antimicrobial susceptibility test results, thereby assisting clinicians in their choice of empiric antimicrobial therapy. Including cost information for the most commonly used antimicrobials also may be appropriate.

Specialized Studies for Outbreak Investigations

When the infection control team detects a cluster or outbreak of nosocomial infection, they must act promptly to define the extent of the outbreak, learn the mode of transmission of the pathogen, and instititute appropriate control measures. The clinical microbiology laboratory must provide appropriate support during this time. Because the demands on the laboratory may be great during outbreak settings, advance preparation should be performed. Laboratory personnel should communicate regularly with infection control practitioners regarding the types of outbreaks that have occurred in the past, and what laboratory resources would be required should such outbreaks occur in the future. The extra costs associated with outbreak investigations also should be anticipated by hospital administration and calculated into annual budgets if possible. Costs should not be borne by the laboratory or charged to individual patients involved in the outbreak.

Molecular Typing in the Outbreak Setting

Outbreaks of nosocomial infection often result from exposure of a number of hospitalized patients to a common source or reservoir of a pathogenic agent (e.g., a healthcare worker colonized with *S. aureus*, water from a hot water tank colonized with *Legionella*

spp.). The organisms causing the outbreak in these cases all derive from a single strain (i.e., they can be said to be clonally related). The hospital epidemiology program therefore may request that the microbiology laboratory characterize potential outbreak strains to determine their genetic relatedness. In the appropriate clinical setting, species identification and antimicrobial susceptibility testing (antibiogram) may provide strong evidence for an epidemiologic link. However, because nosocomial pathogens also are frequently colonizing flora or normal environmental flora, more sensitive methods of strain delineation are often necessary (49,50 and 51). For this reason, most phenotypic typing methods (e.g., AST, biochemical profiles, bacteriophage susceptibility patterns, multilocus enzyme electrophoresis profiles), which discriminate poorly among strains, have been replaced almost completely by genotypic or DNA-based typing methods (49,50 and 51).

The cost-effective application of genotypic typing methods requires that they be used only for well-defined epidemiologic objectives. These objectives include: (a) determination of the extent of an outbreak, (b) determination of the mode of transmission of nosocomial pathogens, (c) evaluation of the efficacy of preventative measures, and (d) monitoring of infection in high risk areas (e.g., intensive care units), where cross-infection is a recognized hazard.

The ideal genotypic typing system should be standardized, reproducible, stable, sensitive, broadly applicable, readily available, inexpensive, and of proven value in epidemiologic investigation. Table 57.4 summarizes the characteristics of several DNA-based typing methods. Pulsed-field gel electrophoresis (PFGE) and certain polymerase chain reaction (PCR)-based typing methods perform well for a wide array of nosocomial pathogens and are the molecular typing methods most frequently used to investigate nosocomial infections and outbreaks. Further discussion of the relative advantages and disadvantages of the many available typing systems is beyond the scope of this chapter and has been well summarized in several recent reviews (49,50,51,52 and 53).

TABLE 57.4. DNA-BASED METHODS FOR EPIDEMIOLOGICAL TYPING OF MICROORGANISMS

Method	Substrate	Principle Characteristics	Examples
Plasmid fingerprinting	Plasmid DNA	Technically simple and inexpensive; potentially unstable because of loss of plasmids; may be augmented by restriction endonuclease digestions	*S. aureus*, coagulase-negative staphylococci, *Klebsiella* spp., *Serratia* spp., *Enterobacter* spp.
Restriction endonuclease analysis of chromosomal DNA with conventional electrophoresis	Chromosomal DNA	Reproducible and broadly applicable; complex banding patterns; very difficult to interpret	*Clostridium difficile, Enterococcus faecium, S. aureus*
Restriction-fragment length polymorphism (RFLP) analysis with nucleic acid probes	Chromosomal DNA	Includes insertion sequence analysis and ribotyping; broadly applicable; multistep process; automated system available; limited discriminatory power (ribotyping)	*M. tuberculosis, S. aureus, C. albicans*, Enterobacteriaceae
Pulsed-field gel electrophoresis	Chromosomal DNA	Broadly applicable; excellent reproducibility and discriminatory power; employs rare cutting restriction enzymes to generate large band DNA fragments (10–800 kb); fewer bands; slow turnaround time; expensive equipment	Staphylococci, enterococci, Enterobacteriaceae, *Pseudomonas* spp., *Candida* spp.
Polymerase chain reaction (e.g., RAPD, rep-PCR, CFLP, AFLP: see reference 52 for details)	Chromosomal DNA	Rapid and relatively easy to perform; universally applicable when random primers are used (RAPD); moderate to good discriminatory power	*S. aureus, C. difficile*, Enterobacteriaceae, *Candida* spp.

From reference 22.

RAPD, Random Amplified Polymorphic DNA; rep-PCR, repetitive extragenic palindromic–polymerase chain reaction; CFLP, cleavase fragment length polymorphism; AFLP, amplified fragment length polymorphism

Regardless of the typing system employed, several general principles should be emphasized. A valid comparison between two or more isolates can be difficult unless the samples are run under the same conditions, and usually on the same day with the same reagents. Additionally, in order to make valid conclusions about the epidemic strain during an outbreak, one must demonstrate that control isolates from epidemiologically unrelated patients and/or environmental sources are different from the outbreak strain (51, 54). Finally, regardless of the method used, if molecular typing is applied indiscriminately or without sound epidemiologic reasoning, the information gleaned often will be conflicting and confusing.

Unfortunately, because the methodology, nomenclature, and reference strains have not been standardized, it is not usually possible to compare molecular typing results obtained by different methods or different laboratories (53). Sophisticated computer-based analysis systems and fully automated molecular typing systems (e.g., Riboprinter, Qualicon, Wilmington, DE) are first steps toward standardization and quantitative analysis of molecular typing results (55), and some early standards and guidelines for the interpretation of typing results have been developed (56).

Of course, it is impossible to provide infection control practitioners with supplemental testing like molecular typing if the appropriate isolates have not been saved. The laboratory should plan ahead with infection control and save all epidemiologically important isolates. Decisions about the number and type of isolates to be banked and the duration of storage will differ from hospital to hospital based upon available resources. Certainly all isolates from blood and other normally sterile sites (e.g., cerebrospinal fluid) as well as multiply antibiotic-resistant strains from any site should be saved for a period of several months to years.

Cultures of Hospital Personnel and Environment

Investigating an outbreak of nosocomial infection may require culture of many specimens from patients, hospital personnel, and the environment. The use of selective media (i.e., media which inhibits growth of organisms other than those of epidemiologic interest) or differential media (i.e., media that results in distinctive colonial features of the target organism) can significantly reduce laboratory workload and expedite specimen processing in these circumstances. Likewise, enrichment cultures may be necessary to detect specific pathogens that may be present in low numbers, such as methicillin-resistant *S. aureus* (57). Given the resources required to process these cultures, it is imperative that they be obtained only when epidemiologically indicated.

Cultures of medical devices and the environment (surfaces, air, water) should be performed rarely and only when the evidence suggests involvement in the transmission of a nosocomial pathogen. Various potential sources and their appropriate culture methods are outlined in Table 57.5. Although the use of such cultures is frequently entertained, it should be emphasized that the cultures are labor intensive, nonstandardized, difficult to interpret, and rarely provide useful information.

TABLE 57.5. CULTURE OF POTENTIAL SOURCES OF CROSS-INFECTION IN NOSOCOMIAL INFECTION OUTBREAK[a]

Source	Culture Method	Comment
Blood products	Broth culture incubated aerobically and anaerobically at 30°–32°C for 10 days	Following transfusion reaction; obtain simultaneous blood cultures by venipuncture
Parenteral fluids and intravenous devices	Broth or membrane filter method	Culture needle, catheter, administration set, fluid, closure; obtain blood culture
Environmental surfaces	Swab-rinse or impression plate	No evidence that any particular level of contamination correlates with nosocomial infection
Tubes and containers	Broth-rinse or swab-rinse with semiquantitative plating	At least two colonies of each morphologic type should be picked for identification
Disinfectants and antiseptics	Plating of serial dilutions of the product with and without specific neutralizers	Organisms usually nonfermenting Gram-negative aerobic bacilli
Respiratory therapy equipment	Broth-rinse or swab-rinse	Only in situations of high endemic or epidemic levels of nosocomial respiratory infection
Air	Mechanical air sampler (preferred); settling (plates (poor)	No uniform agreement on acceptable levels of contamination; lack of correlation with infection
Water and ice	Membrane filter	Poor correlation of culture findings with illness
Hands of personnel	Broth-bag: 10–20 mL nutrient broth in sterile plastic bag. Wash hands in broth and plate semiquantitatively	May confirm the mechanism of cross-infection; impress the importance of hand washing

[a] Cultures to be performed only if clearly indicated by epidemiologic data. From reference 62.

Because the hands of healthcare workers are such an important vehicle for transmission of nosocomial pathogens from patient to patient, hand cultures are sometimes useful in confirming the mechanism of cross-infection during an outbreak investigation (58). A sensitive method of culturing hands to detect transient or resident flora is the broth-bag technique. The subject vigorously washes both hands in 10 to 20 mL of nutrient broth (supplemented with Tween 80, sodium thiosulfate, and lecithin to neutralize any residual antiseptics), and an aliquot of the broth is then plated onto selective and nonselective media and observed for growth of the pathogen of interest (59).

As a general rule, routine cultures of hospital personnel and the environment should not be performed. Exceptions include routine monitoring of sterilization, infant formula, and other hospital-prepared products; blood components prepared in an “open” system; hemodialysis fluid; and disinfected equipment. Sampling activities that have been identified specifically as unnecessary because of high cost and lack of clinical or epidemiologic benefit include routine culturing of patients or hospital personnel, routine sampling of commercial patient-care items, in-use testing of antiseptics and disinfectants, random culturing of blood units, and routine culturing of respiratory therapy equipment, peritoneal dialysate, and air. These routine cultures are a burden to the laboratory and seldom, if ever, provide useful information or lead to specific interventions (60, 61). They only should be done with appropriate epidemiologic indications or as part of an educational program (e.g., routine broth-bag hand culturing in the intensive care unit to reinforce good handwashing practices).

CONCLUSION

Part of "57 - Role of the Clinical Microbiology Laboratory in Hospital Epidemiology and Infection Control"

The clinical microbiology laboratory is an essential component of an effective hospital epidemiology and infection control program. The changing spectrum of nosocomial pathogens and the rapid development and application of new technologies in the diagnostic laboratory require ongoing cooperation and collaboration between the laboratory and infection-control personnel. Appropriate application of the newer methods for detection, identification, and typing of nosocomial pathogens will enhance the effectiveness of infection control efforts. A good working relationship between the clinical microbiologist and hospital epidemiologist has a positive influence on both laboratory and infection control operations and facilitates the investigation and control of nosocomial infection problems.

REFERENCES

1. CDC NNIS System. National Nosocomial Infections Surveillance (NNIS) report, data summary from October 1986–April 1997, issued May 1997. *Am J Infect Control* 1997;25:477–487.

2. CDC NNIS System. National Nosocomial Infections Surveillance (NNIS) system report, data summary from October 1986–April 1998, issued June 1998. *Am J Infect Control* 1998;26:522–533.

3. Pfaller MA. Nosocomial candidaisis: emerging species, reservoirs, and modes of transmission. *Clin Infect Dis* 1996;22(suppl 2):S89–S94.

4. Cormican MG, Jones RN. Emerging resistance to antimicrobial agents in gram-positive bacteria: enterococci, staphylococci, and non-pneumococcal streptococci. *Drugs* 1996;51(suppl 1):6–12.

5. Spengler RF, Greenough WE III. Hospital costs and mortality attributed to nosocomial bacteremias. *JAMA* 1987;240:2455–2458.

6. Gross PA, Neu HC, Aswapokee P, et al. Deaths from nosocomial infections: experience in a university hospital and a community hospital. *Am J Med* 1980;68:219–223.

7. Martin MA, Pfaller MA, Wenzel RP. Mortality and hospital stay attributable to coagulase-negative staphylococcal bacteremia. *Ann Intern Med* 1989;110:9–16.

8. Landry SL, Kaiser DL, Wenzel RP. Hospital stay and mortality attributed to nosocomial enterococcal bacteremia: A controlled study. *Am J Infect Control* 1989;17:323–329.

9. Edmond MB, Ober JF, Dawson JD, et al. Vancomycin-resistant enterococcal bacteremia: Natural history and attributable mortality. *Clin Infect Dis* 1996;23:1234–1239.

10. Wey SB, Mori M, Pfaller MA, et al. Hospital acquired candidemia: attributable mortality and excess length of stay. *Arch Intern Med* 1988;148:2642–2645.

11. Pittet D, Tarara D, Wenzel RP. Nosocomial bloodstream infection in critically ill patients: excess length of stay, extra costs, and attributable mortality. *JAMA* 1994;271:1598–1601.

12. Emori TG, Gaynes RP. An overview of nosocomial infections, including the role of the microbiology laboratory. *Clin Microbiol Rev* 1993;6:428–442.

13. Haley RW, Culver DH, White JW, et al. The efficacy of infection control surveillance and control programs in preventing nosocomial infections in U.S. hospitals. *Am J Epidemiol* 1985;121:182–205.

14. Cruse PJE. Surgical wound sepsis. *Can Med Assoc J* 1970;102:251–258.

15. Emori TG, Haley RW, Garner JS. Technique and use of nosocomial infection surveillance in U.S. hospitals. *Am J Med* 1981;70:933–940.

16. Wenzel RP, Osterman CA, Hunting KJ, et al. Hospital-acquired infections. I. Surveillance in a university hospital. *Am J Epidemiol* 1976;103:251–160.

17. Gross PA, Beaugard A, Van Antwerpen C. Surveillance for nosocomial infections: can the sources of data be reduced% *Infect Control* 1980;1:233–236.

18. Yakoe DS, Anderson J, Chambers R, et al. Simplified surveillance for nosocomial bloodstream infections. *Infect Control Hosp Epidemiol* 1998;19:657–660.

19. Broderick A, Mori M, Nettleman MD, et al. Nosocomial infections: validation of surveillance and computer modeling to identify patients at risk. *Am J Epidemiol* 1990;131:734–742.

20. Wong LK, Barry AL, Horgan SM. Comparison of six different criteria for judging the acceptability of sputum specimens. *J Clin Microbiol* 1982;16:627–631.

21. Ashford DA, Kellerman S, Yakrus M, et al. Pseudo-outbreak of septicemia due to rapidly growing mycobacteria associated with extrinsic contamination of culture supplement. *J Clin Microbiol* 1997;35:2040–2042.

22. Pfaller MA, Herwaldt LA. The clinical microbiology laboratory and infection control: emerging pathogens, antimicrobial resistance, and new technology. *Clin Infect Dis* 1997;25:858–870.

23. Kellog JA. Culture versus direct antigen assays for detection of microbial pathogens from lower respiratory tract specimens suspected of containing the respiratory syncytial virus. *Arch Pathol Lab Med* 1991;115:451–458.

24. Gerding DN, Brazier JS. Optimal methods for identifying Clostridium difficile infections. *Clin Infect Dis* 1993;16(suppl 4):S439–S442.

25. Kearns AM, Freeman R, Steward M, et al. A rapid PCR technique for detecting *M. tuberculosis* in a variety of clinical specimens. *J Clin Pathol* 1998;51:922–924.

26. Plouffe JF, File TM Jr, Breiman RF, et al. Reevaluation of the definition of Legionnaires' disease: use of the urinary antigen assay. Community Based Pneumonia Incidence Study Group. *Clin Infect Dis* 1995;20:1286–1291.

27. Laussucq S, Schuster D, Alexander WJ, et al. False-positive DNA probe test for *Legionella* species associated with a cluster of respiratory illnesses. *J Clin Microbiol* 1988;26:1442–1444.

28. Centers for Disease Control and Prevention. Rapid HIV Tests: Questions/Answers. March 1998.

29. Ferraro MJ, Jorgensen JH. Instrument based antibacterial susceptibility testing. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of Clinical Microbiology* 6th ed. American Society for Microbiology 1995:1379–1384.

30. Stager CE, Davis JR. Automated systems for identification of microorganisms. *Clin Microbiol Rev* 1992;5:302–327.

31. Robinson A, McCarter YS, Tetreault J. Comparison of crystal enteric/nonfermenter system, API 20E system, and Vitek automicrobic system for identification of Gram-negative bacilli. *J Clin Microbiol* 1995;33:364–370.

32. Baker CN, Banerjee SN, Tenover FC. Evaluation of alamar colorimetric MIC method for antimicrobial susceptibility testing of Gram-negative bacteria. *J Clin Microbiol* 1994;32:1261–1267.

33. Kitch TT, Jacobs MR, Applebaum PC. Evaluation of rapid one system for identification of 379 strains in the family enterobacteriacae and oxidase-negative, gram-negative nonfermenters. *J Clin Microbiol* 1994;32:931–934.

34. Doern GV, Brueggemann AB, Pfaller MA, et al. Assessment of laboratory performance with *Streptococcus pneumoniae* antimicrobial susceptibility testing in the United States: a report from the College of American Pathologists Microbiology Proficiency Survey Program. *Arch Pathol Lab Med* 1999;123:285–289.

35. Alonso-Echanove J, Robles B, Jarvis WR, and the Spanish VRE Study Group. Proficiency of clinical laboratories in Spain in detecting vancomycin-resistant *Enterococcus* spp. *J Clin Microbiol* 1999;37:2148–2152.

36. Ribeiro J, Vieira FD, King T, et al. Misclassification of susceptible strains of *Staphylococcus aureus* as methicillin-resistant *S. aureus* by a rapid automated susceptibility testing system. *J Clin Microbiol* 1999;37:1619–1620.

37. Thompson KS, Sanders CC. Detection of extended-spectrum beta-lactamases in members of the family Enterobacteriaceae: comparison of the double disk and three dimensional test. *Antimicrob Agents Chemother* 1992;36:1877–1882.

38. Beidenbach DJ, Jones RN. Interpretive errors using an automated system for the susceptibility testing of imipenem and aztreonam. *Diagn Microbiol Infect Dis* 1995;21:57–60.

39. Philippon A, Arlet G, Lagrange PH. Origin and impact of plasmid-mediated extended spectrum beta-lactamases. *Eur J Clin Microbiol Infect Dis* 1994;13(S1):17–29.

40. Sanders WE, Sanders CS. Inducible beta-lactamases: clinical and epidemiologic implications for use of newer cephalosporins. *Rev Infect Dis* 1988;10:830–838.

41. Jones RN, Sader HS, Erwin ME, et al. Emerging multiply resistant enterococci among clinical isolates. I. Prevalence data from 97 medical center surveillance study in the United States. *Diag Microbiol Infect Dis* 1995;21:85–93.

42. Schwalbe RS, Stapleton JT, Gilligan PH. Emergence of vancomycin resistance in coagulase-negative staphylococci. *N Engl J Med* 1987;316:927–931.

43. Centers for Disease Control and Prevention. Update: *Staphylococcus aureus* with reduced susceptibility to vancomycin-United States, 1997. *MMWR* 1997;46:813–815.

44. Doern GV, Pfaller MA, Kugler K, et al. Prevalence of antimicrobial resistance among respiratory tract isolates of *Streptococcus pneumoniae* in North America: 1997 results from the SENTRY antimicrobial surveillance program. *Clin Infect Dis* 1998;27:764–768.

45. Pfaller MA, Jones RN, Doern GV, et al. Survey of blood stream infections attributable to Gram-positive cocci: frequency of occurrence and antimicrobial susceptibility of isolates collected in 1997 in the United States, Canada, and Latin America from the SENTRY Antimicrobial Surveillance Program. *Diagn Microbiol Infect Dis* 1999;33:283–297.

46. Jones RN, Erwin ME, Anderson SC. Emerging multiply resistant enterococci among clinical isolates: II. Validation of the Etest to recognize glycopeptide-resistant strains. *Diagn Microbiol Infect Dis* 1995;21:95–100.

47. Huang MC, Baker CN, Banerjee S, et al. Accuracy of the Etest for determining antimicrobial susceptibilities of staphylococci, enterococci, *Campylobacter* jejuni, and Gram-negative bacteria resistant to antimicrobial agents. *J Clin Microbiol* 1992;30:3243–3248.

48. Sanchez ML, Barrett MS, Jones RN. Use of the Etest to predict high-level resistance to aminoglycosides among enterococci. *J Clin Microbiol* 1992;30:3030–3032.

49. Sader HS, Hollis RJ, Pfaller MA. The use of molecular techniques in the epidemiology and control of infectious diseases. *Clin Lab Med* 1995;15:407–431.

50. Weber S, Pfaller MA, Herwaldt LA. Role of molecular epidemiology in infection control. *Infect Dis Clin North Am* 1997;11:257–278.

51. Arbeit RD. Laboratory procedures for the epidemiologic analysis of microorganisms. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of Clinical Microbiology*. 7th ed. Washington, D.C., American Society for Microbiology, 1999:107–115.

52. Olive DM, Bean P. Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol* 1999;37:1661–1669.

53. Tenover FC, Arbeit RD, Goering RV, the Molecular Working Group of the Society for Healthcare Epidemiology of America. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. *Infect Control Hosp Epidemiol* 1997;18:426–439.

54. Mayer LW. Use of plasmid profiles in epidemiologic surveillance of disease outbreaks and in tracing the transmission of antibiotic resistance. *Clin Microbiol Rev* 1988;1:228–243.

55. Pfaller MA, Wendt C, Hollis RJ, et al. Comparative evaluation of an automated ribotyping system versus pulsed-field gel electrophoresis for epidemiologic typing of clinical isolates of *Escherichia coli* and *Pseudomonas aeruginosa* from patients with recurrent Gram-negative bacteremia. *Diagn Microbiol Infect Dis* 1996;25:1–8.

56. Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* 1995;33:2233–2239.

57. Kernodle DS, Barg NL, Kaiser AB. Low-level colonization of hospitalized patients with methicillin-resistant coagulase-negative staphylococci and emergence of the organisms during surgical antimicrobial prophylaxis. *Antimicrob Agents Chemother* 1988;32:202–208.

58. Doebbeling BN, Hollis RJ, Isenberg HD, et al. Restriction fragment analysis of a *Candida* tropicalis outbreak of sternal wound infections. *J Clin Microbiol* 1991;29:1268–1270.

59. Doebbeling BN, Pfaller MA, Houston AK, et al. Removal of nosocomial pathogens from the contaminated glove: implication for glove reuse and handwashing. *Ann Intern Med* 1988;109:394–398.

60. Simmons BP. Centers for Disease Control guidelines for hospital environmental control/microbiologic surveillance of the environment and of personnel in the hospital. *Infect Control* 1981;2:145–146.

61. Garner JS, Favero MS. Centers for Disease Control guideline for handwashing and hospital environmental control, 1985.

62. Pfaller MA, Cormican MG. Microbiology: the role of the clinical laboratory. In: Wenzel RP, ed. *Prevention and control of nosocomial infections* Baltimore: Williams & Wilkins, 1997:95–118.

63. Jones RN, Low DE, Pfaller MA. Epidemiologic trends in nosocomial and community-acquired infections due to Gram-positive bacteria: the role of streptogramins and other newer compounds. *Diagn Microbiol Infect Dis* 1999;33:101–112.

58
Autopsy Microbiology

Mary J. Reznicek

Franklin P. Koontz

Autopsies are one of the most powerful learning tools the field of medicine has in its armamentarium for understanding disease. They are performed for a variety of reasons: to gain additional information into the pathogenesis of disease, to identify clinically unsuspected disease processes, and to correlate premortem clinical diagnoses with postmortem diagnoses. The autopsy is a valuable teaching tool and quality control measure that works most efficiently when there is open communication between the primary physician and the pathologist. This ensures that valuable information is exchanged between both parties to facilitate a greater understanding of disease processes.

Despite the advent of the antibiotic era, infection has remained a significant cause of death in the United States. Therefore, the ability to identify an infectious process at autopsy is crucial. The diagnosis of an infectious disease at autopsy, like the diagnoses made in living patients, can be accomplished by positive cultures, tissue histology, and serologic procedures. These tests can be utilized singly or in combination.

Postmortem cultures can be a very useful component of the postmortem examination. They can be utilized to detect infections not suspected clinically, such as mycobacterial or fungal infections. They also can be useful in pinpointing the source of a known sepsis. For example, the discovery at autopsy of a *Staphylococcus aureus* psoas muscle abscess in a patient with known *S. aureus* sepsis and no other potential sites of bacterial seeding would certainly make the abscess the most likely source. Culturing also can be useful as a quality control measure to assess the effectiveness of antimicrobial therapy. As a quality control measure, the identification of an unsuspected infectious process can also help clinicians identify, in a retrospective fashion, other tests that might have aided in the premortem identification of an infection.

All of the knowledge gained by postmortem culturing is dependent on the utilization of proper methods to collect and process samples. This would include proper cooling of the body after death, avoiding lengthy delays in the performance of an autopsy, the use of sterile culturing techniques, and the use of good interpretive skills to correlate the clinical history, culture results, and the gross and histologic findings.

MICROBIOLOGY OF AUTOPSY TISSUE

Part of "58 - Autopsy Microbiology"

The use of cultures in autopsy microbiology is controversial, and opinions vary on its usefulness. Several basic questions concerning the validity of postmortem cultures have been discussed in the literature. Most of these questions center around the basic question of whether a dead human body with its own inherent normal bacterial flora is suitable for culturing. More specifically, these questions are:

1. Does agonal or postmortem bacterial invasion occur?
2. What effect does the postmortem interval have on culture results?
3. Are human viscera sterile?

Two theories exist that support the idea that autopsy material is unsuitable for culture: the theories of *agonal invasion* and *postmortem invasion*. The theory of agonal invasion states that the diminished vitality of the dying body predisposes it to disseminated infection. The theory of postmortem invasion suggests that bacteria multiply and migrate throughout the body after death. Potential sites of dissemination include the normal flora of the oropharynx, which can seed the lung and blood, and normal gastrointestinal (GI) tract flora, which can seed the liver, spleen, and blood.

Both of these theories promote the concept that increasing numbers and types of organisms would be seen after death. Similarly, one would expect that the number of positive cultures would increase in direct proportion to the postmortem time interval. However, this concept does not appear to be valid if one considers several studies that illustrate that the postmortem interval has no effect on the number of positive cultures at autopsy (1,2,3,4,5,6 and 7). The postmortem interval in these studies ranged from 14 to 48 hours, and bodies were placed in a cooler at 4° to 5°C soon after death. There was no increase in positive autopsy cultures as the postmortem interval increased. In contrast, Carpenter and Wilkins (8) conducted a retrospective study of 2,033 autopsy cases that showed a rise in postmortem positive cultures with increasing postmortem intervals up to 18 hours. Unfortunately, because the distribution of cases between short and long delay intervals was not provided, the conclusions of this study are open to question. Overall, the majority of the published evidence suggests that, if proper cooling of the body takes place and the autopsy is performed in a timely manner, agonal or postmortem invasion (if either exists) should not produce false-positive culture results.

The high frequency with which positive cultures are isolated from autopsy tissues that clinically and pathologically appear

normal has raised the question of whether human viscera are sterile. This question has sparked controversy about the increasing numbers of cadaver organs harvested for transplant. A study of 148 surgical biopsies (excluding skin and GI tract) showed that approximately 28% were culture-positive (2). Because the samples were surgical biopsy specimens, one cannot argue that agonal or postmortem invasion played any role. Considering the transient bacteremias that are known to occur in individuals undergoing dental or urologic procedures, it is not surprising that human viscera may not be sterile at any given time. In healthy individuals, these numbers of circulating bacteria are likely to be too small to result in clinical disease. In interpreting a positive postmortem culture result, one must be careful not to equate a positive culture with disease without first correlating the culture results with the clinical history and histologic findings. A positive culture result may be *valid* in light of the above discussion, but the determination of its clinical significance is of major importance in ascertaining its contribution to the cause of death.

Knapp and Kent (1) addressed this issue and suggested that low levels of bacteria may be present in noninfected organs at the time of autopsy. They showed by quantitative cultures that less than 10^5 organisms per milliliter in lung tissue correlated with the absence of pulmonary disease by autopsy histology and, conversely, more than 10^5 organisms per milliliter in lung tissue correlated with clinical and pathological pneumonia.

It is imperative to understand that the data and concepts expressed in the above discussion are totally dependent upon proper culturing techniques at the time of autopsy. Although the “sterile autopsy technique” suggested by O'Toole et al. (4) is impractical for routine use, DeJongh et al. (3) have described a more practical technique. This technique is described in detail elsewhere in this chapter (see Specimen Collection and Transport, below).

Avoiding contamination of tissues with fluids introduced into the body cavities at the time of dissection is an important aspect of culturing. One study by Silva and Sonnenwirth (7) reported that blood cultures obtained from the heart through a closed chest were more likely to yield significant results than those obtained from an exposed right atrium. They also showed that bowel manipulation increased the number of positive cultures. Contamination of postmortem cultures with normal flora from the bowel or nasopharynx can lead to significant problems. However, these can be controlled by strict adherence to protocols or procedures adapted for the care and handling of the cadavers from the moment of death to the completion of the autopsy.

SPECIMEN COLLECTION AND TRANSPORT

Part of "58 - Autopsy Microbiology"

Preparation

Before beginning the autopsy, it is imperative to fully review the patient's chart. The value of the autopsy is directly proportional to the preparation by the prosector. The information contained in the patient's chart can be used to ascertain the clinical questions that need to be answered by the autopsy. Areas of the hospital record that are particularly helpful include current hospital notes, transfer notes, and past discharge notes, which often provide an excellent review of the patient's medical history. Radiology examinations such as radiographs, computed tomography (CT) scans, and white blood cell (WBC) scans will often point out areas of suspected infections. For example, when the clinical note states “suspected pancreatic pseudocyst,” the CT scan may be helpful in describing the exact location of the cyst and the size. If the prosector is unsure of the clinical questions, it can be extremely helpful to consult with the physician who managed the patient. This is strongly recommended in every autopsy case and can often provide more recent data that may not be evident in the chart.

In reading the patient's record, it is important to delineate those cases in which infection was thought to play a clinically important role and those cases in which the cause of death is unknown and infection is a possible cause. In the former situation, the prosector should study clinically suspicious areas of infection by gross and microscopic examination and possibly by culture. Autopsy data, whether gleaned from gross observation, histologic examination, or microbiological culture, serve as quality control mechanisms for antemortem concepts. The autopsy can clarify these concepts regardless of whether the findings are positive or negative. For example, the clinicians and the radiologist may suspect a right lower lobe pneumonia in a patient who dies a few days later. The prosector, by reading the chart, knows this information and, despite the gross findings of dependent congestion and a lack of consolidation in the right lung, can obtain samples of that area for microbiological culture and histologic examination. By doing so, the prosector can adequately address the antemortem diagnosis of pneumonia. Additional information can be gleaned from the hospital record relative to the use of antimicrobial therapy. This information could be critical in evaluating the effectiveness of specific antimicrobial therapy.

Diagnosing an unsuspected infection when the cause of death is unknown is also a responsibility of the prosector. It is important to consider doing cultures in this situation. It is relatively easy to culture during the autopsy; however, the chance is lost after the tissue has been placed in formalin and the cadaver has been taken to the funeral home. For example, a 35-year-old male is admitted to the hospital with severe congestive heart failure, and he dies before a workup is performed. No definite etiology for the congestive failure is known. The clinical differential includes cardiomyopathy, myocardial infarction, *cor pulmonale*, and myocarditis. Histologic sections of the heart show scattered inflammatory cells indicative of a myocarditis. In this case, confirmatory viral cultures would not only add strength to the diagnosis of myocarditis but also would give valuable epidemiologic information. Similar information also can be supplied by specific serologic tests. If antemortem tests were performed for specific antibodies, postmortem serum should be submitted for convalescent titers.

In reviewing the hospital record, it is important to ascertain whether the patient may have had an infectious disease that would place individuals present at the autopsy at risk, such as tuberculosis, hepatitis, and acquired immune deficiency syndrome (AIDS). In cases of suspected tuberculosis, PPD results and sputum culture results may be informative. For hepatitis, elevated liver function tests or a previous history of transfusion or drug abuse should be noted. If AIDS is suspected, the patient's human

immunodeficiency virus (HIV) status and personal risk history should be determined. The presence of multiple opportunistic infections, such as *Pneumocystis carinii* pneumonia and/or atypical mycobacterial infection, in an otherwise normal host, or an unknown cause of death, also should arouse suspicions. Although universal precautions are recommended for all autopsies, particular caution should be exercised in instances where infectious diseases are suspected. In such cases, it is best to limit the number of people in the autopsy suite, thereby decreasing the number of people possibly exposed. In addition, consideration might be given to limiting the autopsy to areas of clinically suspected disease. For example, in a patient who died of fulminant hepatitis, there may be no reason to examine the brain except for completeness. In these cases, limiting the autopsy may be indicated.

Culturing

The extent of postmortem culturing is a problem that has never been specifically addressed. If only one area is to be cultured, the spleen and heart blood are favored as reliable sites. However, additional cultures from other areas such as lungs and kidneys will more accurately reflect the spectrum of bacteria affecting the patient. The clinical history and the gross appearance of the organs at autopsy also can help the prosector to decide how many postmortem cultures are necessary. For example, in a patient with a clinical history of sepsis, cultures of heart blood and spleen should be taken even if gross organ appearance is normal. If both sites grow identical organisms, a diagnosis of sepsis is supported.

In transplant patients and other immunocompromised hosts, special diligence must be shown. In the event of an established disseminated infection, it is clinically important to identify the site of entry of the infection and/or the primary organ of involvement that may be seeding the rest of the body. Entry sites other than the typical respiratory route include indwelling catheters, Hickman catheters, skin wounds, and nasal lesions. Extensive culturing in these patients is recommended to document the dissemination of the infection and to identify possible sources of infections.

The extent of culturing not only refers to the number of cultures but to the types of cultures as well. In immunocompromised patients, the use of special viral, fungal, or mycobacterial cultures in addition to routine bacteriology should be considered. These special cultures must be clearly requested on the test requisition form because they are not routinely performed on autopsy specimens in most laboratories.

The use of cultures should not be discounted because of antemortem antimicrobial therapy (8,9,10,11 and 12). Some studies have shown that antibiotics per se do not influence the number of positive cultures at autopsy. In addition, the availability of resin bottles or other antibiotic removal devices for blood cultures also may help increase this yield. In patients who have received antimicrobial therapy, postmortem cultures can assess the effectiveness of the drug.

As previously discussed, the techniques employed in obtaining postmortem cultures clearly affect the usefulness of the results. A more practical technique, described by DeJongh et al. (3), compares very favorably with those used in autopsies performed under strict sterile operative conditions. This technique, as it applies to blood, urine, CSF, and tissue, includes the following steps:

1. After making the initial incisions, visually inspect the cavities for fluid accumulations that can be cultured immediately before manipulation. Gently inspect organs for areas suspicious for infection. Consider culturing organs that are *clinically* suspected of harboring infection.
2. Dry the area to be cultured by searing with a hot steel spatula while keeping the organ elevated away from any secretions. Avoid dripping gloves that may contaminate the surface after it has been sterilized.
3. For tissue culture, remove a 1-cm^3 portion of tissue with sterile forceps and blade and place it in a sterile Petri dish that has been identified properly with the patient's name, the culture site, and the specific type of culture desired (e.g., viral and fungal).
4. For blood culture, aspirate blood from the right atrium with a sterile syringe and inject 10 mL into aerobic and anaerobic blood culture bottles whose surfaces have been sterilized. Consider using a resin culture bottle if the patient has been on antibiotics. In cases of suspected fungemia, centrifugation lysis blood tubes (Isolator, DuPont) may be used.
5. To obtain urine for culture, aspirate urine directly from the bladder with a sterile syringe and inject it into a sterile tube.
6. CSF fluid for culture can be obtained from the lateral ventricle through the *corpus callosum* after the skull cap is removed or by cisternal tap through the skin. Both methods are adequate when performed aseptically.
7. Cultures of abscesses and granulomata should be obtained from both the center and the peripheral wall of the lesions. Depending on the disease, the organisms can be in either location.
8. Use different sets of sterile instruments for each culture/site.
9. If the microbiology laboratories are not all in one location, divide the specimens aseptically for the bacteriology, mycology, TB, and virology laboratories. The specimens should be delivered promptly to the microbiology laboratory.

Thus, by careful sampling, surface decontamination, and handling of tissues, one can obtain postmortem cultures that are useful in assessing the importance of infection in the death of a given patient.

It is very important to adequately fill out the laboratory requisitions. Tissue from surgical specimens may be automatically plated for bacteriology, anaerobes, and fungi; however, autopsy cultures, because of their low yield, may not get the same treatment. It is important to let the microbiology laboratory know what is suspected. For example, if you clinically suspect an organism that has special growth requirements, like *Haemophilus*, that information must be relayed to the laboratory to ensure proper handling of the specimen.

Tissue specimens for viral culture should be placed in transport media and kept at 4°C if any delay in transport to the laboratory is anticipated. If a longer delay (more than 48 hours) is anticipated, the specimen should be snap frozen in liquid nitrogen. Slow freezing may kill some unstable viruses. In suspected viral syndromes, the virus may be excreted in body fluids and thus may be present

in urine and feces as well as in the organ of primary interest. Culturing multiple areas may result in higher yield (Table 58.1).

TABLE 58.1. SPECIFIC BODY SITES FOR VIRAL CULTURING

Syndrome	Specimens
Respiratory	Lung, tracheal swab
CNS	Brain, spinal cord, CSF, feces
Cardiac	Heart, nasopharyngeal swab, feces, pericardial fluid
Liver	Liver, nasopharyngeal swab, feces, blood
GI	Feces, GI tract

CNS, central nervous system: CSF, cerebrospinal fluid: GI, gastrointestinal

One procedure that can be performed immediately in the morgue is a touch preparation of the infected organ. A Diff-Quick (Fisher Scientific) or Giemsa stain set-up and a microscope can be kept in the morgue. This procedure can permit the rapid identification of yeast, molds, or granulomatous inflammation and thereby help direct the extent and type of culturing (see Fig. 58.6). It also presents the opportunity to give rapid feedback to the clinician while the clinical situation is still fresh in his or her mind.

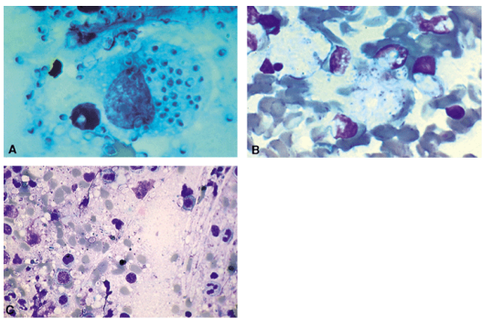


FIGURE 58.6. A: Touch preparation from a lymph node with internal structure of Histoplasma (Diff-Quick, ×1,000, oil). **B:** Touch preparation from a lung with internal structure of *Pneumocystis carinii* (Diff-Quick, ×1,000, oil). **C:** Touch preparation from a spleen containing *Candida* showing effects of amphotericin therapy (ghost pseudohyphae) (Diff-Quick, ×400).

Specimen Processing

Autopsy culture specimens should be processed by the laboratory in much the same way as clinical specimens. These techniques already have been described in previous chapters. However, in processing these specimens, the general philosophy is to avoid picking up less meaningful data. Broth cultures can be avoided because the low levels of organisms at autopsy are unlikely to be clinically significant. Culture plates that grow multiple organisms also are unlikely to be significant in most clinical situations.

Additional microbiological methods that can be applied include print culture (13), immunofluorescence, DNA probes, and polymerase chain reaction (PCR) (14). In general, these methods involve additional laboratory time and expensive reagents. Print cultures involve pressing culture media directly against the block of tissue to be cultured. The print cultures then can be directly correlated to the presence of an inflammatory response in the tissue section, allowing a rough quantitation of organisms to be made. A similar principle was advocated by Knapp and Kent (1) 12 years earlier when they compared quantitative cultures of lung tissue with Gram stains of tissue imprints and histologic sections. The advantage of this technique is the ability to directly correlate culture information to histologic sections. In both studies, contaminants were readily identified by the low numbers of organisms grown in culture (1) or by the pattern of growth. For example, in the print culture method, contaminating bacteria are generally restricted to the edges of the culture imprint (13).

At present, rapid diagnostic methods such as DNA probes and immunofluorescence should be reserved for autopsy cases in which an organism is suspected clinically and pathologically despite negative routine laboratory cultures and tissue sections. Special techniques also may be necessary for organisms that cannot be easily cultured or identified by routine histologic stains, such as *Rickettsia*.

Special histologic stains on formalin-fixed autopsy tissue also can be performed alone or in combination with cultures to diagnose an infectious process. Which special stains to order can often be determined by the routine hematoxylin and eosin (H&E) stain, which provides information as to the type and degree of cellular host response. These host responses can be divided into the broad categories of acute inflammation, chronic inflammation, and granulomatous inflammation, and the identification of one of these can alert the prosector to search for the presence of an organism that characteristically elicits that type of host response (Table 58.2).

TABLE 58.2. HOST TISSUE REACTIONS TO ORGANISMS

Host Tissue Reaction	Organisms
Acute inflammation	Routine bacteria, actinomyces, *Nocardia, Blastomyces, Candida*, and herpes virus
Chronic inflammation	Viruses, *Pneumocystis*
Granulomatous inflammation	Mycobacteria, *Cryptococcus, Histoplasma, Coccidioides, Candida*
Caseous necrosis	Mycobacteria, *Histoplasma, Coccidioides*
Thrombosis	*Aspergillus*, Zygomycetes, *Fusarium*

Several characteristic host reactions can be identified. Infection with *Pneumocystis* can produce honeycombed, intraalveolar material seen in immunocompromised patients as well as the interstitial plasma cell infiltrates seen in patients with intact immune systems. Viruses such as *cytomegalovirus* and herpes simplex, when they occur on a mucosal surface, may cause necrosis and ulceration with acute inflammation. Other viruses may only elicit chronic inflammation. Areas of infarction with vascular thrombosis should suggest the possibility of *Aspergillus* or *Zygomycetes*.

In general, the host response is a reflection of the status of the patient's immune system. For example, a patient with a competent immune system forms a *caseous granuloma* around an organism such as *Mycobacterium* tuberculosis. Few organisms, if any, can be identified in these lesions. In patients with AIDS or chronic immunosuppression, well-formed granulomata usually are not found. The response may range from ill-formed collections of histiocytes to only scattered histiocytes. In these cases, abundant organisms usually are seen within histiocytes.

After developing a differential diagnosis as to the type of infectious process, the proper histologic stain can be selected. Table 58.3 lists the most commonly employed histologic stains and the organisms they demonstrate (Fig. 58.1,Fig. 58.2,Fig. 58.3,Fig. 58.4,Fig. 58.5,Fig. 58.6 and Fig. 58.7). Under ideal circumstances, these results can be correlated to positive culture results.

TABLE 58.3. COMMONLY USED HISTOLOGICAL STAINS TO DEMONSTRATE ORGANISMS IN TISSUES

Stain	Organism	Example
Hematoxylin and eosin	Viral inclusions (specifically CMV, *herpes*, adenovirus, rabies); Toxoplasma; dematiaceous (pigmented) fungi; other fungi (*Aspergillus*, Zygomycetes)	CMV (Fig. 53.1 *A*) *Herpes* (Fig. 53.1 *B*) Adenovirus (Fig. 53.1 *C*) *Toxoplasma* (Fig. 53.1 *D*)
	Demonstrates cellular response	*Aspergillus* (Fig. 53.1 *E*)
Gomori's methenamine silver nitrate (GMS)	Fungi (yeast and mycelial forms); *Pneumocystis*	*Histoplasma* (Fig. 53.2 *A*) *Candida* (Fig. 53.2 *B*) *Aspergillus* (Fig. 53.2 *C*) *Blastomyces* (Fig. 53.2 *D*)
Periodic acid–Schiff (PAS)	Fungi (yeast and mycelial forms), especially good for sporotrichosis, *Blastomyces*, and Zygomycetes	*Blastomyces* (Fig. 53.3 *A*) *Zygomycosis* (Fig. 53.3 *B*) *Coccidiodes* (Fig. 53.3 *C*)
Acid fast stains (AFB, Fite, Kinyoun)	Mycobacteria, Nocardia, (Fite—*Mycobacterium* leprae)	Mycobacteria (Fig. 53.4)
Gram	Most bacteria, including Actinomycetes Also stains fungi, especially on touch preparations	*Nocardia* (Fig. 53.5)
Giemsa (Diff-Quick)	*Ricksettsia, Pneumocystis* (will stain internal structure), fungi	*Histoplasma* (Fig. 53.6*A*) *Pneumocystis* (Fig. 53.6*B*) *Candida* (Fig. 53.6*C*)
Mucicarmine	*Cryptococcus* (mucin positive capsule)	*Cryptococcus* (Fig. 53.7)
Warthin-Starry	Spirochetes, *Bartonella* spp.	

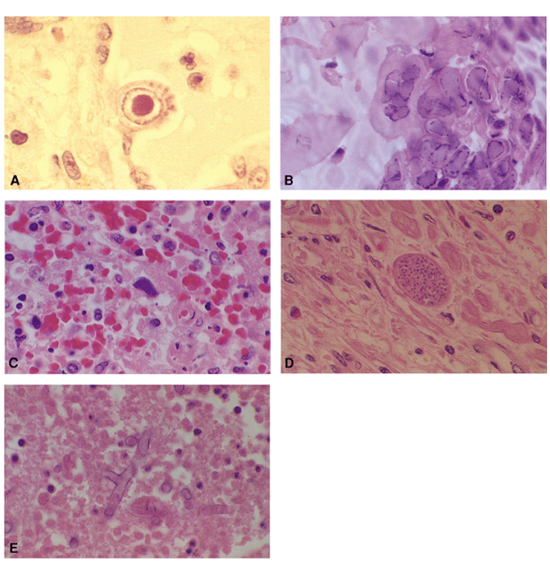


FIGURE 58.1. A: Lung tissue with characteristic “owl's-eye” *cytomegalovirus* nuclear inclusion and cytoplasmic inclusions (H&E, ×1,000, oil). **B:** Esophageal tissue with herpes virus intranuclear inclusions with multinucleation (H&E, ×1,000, oil). **C:** Lung tissue with a “smudged” intranuclear inclusion characteristic of adenovirus (center H&E, ×1,000, oil). **D:** Heart tissue with an intramyocardial *Toxoplasma* cyst (H×E, ×1,000, oil). **E:** Lung tissue with invasive *Aspergillus* (H&E, ×1,000, oil).

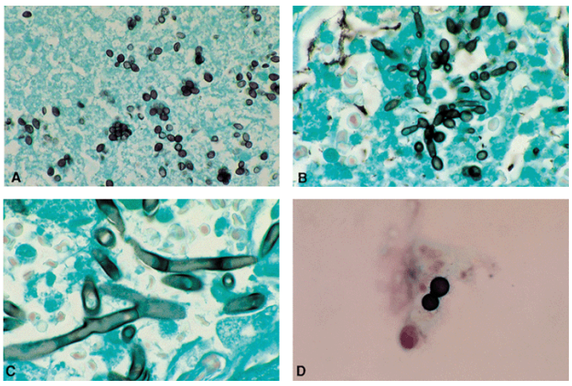


FIGURE 58.2. A: Lung granuloma with *Histoplasma* organisms (GMS, ×1,000, oil). **B:** Lung tissue with invasive *Candida* and characteristic pseudohyphae with pinching at the septae (GMS, ×1,000, oil). **C:** Lung tissue with invasive *Aspergillus* and characteristic true septae (no pinching) and 45° angle branching (GMS, ×1000, oil). **D:** Cytology bronchial specimen with broad-based bud of Blastomyces (GMS, ×1,000, oil).

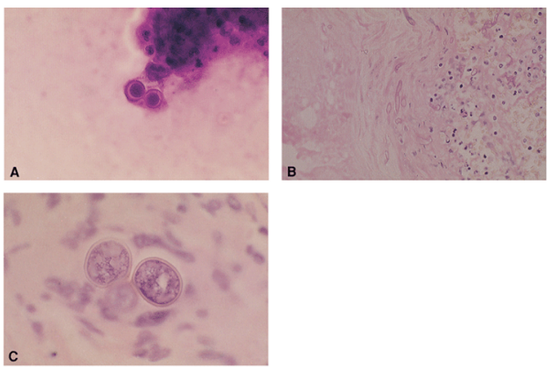


FIGURE 58.3. A: Cytology bronchial specimen with internal structure of *Blastomyces* (PAS, ×1,000, oil). **B:** Sinus vessel wall containing characteristic pleomorphic ribbon appearance and absence of septae in zygomycosis (PAS, ×2,000). **C:** Lung granuloma with *Coccidiodes*spherule (PAS, ×1,000, oil).

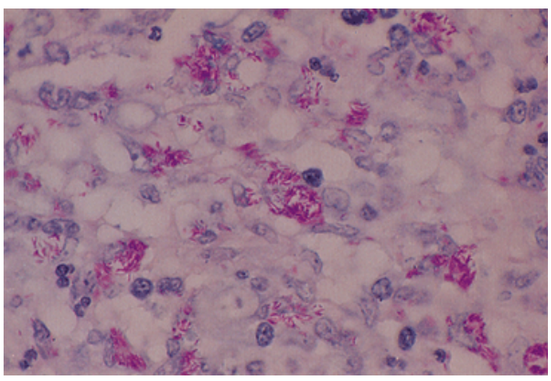


FIGURE 58.4. Lymph node tissue with acid-fast positive rods within histiocytes characteristic of *Mycobacterium avium-intracellulare* (AFB, ×1,000, oil).

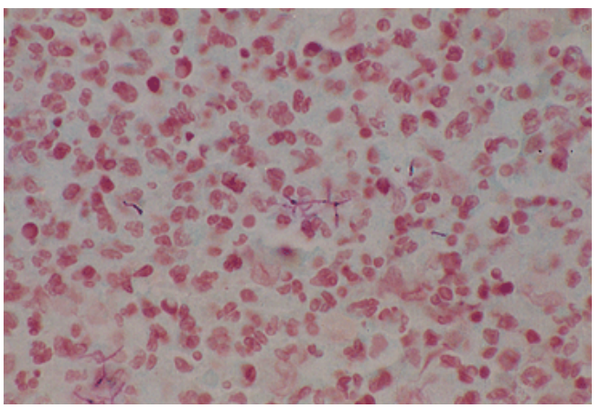


FIGURE 58.5. Skin abscess tissue with Gram-variable, right-angle branching, filamentous bacteria characteristic of *Nocardia* (Gram, ×1,000, oil).

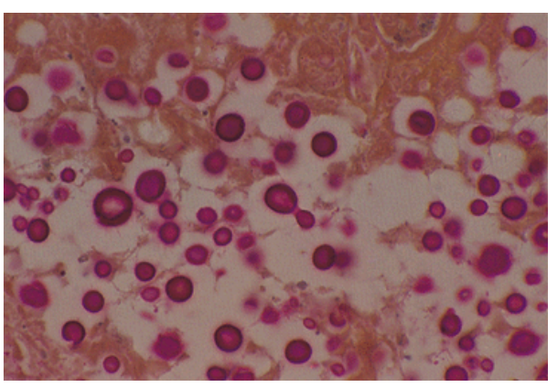


FIGURE 58.7. Liver granuloma with characteristic variably sized budding yeasts surrounded by a mucoid capsule (mucicarmine positive) characteristic of Cryptococcus (mucicarmine, ×1,000, oil).

It is important to remember that histological stains complement each other. The H&E stain provides information about the type of host response, and the Gomori methenamine silver (GMS) stain acts as a good screening tool for the presence of fungi. With the GMS stain, the fungus stains black on a green background, but host response is not demonstrated. Therefore, correlation between the GMS and H&E stains will provide information about both host response and the presence of organisms. A combination stain utilizing GMS with an H&E counterstain will provide both sets of information and, coincidentally, photographs well (Fig. 58.8).

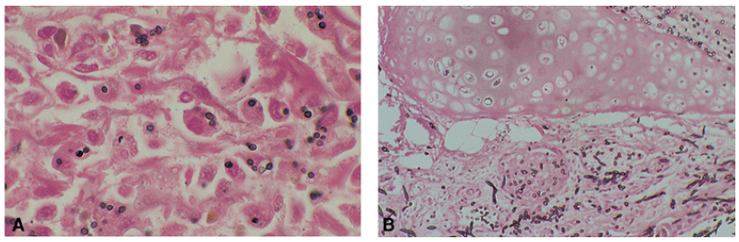


FIGURE 58.8. GMS with H&E counterstain preparation showing: **A:** *Histoplasma* organisms within histiocytes (×400). **B:** *Aspergillus* invading through bronchial cartilage (×200).

INTERPRETATION OF POSTMORTEM CULTURE RESULTS

Part of "58 - Autopsy Microbiology"

A common sense approach to the evaluation of autopsy microbiology includes the correlation between clinical history, culture results, and tissue histology. Determining which results represent contamination versus true infections can be difficult; however, when the patient's clinical history, culture results, and tissue histology are considered jointly, this job becomes much simpler. The literature indicates only fair correlation, at best, between premortem and postmortem cultures. In a study by Wood et al. (12), a positive correlation of 76% was noted when antemortem and postmortem blood cultures were compared with anatomical evidence of infection in 62 autopsy cases. In a study by Koneman and Davis (10) comparing only antemortem and postmortem cultures, the results were less inspiring; no correlation was demonstrated between antemortem and postmortem cultures in 42%, 63%, and 52% of sputum-lung, urine-kidney, and blood-heart cultures, respectively. A more recent publication by Wilson et al. (15) on the utility of postmortem blood cultures created some controversy. The results of their study of 111 autopsies reported that better than half of the positive postmortem blood cultures were from patients whose cause of death was not related to an infection. Conversely, of the 20 patients with antemortem microbemia, seven yielded the same organisms at autopsy, 10 yielded multiple contaminants, and three had blood culture isolates different than antemortem blood cultures. They believe that postmortem blood cultures rarely, if ever, provide information not previously known, or provide any pathophysiologic clarification, or detect errors in therapy or patient management. These results were subsequently challenged by Silver (16) and Nichols (17) in the Letters to the Editor Section of that journal. In practice, the degree of correlation is largely dependent on the ability of the prosector to select the correct areas to sample for culture, to use the correct technique to procure the samples, and to delineate the insignificant results from the clinically important ones. The following sections suggest an approach to the interpretation of positive culture results from blood, tissues, and fluids.

Positive Blood Cultures

1. Interpret positive cultures carefully. Could the organism be a contaminant? Some of the common contaminants in morgues include *Staphylococcus epidermidis, Streptococcus* species, *Corynebacterium* species, *Pseudomonas fluorescens/putida,* and other environmental organisms, such as *Flavobacterium* species. Contamination also is indicated if an excessive number of days (more than 5 to 7) is required for the blood culture bottle to become positive or if only one of two bottles is positive.
2. Did the spleen culture or other tissue cultures grow the same organism as the blood culture? Having the same organism cultured from multiple sites certainly strengthens an argument for sepsis.
3. Could the antemortem clinical course be consistent with sepsis? Patients with sepsis often have a rapid downward course with fever and hypotension.

Positive Tissue or Fluid Cultures

1. Were organisms demonstrated on tissue sections or on Gram stains of homogenized tissue at the time of culture? Given the large number of organisms required to be able to demonstrate even one organism histologically per high power field, it is not surprising that one is frequently unable to demonstrate organisms in tissue stains. However, if organisms that are detected in tissue or on a Gram stain from culture material correlate with the culture results, it will significantly strengthen a diagnosis of infection.
2. Is a host response present that is consistent with the proposed infection? Growth of multiple organisms in a lung culture that is confirmed by tissue Gram stain without evidence of tissue reaction or inflammation most likely indicates agonal aspiration of stomach contents and not a pneumonia. Gross findings of a consolidated lobe of lung with microscopic findings of an acute intraalveolar inflammatory infiltrate and a positive culture of *Streptococcus pneumoniae* are classic findings typical of pneumonia. Limited inflammatory responses may occur in immunosuppressed patients; in these cases, tissue responses such as cell necrosis, fibrinous exudate, and cellular atypia may be indicative of infection.
3. How well do the autopsy findings correlate with the antemortem diagnosis? For example, the isolation of poliovirus from the stool of a patient who died at age 2 of congenital heart disease may raise the question of premortem administration of polio vaccine rather than actual infection. Polio vaccine is a live attenuated vaccine, and virus can be cultured from stool for several weeks in patients receiving the oral vaccine.

By correlating the clinical findings, the autopsy histology findings, and the results of microbiological cultures, under- or overinterpretation of the findings can be avoided. Items that should definitely be addressed in the autopsy report include the confirmation of clinically suspected infections, the adequacy of antemortem antimicrobial therapy, the findings of a clinically unsuspected infection, and the inability to demonstrate a clinically suspected infection.

Methods to determine the adequacy of antimicrobial therapy may include susceptibility testing of the postmortem isolate to the antibiotic given antemortem. This information may be very important epidemiologically, especially in an intensive care unit situation in which multiple patients may be colonized or may later develop infections with the same organisms.

CONCLUSIONS

Part of "58 - Autopsy Microbiology"

The use of microbiological techniques in autopsy pathology can be an important part of the postmortem examination. Infections continue to play a major role in causing significant morbidity and mortality. The confirmation or discovery of infections at autopsy is an important duty of the pathologist. While it can be argued that at times the results of microbiological studies are more confusing than helpful, careful correlation with the patient's clinical history and microscopic pathology often clarifies difficult

cases. It is important to recall that postmortem culturing is affected by many variables, including culturing techniques, gross observational skills, adequate tissue sampling, antimicrobial therapy, possible contamination by normal flora, and the possibility that viscera may not be sterile. Thus, the finding of positive cultures should not be equated with infection unless supported by clinical and histopathological evidence. Understanding the strengths and weaknesses of postmortem culturing will help ensure that results are interpreted correctly. Although postmortem culturing has certain limitations, this should not preclude its use by pathologists.

ACKNOWLEDGMENT

Part of "58 - Autopsy Microbiology"

The authors wish to thank Ms. Linda Schneekloth for her excellent secretarial assistance.

REFERENCES

1. Knapp BE, Kent TH. Postmortem lung cultures. *Arch Pathol* 1968;85:200–203.

2. Minckler TM, Newell GR, O'Toole WF, et al. Microbiology experience in collection of human tissue. *Am J Clin Pathol* 1966;45:85–92.

3. DeJongh DS, Loftis JW, Green GS, et al. Postmortem bacteriology. *Am J Clin Pathol* 1968;49:424–428.

4. O'Toole WF, Saxena HM, Golden A, et al. Studies of postmortem microbiology using sterile autopsy technique. *Arch Pathol* 1965;80:540–547.

5. Kurtin JJ. Studies in autopsy bacteriology. *Am J Clin Pathol* 1958;30:239.

6. Wilson WR, Dolan CT, Washington JA, et al. Clinical significance of postmortem cultures. *Arch Pathol* 1972;94:244–249.

7. Silva H, Sonnenwirth AC. A practical and efficacious method for obtaining significant postmortem blood cultures. *Am J Clin Pathol* 1969;52:433–437.

8. Carpenter HM, Wilkins RM. Autopsy bacteriology: review of 2,033 cases. *Arch Pathol* 1964;77:73–81.

9. Koneman EW. Postmortem bacteriology. *Crit Rev Clin Lab Sci* 1970;1:5–23.

10. Koneman EW, Davis MA. Postmortem bacteriology. III. Clinical significance of microorganisms recovered at autopsy. *Am J Clin Pathol* 1974;61:28–40.

11. Roberts FJ. A review of postmortem bacteriological cultures. *Can Med Ass J* 1969;100:70–74.

12. Wood WH, Oldstone M, Schultz RB. A re-evaluation of blood culture as an autopsy procedure. *Am J Clin Pathol* 1965;43:241–247.

13. Zanen-Lim OG, Zanen HC. Postmortem bacteriology of the lung by print culture of frozen tissue. *J Clin Pathol* 1980;33:474–480.

14. DuMoulin GC, Paterson DG. Clinical relevance of postmortem microbiologic examination: a review. *Hum Pathol* 1985;16:539–548.

15. Wilson SJ, Wilson ML, Reller LB. Diagnostic utility of postmortem blood cultures. *Arch Pathol Lab Med* 1993;117:986–988.

16. Silver H. Postmortem blood cultures (Letters to the editor). *Arch Pathol Lab Med* 1994;118:594–595.

17. Nichols L. Postmortem blood cultures (Letters to the editor). *Arch Pathol Lab Med* 1994;118:594–596.

59

Diagnostic Virology

Karen C. Carroll

James C. Overall Jr.

Viruses continue to be among the most intriguing and important emerging pathogens. Molecular techniques have clearly added to our understanding of viral pathogenesis and have made rapid viral diagnosis and viral quantitation possible. In addition, the past decade has seen an explosion in new antiviral therapies. All these advances should dispel myths that viral diagnosis is neither timely nor clinically useful.

In 1998, 19,916 specimens were submitted for viral culture or antigen detection to the diagnostic virology laboratory at the Associated Regional and University Pathologists, University of Utah Medical Center. Twenty percent were positive, with a mean turnaround time for positive results of 1.4 days. The 20% of specimens positive for a virus compares favorably with the 5% reported rate for positive blood cultures (1). The 1.4-day mean time to positive test results includes confirmed specific identification of herpes simplex virus (HSV), cytomegalovirus (CMV), influenza A and B viruses, respiratory syncytial virus (RSV), adenoviruses, and varicella-zoster virus (VZV) and presumptive identification of an enterovirus. The particular methods used to detect and identify these viruses are described in this chapter. For more technical detail about specific viruses, please see references 2,3,4,5,6,7,8,9,10 and 11.

Table 59.1 lists the antiviral agents [other than those used for the treatment of human immunodeficiency virus type 1 (HIV-1)] licensed for use as of 2001 and the viruses against which the drugs are active. The most recent agents to be approved are the neuraminidase inhibitors for the treatment of influenza A and B (12). Equally exciting is the agent pleconaril, currently in clinical trials, with activity against the enteroviruses (13). Note that the drugs are very selective, i.e., clinically active against only one or a few viruses. Therefore, a specific viral diagnosis must be made to select the proper drug. For more detailed information about antiviral agents, the reader is referred to references 14,15 and 16.

TABLE 59.1. LICENSED ANTIVIRAL AGENTS (NON–HIV-1)—1999

Generic Drug	Brand Name	Viruses	
		Proved	Possible
Amantidine	Symmetrel	Influenza A	
Rimantidine	Flumadine	Influenza A	
Zanamivir	Relenza	Influenza A, B	
Oseltamivir	Tamiflu	Influenza A, B	
Acyclovir	Zovirax	HSV, VZV, CMV	EBV, herpes B
Valacyclovir	Valtrex	HSV, VZV, CMV	
Ganciclovir	Cytovene	CMV	HSV, VZV, EBV, HHV-8, herpes B
Penciclovir		HSV	
Famciclovir	Famvir	HSV, VZV	HBV
Foscarnet	Foscavir	CMV, HSV, VZV including acyclovir resistant	HHV-8, HIV-1
Cidofovir	Vistide	Ganciclovir resistant CMV	Acyclovir resistant HSV, HPV
Ribavirin	Virazole	Lassa fever, hantavirus, HCV	RSV, parainfluenza, influenza A, B, measles, Sin Nombre virus
Lamivudine	Epivir	HBV, HIV	
Interferon alfa	Several	HBV, HCV, HHV-8, HPV	Hepatitis D

HBV, hepatitis B virus; HSV, herpes simplex virus; VZV, varicella-zoster virus; CMV, cytomegalovirus; EBV, Epstein-Barr virus; HHV-8, human herpesvirus type 8; HIV, human immunodeficiency virus; HPV, human papilloma virus; HCV; hepatitis C virus.
Modified from references 12–16.

VALUE OF SPECIFIC VIRAL DIAGNOSIS

Part of "59 - Diagnostic Virology"

There is value in making a specific virologic diagnosis, not only for the individual patient but also for the community and hospital. For the patient, assistance in the decision to initiate antiviral therapy has already been mentioned. In addition, identification of a particular virus in a patient can result in the discontinuation of unnecessary antibiotic therapy and shorten the duration of hospitalization. For example, in the infant admitted with a diagnosis of meningitis and given oral antibiotic therapy before admission, culture of an enterovirus from the cerebrospinal fluid (CSF) can discriminate between (a) partially treated bacterial meningitis requiring as long as 10 days of hospitalization and intravenous antibacterial therapy and (b) viral meningitis, for which patients can be managed symptomatically and often discharged after a few days (17,18). Identification of a viral agent often enables cessation of unnecessary diagnostic tests (e.g., in a patient with a fever of unknown origin) and provides a more accurate prognosis, both acute and long term.

At the community level, identification of a specific viral disease (e.g., measles) can aid in the recognition of illness in subsequent patients and in the institution of preventive measures (e.g., vaccine). In the hospital setting, detection of a particular virus can allow more rational use of infection control measures. For example, influenza virus is known to be transmitted by small particle aerosol, requiring a separate room and respiratory isolation (19). In contrast, RSV is transmitted by direct contact or large droplet aerosol and needs only good hand washing (20) and use of gloves and gown (21). Diagnosis of a particular virus will enable rational use of antiviral prophylaxis in selected situations. For example, if an index case of influenza A is identified in a nursing home patient, use of amantadine or rimantadine in the other residents and staff of the facility for a 2- to 3-week period can prevent serious illness (22).

Finally, at least two studies have demonstrated the cost-effectiveness of rapid diagnosis of respiratory virus infections in hospitalized pediatric patients (23,24). A reduction in antibiotic use was seen in both studies and a decrease in hospital stay and laboratory utilization was found in the study by Woo et al. (24).

APPROACH TO VIRAL DIAGNOSIS

Part of "59 - Diagnostic Virology"

There are two major approaches to the diagnosis of a viral infection: virologic and serologic. The virologic approach includes (a) isolation of infectious virus in cell culture or in laboratory animals, (b) detection of viral antigen by immunologic methods such as immunofluorescence assay (IFA) or enzyme immunoassay (EIA), (c) identification of viral particles by electron microscopy (EM), and (d) detection of viral nucleic acid. The serologic approach includes (a) demonstration of a fourfold or greater rise in antiviral antibody using acute and convalescent sera and a variety

of methods that measure predominantly immunoglobulin (Ig) antibody and (b) demonstration of virus-specific IgM antibody in a single late acute or early recovery phase serum.

As is true for any diagnostic test, the timing, quality, and handling of the specimen are critical for an accurate diagnosis. For the virologic approach to diagnosis, it is important to obtain specimens during the acute infection from the site of disease. For example, samples collected after bacterial and fungal cultures have returned negative results and the patient has begun to recover are unlikely to yield infectious virus. A positive viral culture from a stool specimen in a patient with suspected viral meningitis (instead of CSF) may provide misleading results.

SPECIMEN COLLECTION AND TRANSPORT

Part of "59 - Diagnostic Virology"

The optimal specimens vary depending on the site of disease. Lesions that can be sampled with a swab include vesicles or ulcers on the skin or mucous membranes, the urethra, and the cervix. Several different synthetic swabs, such as Dacron or rayon, on a plastic or aluminum shaft with an accompanying transport tube are commercially available (25,26). For culture of infectious virus, calcium alginate, charcoal-impregnated swabs and swabs on wooden shafts should be avoided (8,25,26). Calcium alginate binds to the virus, rendering it noninfectious, particularly if the swab is left in the transport medium (25,26 and 27). Calcium alginate also inhibits the polymerase chain reaction (PCR) (26).

For pediatric respiratory illnesses, several investigations have shown that the nasopharyngeal (NP) aspirate or nasal wash provides a better yield of virus than a NP swab (28,29 and 30). The NP aspirate is obtained by passing a suction catheter into the posterior nasopharynx, a procedure that usually elicits coughing up secretions from the lower respiratory tract. Gentle suction is applied with a device that allows the respiratory secretions to be delivered into a trap vial or tube. If insufficient respiratory secretions are obtained, 2 mL of nonpreservative saline can be instilled in the opposite nostril, and the suction process repeated. Nasal-wash specimen collection is accomplished by aspirating 3 mL of saline into a suction bulb with a snout. The saline is instilled into one nostril and immediately aspirated back into the suction bulb. In adults, the usual recommended specimens for respiratory viruses are a throat swab or a throat wash with viral transport media (8,11). It might be, however, that a NP aspirate or nasal wash is the preferred specimen in adults also because most respiratory viruses replicate preferentially in columnar epithelial cells located primarily in the posterior NP and in the lower respiratory tract. Aspirates, sputum, bronchoalveolar lavage (BAL) samples in volumes exceeding 0.5 mL can be placed directly in a sterile container and sent to the laboratory. Volumes less than 0.5 mL should be placed into 2 to 3 mL of transport media (8).

For stool specimens, fresh stool is preferable to a rectal swab (8,11). Body fluids other than blood, such as urine specimens and CSF, should be obtained using sterile procedures. Small volumes (0.5 mL) should be placed in viral transport media.

Several viruses are present in the blood during acute infection, and isolation/detection of virus from this body fluid can be an important indicator of disease. Several viruses are associated with white cells, including CMV, HSV, HIV, and the enteroviruses (Coxsackie, echo, polio). HIV is found in lymphocytes and macrophages (31), whereas CMV is associated with neutrophils and, to a lesser extent, macrophages (32). The enteroviruses can be isolated from plasma as well as from white cells (33). The optimal blood fraction for HSV has not been well delineated. Because the particular virus being sought may not be clearly known, it may be optimal to process blood specimens to obtain both neutrophil and mononuclear cell fractions by gravity sedimentation

or by separation techniques that provide both fractions (34). There is limited direct information on the optimal anticoagulant for blood specimens: heparin, ethylenediaminetetraacetic acid (EDTA), or acid citrate dextrose. One investigation demonstrated no difference between heparin and EDTA-anticoagulated blood in the recovery of CMV (34). Recovery rates of HIV from blood were higher with EDTA than heparin-anticoagulated blood (35). Because it is also known that heparin can inactivate the herpesviruses *in vitro* (36), it seems reasonable to recommend the use of EDTA (purple top) tubes. EDTA is also an acceptable transport for molecular assays such as hybridization and PCR.

For small-volume samples and biopsies of tissues or when there is concern about labile infectivity for particular viruses (e.g., RSV or VZV), commercially available tubes or vials with transport media should be used. Ideally, the viral transport system should be isotonic, contain protein, antibiotics, and buffer to control pH and should be nontoxic to cell culture, antigen detection, and molecular assays. Two reports (37,38) evaluate the viability of frequently isolated viruses in several commercially available viral transport systems. Although several systems are comparable in terms of virus stability, Multi-Microbe Media (M-4) (MicroTest, Inc., Snellville, GA) can be used to culture chlamydia, urea plasma, and mycoplasma as well.

The optimal temperature for storage and transport of specimens for viral culture and antigen detection is 4°C (refrigerator or wet ice temperature). Most viruses will be stable for 2 to 5 days at this temperature (25,37,39). Freezing at -20°C (ordinary freezer temperature) destroys or significantly reduces infectivity of virtually all viruses. If specimens must be stored for longer than 3 to 4 days, an ultra-low temperature freezer (-70°C) should be used, and the specimen transported to the virology laboratory on dry ice.

VIROLOGIC METHODS IN VIRAL DIAGNOSIS

Part of "59 - Diagnostic Virology"

Viral Isolation

Isolation of infectious virus in tissue cell cultures remains one of the most commonly used methods to detect viruses. Because not all culturable viruses will grow in a single cell line, several different types of cells are required. The cells that may be used during the winter season to isolate respiratory viruses include primary rhesus monkey kidney, Madin-Darby canine kidney, Hep-2 (a human epithelial carcinoma cell line), buffalo green monkey kidney continuous cell line, and a human fibroblast cell. These are different from the cells used for the summer/fall enterovirus season: primary rhesus monkey kidney, Hep-2, buffalo green monkey kidney, RD (a human rhabdomyosarcoma cell line), and a human fibroblast cell. Many viruses will demonstrate a cytopathic effect (CPE) within a few days, whereas some may take as long as a week or two. Some viruses do not cause typical CPE and are identified by the adsorption of red blood cells to the surface of virus-infected cells in the culture (e.g., hemadsorption for the influenza and parainfluenza viruses). For some viruses, such as several of the Coxsackie A viruses and Colorado tick fever virus, mouse inoculation is required.

With cell culture systems, presumptive identification of a particular virus or virus group (e.g., HSV or RSV or an enterovirus) can be performed within a few or several days based on the CPE: characteristics, time of onset, and cell type in which it occurs. This presumptive identification can be greatly facilitated if the test requisition slip contains the important information requested: source of the specimen (e.g., swab of a genital ulcer or a conjunctival swab rather than an unidentified swab source) and clinical diagnosis (e.g., genital herpes, influenza, aseptic meningitis).

Confirmation of a specific virus or serotype of a virus (e.g., HSV-2, influenza A H3N2, echovirus 11) requires use of immunologic methods with antibody of known specificity. In recent years, fluorescein or peroxidase-conjugated monoclonal antibodies (MABs) have become available commercially to detect viral antigen in cell cultures, e.g., HSV (40,41,42,43 and 44), CMV (45,46,47 and 48), VZV (49), RSV (50,51 and 52), influenza A and B (53), the parainfluenza viruses (54), and the adenoviruses (55). To identify the specific serotype of influenza A or B, inhibition of hemagglutination by specific serotype antisera is used. For confirmation of a specific Coxsackie or echovirus serotype, neutralization of CPE with pools of antisera and then individual antisera is still required (56).

The technique of centrifugation of the patient specimen onto the cell monolayer on a coverslip in the bottom of a shell vial followed by staining for viral antigen with MAB after 1 or 2 days of incubation has significantly reduced the time required to detect and confirm a number of viruses (44,47,52,53,55). The centrifugation step shortens the time required for replication of the virus and production of viral antigen in cell cultures. For more slowly growing viruses such as CMV, the use of MAB against nonstructural proteins produced early in the replication cycle [i.e., immediate early antigen (EA) or EA] allows detection of virus even before CPE can be observed (47). Table 59.2 compares the times to detection for several viruses using the shell vial centrifugation/viral antigen staining (SVC/VAS) method versus conventional CPE. The time to detection and confirmation is much faster with the SVC/VAS method with every virus listed. For example, with CMV, the SVC/VAS method detects 50% of the total positives in 1 day and 90% in 2 days, whereas it requires 10 days to detect 50% with conventional CPE. Many diagnostic virology laboratories, therefore, use the SVC/VAS technique, particularly for detection of CMV.

TABLE 59.2. VIRUS IDENTIFICATION: CENTRIFUGATION PLUS STAIN VERSUS CONVENTIONAL CYTOPATHIC EFFECT

	Centrifugation + Stain		Conventional CPE	
Virus	**Days in Culture**	**Percentage Detected**	**Days in Culture**	**Percentage Detected**
HSV	1	96, 99[a, c]	1	50[b]
CMV	1	55[c]	10	50[c]
CMV	2	95[c]		
VZV	2	90[d]	5	50[d]
Adenovirus	2	97[e]	4	50[d]
Influenza	2	60[f]	4	50[e]

[a] Salmon VC, Turner RB, Speranza MJ, Overall JC Jr. Rapid detection of herpes simplex virus in clinical specimens by centrifugation and immunoperoxidase staining. *J Clin Microbiol* 1986;23:683–686.

[b] Salmon VC, Stanberry LR, Overall JC, Jr. More rapid isolation of herpes simplex virus in continuous line of mink lung cells than in Vero or human fibroblast cells. *Diagn Microbiol Infect Dis* 1984;4:317–24.

[c] Salmon VC, Overall JC. Unpublished observations.

[d] Landry ML. *VA Practitioner*
1988;35.

[e] Espy MJ, Hierholzer JC, Smith TF. The effect of centrifugation on the rapid detection of adenovirus in shell vials. *Am J Clin Pathol* 1987;88:358–60.

[f] Espy MJ, Smith TF, Harmon MW, Kendal AP. Rapid detection of influenza virus by shell vial assay with monoclonal antibodies. *J Clin Microbiol* 1986;24:677–9.

CPE, cytopathic effect; HSV, herpes simplex virus; CMV, cytomegalovirus; VZV, varicella-zoster virus.

The major viruses detected by isolation in cell culture include HSV, CMV, RSV, influenza A and B, the parainfluenza viruses, the adenoviruses, the enteroviruses (Coxsackie A and B, echo, polio), and VZV. Although animal isolation is rarely done at commercial diagnostic virology laboratories, this may be required to isolate some viruses (e.g., Colorado tick fever and many of the Coxsackie A viruses). Animals for isolation of viruses are available at some state health laboratories.

There are several advantages to using isolation of infectious virus as the method of diagnosis. First, a broad range of viruses can be detected. For example, from a respiratory secretion specimen, a number of different respiratory viruses can be isolated, whereas an antigen detection test may identify only one virus (e.g., RSV). Second, isolation of infectious virus makes the agent available for further characterization, such as antiviral susceptibility testing and serotyping/subtyping for epidemiologic purposes. Third, viral isolation is a sensitive and specific standard with which more rapid methods are compared.

There are also several disadvantages to viral isolation. It requires the availability of a cell culture laboratory, with the need for specialized equipment and supplies and trained personnel. Second, isolation is slower than antigen detection, with results in days rather than hours. Finally, specimen collection and transport conditions are more critical for isolation than antigen tests.

Viral Antigen Detection

Viral antigen detection tests may be performed on cell cultures for early detection/confirmation of viral antigen, such as with the SVC/VAS technique previously mentioned, or on specimens directly from patients: nasal or nasopharyngeal secretions, BAL fluids, scrapings of vesicles or conjunctivae, swabs of the cervix or urethra, stool samples, or tissue biopsies. Because viral antigen is present in cells, it is important that specimens from patients contain a sufficient number of cells. Methods to detect viral antigen may be either direct or indirect. With direct methods, fluorescein, an enzyme, or a radiolabel (the indicator system) is conjugated to the antibody used to detect the virus (primary antibody). With indirect methods, the indicator system is conjugated to a secondary antibody (e.g., rabbit anti-mouse), which in turn is directed against the primary antibody (e.g., mouse MAB). The indirect method can be constructed to react with human IgG or IgM molecules for use in antibody assays. The major indicator systems include IFA, EIA, and radioimmunoassay (RIA).

In the EIA procedure, an enzyme (e.g., peroxidase) conjugated to the detector antibody results in the change of a colorless substrate to a colored product that can be read by the naked eye or quantitated in an automated spectrophotometer. The EIA procedure can be adapted to perform (a) an enzyme-linked immunosorbent assay (ELISA) in which an unlabeled antibody (the capture antibody) is bound to a solid phase (e.g., a latex bead or the bottom of a microtiter plate well) to capture the antigen and a second enzyme-labeled antibody is added to detect the captured antigen or (b) immunoperoxidase methods to demonstrate specific viral antigen by light microscopy in tissues from biopsy or autopsy specimens or cells from viral cell cultures. Many commercial ELISA and IFA kits are available to detect viral antigen in patient specimens or cell cultures or IgG or IgM antiviral antibody.

The major viral antigens that can be detected in direct patient specimens include (a) rotavirus and enteric adenovirus in stool specimens; (b) RSV, influenza A and B, the parainfluenza viruses, and the adenoviruses in respiratory specimens; (c) hepatitis B virus (HBV) surface antigen (HBsAg) and HIV p24 antigen in serum; (d) HSV and VZV in vesicle/ulcer swab specimens; and (e) CMV in BAL specimens. The major viruses detected by the SVC/VAS method include CMV, HSV, RSV, the influenza viruses, and adenoviruses.

There are several advantages to viral antigen detection. First, specimen collection and transport conditions are less critical than for viral isolation. Second, antigen detection test results are available sooner than viral isolation, usually within hours. Third, the antigen tests can detect viruses that will not grow in commercially available cell culture systems (e.g., rotavirus, enteric adenovirus, HBV, HIV). Fourth, the antigen detection tests do not require cell culture laboratory equipment and trained personnel for performance. Finally, there may be specimens in which viral infectivity has been lost (negative viral isolation results) but antigenicity has been preserved (true antigen-positive test).

There are also several disadvantages to viral antigen detection. First, there must be a good antiserum and test kit commercially available for the virus sought in the specimen. Examples of medically important viruses for which antigen detection systems are not yet available are Epstein-Barr virus (EBV), hepatitis A and C viruses, the enteroviruses, rubella, mumps, the arboviruses, and parvovirus B19. Second, a separate antigen detection test must be performed for each virus being sought. For example, with respiratory secretions, one test must be performed for RSV, another for influenza A, still another for influenza B, and so on. Finally, depending on the particular virus and antigen detection system available, the antigen test may be less sensitive and is always less specific than viral isolation.

Electron Microscopy

EM remains an important adjunct to other viral diagnostic methods, especially for the identification of agents of viral gastroenteritis. Antigen detection tests are available for two of the three most common causes of viral diarrhea: rotavirus (57,58) and enteric adenovirus (59). Aside from research laboratory serologic tests, alternative diagnostic methods are not yet commercially available for other enteric viral pathogens such as Norwalk virus, astrovirus, calicivirus, or small round viruses and small round structured viruses (60,61). Other uses for EM include detection of polyoma viruses in urine (62,63) and herpesvirus particles in vesicle fluid (HSV or VZV) (64) or in brain tissue (HSV, CMV, EBV), among others (65). Although EM allows detection of viral particles within hours and without the use of specific reagents, there are several disadvantages. A large enough concentration of particles must be present in the specimen to allow detection, and the procedure is expensive and not widely available. More recent methods to enhance EM such as concentrating virus by ultracentrifugation and adding specific antibody to facilitate aggregation of virus particles have been described (63,66). In general, EM is useful when no other viral detection method is available and/or in institutions where there are expertise and interest in EM viral diagnosis.

Nucleic Acid Hybridization and Amplification

Advances in molecular biology in the 1980s led to the development of probe technology using recombinant DNA techniques (67,68). Initial probes were double-stranded DNA containing large numbers of base pairs that led to problems such as nonspecific binding and prolonged incubation periods (7). Since that time, single-stranded nucleic acid probes that hybridize with DNA or RNA targets in clinical samples have been developed. Standardization of hybridization techniques including commercialization of some kits has eliminated the original problems.

After extraction of nucleic acid and denaturation, detection can occur by several hybridization formats: solid phase (slot/spot/dot-blot, microwell/bead capture, Southern/Northern blots), solution phase, and *in situ* hybridization (8,68,69,70 and 71). Solid phase is similar to standard antigen-antibody ELISA formats in which the viral target is immobilized on a membrane or filter and detected by a labeled probe. Initially probes were labeled with ^{32}P, ^{35}S, or ^{131}I. However, biotin and digoxigenin have replaced the radioactive labels. Bound probe is detected by an enzyme-labeled molecule specific for biotin and digoxigenin. In some assays, the substrates are chemiluminescent and emit light, which is detected photographically or is measured by a luminometer when hybridization occurs (72).

The hybridization protection assay (HPA) is an example of liquid-phase hybridization. An oligonucleotide probe labeled with a chemiluminescent acridinium ester is incubated with target nucleic acid. The reaction is subjected to alkaline hydrolysis. Bound probe is protected from hydrolysis, which degrades unbound probe causing it to lose its chemiluminescence. After addition of peroxides, chemiluminescence is detected by a luminometer (73). This method has been developed by GenProbe, Inc. (San Diego, CA) as a means of direct nucleic acid detection or detection of amplified DNA products (74).

The hybrid capture system is a U.S. Food and Drug Administration (FDA)–approved assay produced by Digene Diagnostics, Gaithersburg, MD, which is also based on liquid phase hybridization (75). In the case of human papilloma virus (HPV) detection, DNA is extracted from cells and denatured. Two groups of RNA probes, one containing four low-risk HPV types and the other containing nine high-risk types, are added to tubes coated with alkaline-phosphatase conjugated antibody that captures the RNA/DNA hybrids. Unbound probe is enzymatically degraded. Bound probe is detected with a chemiluminescent substrate (76). Assays for CMV (77) and HBV (78) are also available.

In situ hybridization is performed by hybridizing a labeled probe to fixed paraffin-embedded tissue. An advantage to this technique is the ability to determine the distribution and amount of viral DNA because cellular architecture is preserved. Limitations include cost and complexity of the procedure. Kits for detection of HPV DNA in tissues by *in situ* hybridization are available from several commercial sources.

Probe technology has most often been applied to detection of DNA viruses. However, there are some problems with DNA probe technology. In the absence of some form of amplification, such as PCR, the concentration of viral genomes in direct patient specimens is too low to allow detection with adequate sensitivity. For example, the commercially available probes for HSV and CMV detected only 70% to 90% of virus isolation–positive specimens (40,79).

The remarkable development of PCR technology (80,81,82 and 83), has allowed 10^5- to 10^6-fold amplification of a defined segment of a viral genome (the target). Nucleic acid amplification techniques have exploded in the past decade and are rapidly becoming the diagnostic tests of choice for many viral syndromes. A detailed discussion is beyond the scope of this chapter, and the reader is referred to other chapters in the text and to several reviews (68,69,84,85 and 86).

Amplification techniques can be categorized as (a) target amplification (PCR, transcription–mediated amplification, strand displacement, sequence based amplification); (b) probe amplification (ligase chain reaction, Qβ-replicase); and (c) signal amplification (branched-chain DNA). All techniques involve isolation and denaturation of viral nucleic acid, amplification step(s) and amplicon detection. Commercial assays are available for detection and quantification of hepatitis C virus (HCV) (84,87,88) and HIV-1 (89,90 and 91) and for the detection of nonviral diseases such as chlamydia, gonorrhea, and tuberculosis. For many viral illnesses, “home brew assays” are available in research and reference laboratory settings. Rigorous attention to contamination control is essential and there are evolving guidelines for standardization and quality control (92,93). Interlaboratory performance of these noncommercial assays, however, may vary substantially (94). In addition, clinicians should become aware of some inherent problems with nucleic acid detection, such as false-positive results related to contamination of reagents, samples, and carry-over of amplified product. False-negative results may be owing to the presence of inhibitors, inadequate or poor specimen quality, or poor nucleic acid extraction techniques. In addition, for some diseases, the assays are too sensitive because of indiscriminate detection of dead or live organisms and asymptomatic shedding of some latent viruses.

Despite the above, nucleic acid detection methods are superior to existing diagnostic methods and have emerged as the tests of choice for diagnosis of central nervous system (CNS) infections caused by the herpes viruses and enteroviruses (95,96). The tests are more sensitive and provide more rapid results. Detection of HCV and HIV before seroconversion is another important application (87,89). Finally, quantification of virus (viral load) has provided useful prognostic and therapeutic information in the management of patients with HCV (88), HBV (97), and HIV (90,91) as well as CMV in the susceptible immunocompromised host (98,99). Future technologies will provide improvements in nucleic acid extraction techniques, automation and ability to coamplify multiple targets in a single reaction (84).

SEROLOGIC METHODS IN VIRAL DIAGNOSIS

Part of "59 - Diagnostic Virology"

There are two major uses for serologic methods in viral diagnosis: (a) to diagnose a current or recent acute viral infection and (b) to determine susceptibility or immunity to a particular virus (e.g., VZV in a negative history health care worker exposed to a patient with chickenpox). Measurement of IgG antiviral antibodies is used to determine immunity, whereas quantitation of IgG or IgM antibodies can diagnose current or recent infection. Older serologic methods, such as neutralization, complement fixation (CF), or hemagglutination inhibition (HI), measure only IgG or total antibodies, whereas the newer techniques, such as IFA or ELISA, can be modified to detect either IgG or IgM.

From a clinical or virologic standpoint, there a several situations in which a serologic diagnosis is necessary or more useful than the virologic approach. First, the infectious virus or viral product (antigen, nucleic acid) is not readily available from the patient. A good example would be arbovirus infections of the CNS, in which viremia has cleared by the time the patient presents with encephalitis, and clinicians prefer not to perform a brain biopsy to define the etiologic agent. Second, specimens from the patient that contain virus are available, but the virus will not grow or is difficult to grow in cell culture and tests for viral antigen or nucleic acid are not readily available. Good examples include EBV, hepatitis A (HAV), rubella, and parvovirus B19. Finally, serologic tests are more useful when the incubation period is prolonged (e.g., 3 to 6 weeks) and antibody is already present in the circulation when the patient presents with illness, such as occurs with EBV and CMV mononucleosis and viral hepatitis.

Figure 59.1 shows a typical antibody response to acute measles. A similar response would be expected with virtually any acute, short incubation period (i.e., several days to 2 weeks) viral illness. At the onset of the rash or other symptoms, antibody is undetectable or in low titer. Within 10 days to 2 weeks, appreciable titers of antibody are detectable. With older serologic assays that measure IgG antibody, such as HI or CF, a fourfold or greater rise in titer between acute and convalescent sera (e.g., less than 1:8 to 1:256) enables a diagnosis of measles.

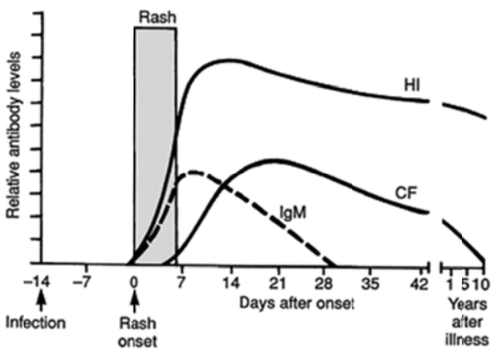


FIGURE 59.1. Antibody responses during acute measles. HI, hemagglutination inhibition antibody; CF, complement fixation antibody. (From Centers for Disease Control. Measles: United States. *MMWR Morb Mortal Wkly Rep* 1982;31:402.)

With IgG antibody assays, a fourfold fall in titer is presumptive evidence of a relatively recent infection, whereas unchanging low titers would be evidence against recent infection and indicate infection in the past (immunity). The presence of high titer antibody in a single convalescent phase serum specimen is not usually sufficient to enable a definitive diagnosis. Note that CF antibody titers rise later, reach lower levels, and may disappear several years after the acute illness. Alternatively, the presence of measles-specific IgM antibody in a single serum obtained 1 to 2 weeks after onset of illness also enables a diagnosis of measles. Typically, the IgM antibody disappears from the serum within a few weeks to a few months after the acute illness.

Many laboratories are now using ELISA kits to measure antibody titers. Results are reported in optical density units rather than dilutions of serum (e g., 1:8, 1:256). The OD units may vary, depending on the particular ELISA kit used and the given virus. To interpret results, therefore, reference range units must be provided by the performing laboratory: no detectable antibody or negative, low, mid, or high positive.

The traditional dogma about the IgM test is that a positive result on a single serum specimen obtained during convalescence is diagnostic of a current or recent infection. Importantly, however, there can be false-positive and false-negative results. False-positive results can occur when there is cross-reactivity, particularly among the herpesviruses (100). Another major factor causing false-positive results is the presence of rheumatoid factor (RF) in the serum (101,102). RF is an IgM antibody that binds to the Fc portion of IgG. So if RF and specific IgG antibody are present in the same serum sample, the IgG binds to the viral antigen and the IgM RF to the IgG, giving a positive test result. Because of the known false-positive IgM test results from RF, several of the commercially available kits either contain or offer reagents that will remove IgG from the serum. With some viruses (e.g., EBV), IgM antibody can persist in the serum of some patients for several months after the acute illness (103). Finally, with viruses known to cause latent infection (e.g., HSV and CMV), reactivation can result in the production of IgM antibody (104,105). So, in these latter two examples, the presence of IgM antibody can represent acute infection months earlier or reactivated latent infection, rather than an acute, recent, primary infection.

False-negative IgM tests can occur because of no or a low IgM response or a delay in the IgM response, particularly in the immunologically immature host (e.g., the neonate with congenital CMV or HIV infection) or in immunosuppressed patients (e.g., HIV-infected patients) (105,106 and 107). False-negative IgM tests can also occur when high-titer IgG antibody in the serum binds to all the viral antigen sites in the kit and prevents adsorption of IgM molecules (competitive inhibition) (101). In summary, although IgM tests may be very useful in the diagnosis of an acute viral infection, both positive and negative results must be interpreted with caution.

The major viruses for which serologic tests provide the most useful or only means for diagnosis of acute infection are listed in Table 59.3.

TABLE 59.3. MAJOR VIRUSES–SEROLOGIC DIAGNOSIS

Virus	Test	Comment
EBV	Monospot; VCA IgM	Nucleic acid detection in research labs
Hepatitis A, B, C, D	ELISA tests IgG, IgM; RIBA confirmation for C	Antigen tests for B, D; qualitative, quantitative nucleic acid detection assays for B, C
Measles, rubella mumps	ELISA tests IgG, IgM	Culture available
Arbovirus	IgG tests	Nucleic acid detection in research labs
Parvovirus B 19	IgG, IgM tests	Nucleic acid detection in research labs
Respiratory viruses	IgG tests	Useful when antigen and culture negative
HIV	ELISA tests; confirmatory Western blot assays	Antigen; Nucleic acid detection methods useful for newborns, acute seroconversion and monitoring viral load
Hantavirus	IgM, IgG tests	Nucleic acid detection in research labs

EBV, Epstein-Barr virus; VCA, viral capsid antigen; Ig, immunoglobulin; ELISA, enzyme-linked immunosorbent assay; HIV, human immunodeficiency virus.

When using IgG antibody tests to determine susceptibility or immunity to a particular virus, it is important to use the most sensitive method to detect antibody. As shown in Fig. 59.1, CF antibody titers are much lower than HI levels and may disappear after several years. Therefore, no detectable CF antibody may not indicate susceptibility. Studies on patients with VZV infections (108) demonstrated this finding.

Some major advantages of serologic diagnosis of viral infections are that the serum is easy to obtain, the serologic tests are widely available, and the specimen can be transported easily. Some major disadvantages are that tests for IgG antibody require acute and convalescent sera, tests for IgM antibody can have false-positive and false-negative results, and, most important, there is usually a delay of 2 to 3 weeks before a diagnosis can be made with short incubation period infections.

OPTIMAL TESTS FOR SPECIFIC VIRUSES

Part of "59 - Diagnostic Virology"

Table 59.4 lists the medically important viruses, together with the recommended tests for viral diagnosis, for (a) acute infection diagnosis—appropriate specimens, the optimal test for rapid and accurate results, and the usual turnaround time and (b) determination of immunity—the preferred antibody assay. The recommended test for diagnosis of acute infection takes into account the shortest turnaround time with acceptable sensitivity (greater than 90%) and specificity (greater than 95%). In general, tests for detection of viral antigen or viral isolation are preferred because of the shorter turnaround time. Serologic tests are used when isolation/antigen detection tests are not available.

TABLE 59.4. MOST RAPID, EXPEDIENT, SENSITIVE, AND SPECIFIC TEST FOR VIRAL DIAGNOSIS

	Diagnose Acute Infection				
Virus Group Virus	Specimen	Test	Time[a]	Determine Immunity[b]	Comments
Herpesviruses					
Herpes simplex virus	Vesicle/ulcer/eye/NP/mouth swab, WBCs, tissue biopsy	Culture in shell vial	16 hr	ELISA FA	Encephalitis-nucleic acid detection becoming standard
	CSF	Nucleic acid detection	6-8 hr		
Cytomegalovirus	Urine, BAL; NP aspirate, tissue biopsy	Culture in shell vial	24-48 hr to 10 d	ELISA FA	Asymptomatic shedding vs. infection
	WBCs	Antigenemia	4-5 hr		
	CSF	Nucleic acid detection	6-8 hr		
Epstein-Barr virus	Serum	Monospot, VCA IgM	1-2 d	VCA	May need to repeat test
Varicella-zoster virus	Vesicle/ulcer swab	Antigen detection by FA	2-4 hr	ELISA, FAMA	CF antibody low sensitivity to determine immunity
		Culture in shell vial	3 d		
Human herpes virus type 6 (HHV-6)	Serum	Anti-HHV-6 IgM	1-3 d	ELISA	
Human herpes virus type 7 (HHV-7)	Serum	Anti-HHV-7 IgM	1-3 d	ELISA	
Human herpes virus type-8 (HHV-8)	Tissue	Histopath			

Respiratory viruses					
Respiratory syncytial virus	NP aspirate Trach aspirate Throat swab	Antigen ELISA DFA Culture	4 hr 6 d	ELISA	Multiplex PCR assay available as RVO assay (Hexaplexi Prodesse, Inc., Wau Kesha, WI)
Influenza A, B	NP aspirate Trach aspirate Throat swab	Antigen DFA Culture	3 hr 7 d	ELISA HI	
Parainfluenza 1, 2, 3	NP aspirate Trach aspirate Throat swab	Antigen DFA Culture	3 hr 7 d	ELISA HI	
Adenovirus	NP aspirate Trach aspirate Throat/eye swab	Culture Antigen DFA, ELISA	7 d 3 hr	ELISA	
Hepatitis viruses					
Hepatitis A (HAV)	Serum	Anti-HAV	2 d	Anti-HAV IgG	Culture not available
Hepatitis B (HBV)	Serum	HBsAG, anti-HBc IgM	2 d	Anti-HBVs	
Hepatitis C (HCV)	Serum	Anti-HCV, ELISA, confirmatory immunoblot	2 d		Quantitative RT-PCR and genotyping useful in patient management
	Serum, EDTA-plasma	RT-PCR	2-5 d		
Hepatitis D (HDV)	Serum	HDV antigen Anti-HDV IgM	2 d		
Hepatitis E (HEV)	Serum	Anti-HEV IgM ELISA			No antibody tests licensed in U.S.
Gastroenteritis viruses					
Rotavirus	Stool	Antigen, ELISA latex	3 hr		Culture not available
Enteric adenovirus	Stool	Antigen, ELISA	3 hr		
Norwalk (SRSV)	Stool	EM	No fixed time period		EM, research ELISA available in some centers
Astroviruses	Stool	EM	No fixed time period		EM only available in some centers
Enteroviruses					
Coxsackie	NP/throat swab, stool	Culture	3-5 d	Neutral	
Echo	WBCs,			antibody	
Polio	CSF	RT-PCR	6-8 hr		
Measles	Serum	Anti-measles IgM, IgG	2 d	ELISA, IFA	Culture available but time consuming
Rubella	Serum	Anti-rubella IgM, IgG	2 d	ELISA, IFA	Culture difficult, time consuming
Mumps	Serum	Anti-mumps IgM, IgG	2 d	ELISA IFA	Culture available but time consuming
Human Immunodeficiency (HIV)	Serum	Anti-HIV ELISA Western immunoblot	2 d 2 d		PCR for proviral DNA, p24 antigen useful for acute disease and neonatal evaluation
	EDTA-plasma	RT-PCR quant (viral load)	2-3 d		
Arbovirus	Serum	Antibody CF	2-5 d	IFA CF	Paired acute and convalescent sera required; PCR available in some centers
Colorado tick fever	Blood clot	Culture:mouse inoculation IgG paired sera	2 wk 2-3 d	Research	Available in state labs; culture is most sensitive, serology, DFA on blood smears less sensitive
Parvovirus B 19	Serum Periph WBCs	Anti-parvo IgM PCR	2-3 d	ELISA, RIA	
Hantavirus	Serum	IgM, IgG	1-2 d		Available in most state labs

[a] Mean time from receipt of specimen in the lab to positive result.

[b] Serum specimen for IgG antibody, using the test method indicated.
NP, nasopharyngeal; WBCs, peripheral white blood cells; ELISA, enzyme-linked immunosorbent assay; FA, immunofluorescence assay (DFA direct FA; IFA, indirect FA); CSF, cerebrospinal fluid; BAL, bronchoalveolar lavage; VCA, viral capsid antigen; Ig, immunoglobulin; FAMA, immunofluorescence assay against membrane antigen; HI, hemagglutination inhibition; HBsAg, hepatitis B surface antigen; HBC, hepatitis B core; EDTA, ethylenediaminetetraacetic acid, anticoagulated peripheral blood for culture of white cells (buffy coat); RT-PCR, reverse transcriptase polymerase chain reaction; EM, electron microscopy; SRSV, small round structured viruses; CF, complement fixation.

With many viral infections, isolation in cell culture has been considered the gold standard. In many publications, therefore, sensitivity and specificity of antigen detection tests have been calculated based on the assumption that the viral isolation data are correct (100% sensitive and 100% specific). However, this may not always be the case because specimen collection, storage, and/or transport conditions may not have been optimal with resultant loss in viral infectivity but preservation of antigenicity. Assuming viral isolation results are the gold standard, therefore, would overestimate the sensitivity of isolation but underestimate the sensitivity of antigen detection. In the results summarized for individual viruses that follow, there has been an attempt (where the reporting of the data allows) to determine total true positive specimens: virus isolation positive plus virus isolation negative but two antigen detection tests positive (e.g., DFA and ELISA). True negative specimens would be the remainder of the samples. Sensitivity and specificity calculations are based on the true positive and true negative specimens.

Although few commercial amplification assays are FDA approved, data supporting their use are available as mentioned previously. Where appropriate, this information has been added to Table 59.4.

Herpes Simplex Virus

The major illnesses associated with HSV are herpes gingivostomatitis (primary oral), herpes labialis (recurrent oral), primary and recurrent genital herpes, primary and recurrent herpes conjunctivitis/keratitis, herpes encephalitis, and neonatal herpes. Most lesions can be sampled with a swab: vesicles or ulcers on the skin or mucous membranes, inflammation of the conjunctiva, cervix, or urethra. Other samples include buffy coat when viremia is suspected (e.g., neonatal herpes, disseminated disease in immunosuppressed hosts), CSF in meningitis complicating primary genital herpes and in neonatal herpes, and tissue biopsies in selected situations. With mucocutaneous vesicles or ulcers, it is important to sample lesions as early as possible in the course of the lesion (109). In our diagnostic virology laboratory, use of the SVC/VAS method with mink lung cell enables detection by the next day (16 to 24 hours) of 96% of the culture positives, with 100% specificity, compared with total cultures positive by CPE over 5 days (44; K. C. Carroll and J. C. Overall, Jr., unpublished observations). Other supportive cell lines include MRC-5, WI-38, rhabdomyosarcoma cells and primary rabbit kidney cells. Routine CPE in a sensitive cell line detects 50% of the positives in 24 hours; 80% in 48 hours; and 95% at 72 hours (110). A novel genetically engineered baby hamster kidney cell line, BHK21ICP6/LacZ (ELVIS, BioWhittaker, Inc., Walkersville, MD), which utilizes a histochemical ß-galactosidase assay, is reported to have a 100% sensitivity for HSV- 1 and -2 detection within 16 to 24 hours compared with detection of positive CPE (111). These cells are specific for HSV and are more costly than typical SVC. Regardless of the chosen culture method, it must be remembered that culture is not 100% sensitive. In a study of genital herpes lesions, culture was 80% sensitive, whereas DFA on cells scraped from the base of lesions was 79% sensitive (112). Both were 100% specific.

Antigen detection systems that can provide results within hours have been marketed. Sensitivity is approximately 80% and specificity 100% compared with culture (113,114). Unfortunately, results with asymptomatic shedding and resolving lesions that contain less virus are not as promising (60% to 75% sensitivity), so these kits would be less useful in situations in which there is a genuine need for rapid and accurate testing for HSV. For example, sampling vaginal secretions at the time of labor assists in decisions about cesarean section to prevent neonatal herpes.

HSV is the most common cause of acute sporadic encephalitis in the United States. In untreated patients, the mortality rate is greater than 70%, and severe neurologic sequelae occur in more than 95% of survivors (115,116 and 117). Intravenous acyclovir therapy has a reduced mortality rate of 30% with 50% of survivors recovering normally (115). Patients with herpes simplex encephalitis (HSVE) typically have focal neurologic clinical features or evidence of focal temporal lobe lesions on neurodiagnostic studies. However, approximately 50% of patients who present with focal disease have diagnoses other than HSVE (e.g., stroke, brain tumor, enteroviral meningoencephalitis, tuberculous meningitis) (118,119).

Diagnosis of HSVE in the past has been problematic. Routine virologic studies of CSF have not been rewarding. HSV has been isolated from only 4% of CSF specimens from patients with biopsy-proven disease (65). Studies examining altered CSF/serum anti-HSV antibody ratios (65,120,121,122 and 123), HSV-specific IgM in the CSF (122), antibody against HSV-1 glycoprotein B in CSF (124,125), and HSV-1 antigen in CSF (126) have yielded results diagnostic of HSV infection of the CNS, but CSF specimens obtained very early in the course of illness were unlikely to be positive. In general, CSF samples obtained 10 days to 2 weeks into the illness were required to have positive results in the majority of patients. Until recently, the most definitive means for establishing a diagnosis of HSVE has been brain biopsy of the involved site with isolation of infectious virus or demonstration of viral antigen (65). Most clinicians opted for empiric acyclovir therapy without a biopsy unless the clinical and neurodiagnostic studies were atypical or the patient worsened on therapy. PCR now provides a noninvasive, rapid, sensitive, and specific assay for diagnosis of HSVE (95,127,128,129,130,131,132,133,134 and 135). Oligonucleotide primers are employed that amplify gene segments from HSV-1 and -2. Although techniques vary among laboratories, amplification of viral DNA is emerging as the gold standard for diagnosis of this disease, as well as neonatal herpes and herpes keratitis. Overall sensitivity and specificity of numerous published studies are 96% and 99%, respectively (132). Quantification of virus may also have prognostic significance and may be useful to monitor response to therapy. In one study, patients with more than 100 copies of HSV DNA per microliter had more severe disease than patients with lower viral loads (133). In addition, mild or atypical cases of HSVE have been increasingly recognized as a result of amplification tests. In a study by Fodor et al. (134), 17% of HSVE patients had mild or atypical disease. Observations included absence of focal findings, subacute course, and initial normal CSF parameters or computed tomography scan. Patients with mild or atypical diseases were more likely to be immunocompromised or to have disease of the nondominant temporal lobe (134). A consensus report providing a strategy for optimal use of PCR in combination with intrathecal antibody measurement is provided in reference 135.

A recent study by Tebas et al. (132) evaluated the cost-effectiveness and patient outcomes of PCR versus empiric acyclovir therapy. When used in conjunction with careful clinical assessment, PCR can lead to good patient outcome and can help decrease costs associated with reduction in acyclovir usage (132).

Although HSV antibody can be demonstrated in the serum of patients who have recovered from primary infection, it has little relevance as a measure of immunity because of latency and reactivation in recurrent lesions of the lips, genitalia, eyes, and other skin sites (136). Conversely, studies of HSV-2 type-specific antibody (directed against glycoprotein G, a type-specific antigen) have provided important information about the frequency of and risk factors for infection with this virus, and the likely occurrence of asymptomatic primary infections (137). Importantly, commercially available IFA and ELISA antibody assays, even when tested against HSV-1 and HSV-2 antigens, do not discriminate cleanly among individuals infected with HSV-1 only, HSV-2 only, or both viruses (138,139). HSV-2 type-specific antibody assays could be used to screen at-risk pregnant women; those seropositive could be tested at the time of labor for shedding of HSV-2 in vaginal secretions (137).

Cytomegalovirus

Significant clinical disease with CMV usually occurs only in immunologically immature hosts (congenital CMV disease) or immunosuppressed patients (pneumonitis, chorioretinitis, colitis, hepatitis) (140,141). Because CMV can be shed asymptomatically for months to years after primary infection (140) and can be reactivated asymptomatically during periods of immunosuppression (142), it is important to be able to distinguish between such asymptomatic shedding and true CMV disease. The major sites for asymptomatic shedding are urine, cervical secretions, semen, and, to a lesser extent, saliva and respiratory secretions (143). Although higher titers of virus are generally present during disease than during asymptomatic shedding, there may be considerable overlap. Therefore, quantitation or semiquantitation of the amount of CMV present in clinical specimens can be helpful (but not absolute) in the diagnosis of CMV disease (144).

The only situation in which isolation of CMV from urine is of undisputed diagnostic value is specimens obtained from neonates during the first several days of life as evidence of congenital infection. In virtually all other situations, it is not possible to be certain whether CMV in urine represents primary infection, reactivation disease, or asymptomatic shedding. Hence, other types of clinical samples have been assessed to determine suitability for predicting CMV disease: BAL specimens, peripheral white blood cells (buffy coat), and plasma.

In immunosuppressed patients, such as organ transplant recipients or patients with the acquired immunodeficiency syndrome (AIDS), diffuse pneumonitis is a frequent opportunistic infection associated with appreciable mortality and morbidity (145,146). Etiologic agents include CMV, *Pneumocystis carinii*, HSV, the respiratory viruses, *Legionella pneumophila*, and a number of other bacterial and fungal agents. Because each agent requires different antimicrobial therapy, it is important to define the etiology. Ganciclovir (GCV) is indicated for treatment of CMV disease in immunosuppressed patients, and therapy of CMV pneumonitis with GCV in combination with CMV hyperimmune intravenous immune globulin (IVIG) has been successful in some centers (147,148,149 and 150). Because GCV has significant adverse effects (151), it is important to determine whether CMV in respiratory secretions represents lung disease or asymptomatic shedding.

Most medical centers perform BAL as the diagnostic procedure of first choice for defining the etiology of pneumonitis in immunosuppressed patients. Centrifugation culture of BAL has reported sensitivities of 70% to 95%, approximating that of culture of lung tissue (144,146,150,152). However, CMV may be isolated in the absence of disease in as many as one third of patients (150). PCR for detection of CMV in BAL samples is more sensitive than virus isolation (150,153). A negative result by PCR strongly supports another etiology for the pneumonia (153). However, the positive predictive value of PCR is low (153). A combination of PCR and immunostaining of alveolar cells may improve the specificity of PCR (150,153). Quantification

of virus in BAL holds promise in solid organ transplant recipients (150,154), but there was no correlation between virus load and pneumonia in bone marrow transplant recipients in one study (155).

As a practical matter, clinicians will administer GCV to severely immunosuppressed patients (such as bone marrow transplant recipients) with pneumonitis and positive CMV cultures from BAL specimens because the adverse consequences of withholding therapy in the presence of true CMV disease are unacceptable. Because mortality of CMV pneumonitis is so high despite GCV therapy, strategies to prevent infection have become the standard of care among bone marrow transplant and solid organ transplant programs (150). Antiviral prophylaxis in the form of either high-dose acyclovir or GCV is given in various regimens within the first 100 days posttransplantation (145,150). Such extensive prophylaxis regimens are costly and not without side effects. A preferred strategy by some programs is the use of early or preemptive therapy based on the detection of CMV infection and treating it before disease is apparent.

In several studies, isolation of CMV from buffy coat has been useful as a predictor of future CMV pneumonitis in organ transplant recipients (142,155). However, there may be both false-negative and false-positive results. For example, in renal transplant recipients, viremia was present in only 67% of patients with CMV disease and was found in 12% of subjects without disease (156,157). Certainly, the correlation of CMV isolation from blood with disease is much better than isolation from throat swabs or urine (142), and the association may be stronger in patients with more severe immunosuppression (e.g., bone marrow and cardiac transplant recipients, AIDS patients) (141,142). Because CMV is found mainly in neutrophils and, to a lesser extent, in monocytes (32), it is important to use leukocyte separation procedures that recover both neutrophils and mononuclear cells (34) in efforts to isolate CMV from the buffy coat. The drawback to using even rapid culture techniques (SVC/VAS) is that a proportion of patients develops disease simultaneously with the appearance of CMV in surveillance cultures.

More recently, investigators demonstrated the diagnostic value of antigenemia or the detection of CMV DNA by PCR as the prompt for beginning antiviral prophylaxis (98,99,158,159,160,161,162 and 163). These techniques are reviewed extensively by Boeckh and Boivin (158). The antigenemia test consists of direct staining of polymorphonuclear neutrophils from the peripheral blood of patients with monoclonal antibodies directed against the lower matrix protein pp65 (99,158,164). Results are reported as the number of antigen-positive cells compared with the number of cells used to make the slide. Quantitation of antigenemia provides an assessment of CMV viral load and seems to correlate well with the results obtained by quantitative PCR (158). Antigenemia has the advantages of a short processing time and some standardization compared with PCR assays and is amenable to testing in a routine clinical laboratory. The disadvantages include the need to process specimens in a timely manner because of loss of leukocyte viability (165), its time-consuming steps, and the difficulty in performing in the early transplant period when the patient is neutropenic. Commercial quantitative amplification assays are in development but none are yet FDA approved (99,158). Numerous studies demonstrate a correlation between high CMV viral load as detected by antigenemia or PCR and a greater risk of progression to disease, especially in HIV-infected patients and solid organ transplant recipients (98,99,161,166,167,168 and 169). However, the threshold for significance may vary depending on the type of transplant and other factors, such as graft versus host disease in the allogeneic bone marrow transplant patient, in whom any detectable level of CMV may be significant. CMV viral load detection may also be useful in monitoring response to therapy. More studies are needed to assess treatment strategies based on quantities of virus in different patient settings.

Testing for CMV IgG antibody has not been particularly useful in the diagnosis of acute or reactivation infection because of the delay in results: seroconversion takes several days to a few weeks after onset of illness (170). In immunologically immature hosts or immunosuppressed patients, the CMV IgM response during acute infection may also be delayed or may not occur at all. For example, CMV IgM antibody was present in the serum of only 70% of neonates proven to be congenitally infected by demonstration of viruria during the first few days of life (105).

The major uses of CMV serology in diagnostic virology laboratories, therefore, appear to be to (a) determine susceptibility or immunity in exposed or potentially exposed health care workers (170) or day care center personnel (171), (b) identify CMV seronegative blood donors for high risk patients (e.g., premature neonates, organ transplant recipients), and (c) to diagnose CMV infections in organ transplant recipients who were seronegative before transplant (172,173). Several of the commercially available ELISA or fluorescence-based IgG tests have acceptable sensitivity and specificity for this use (174,175,176 and 177).

Epstein-Barr Virus

The mainstay in the diagnosis of EBV infections is the Davidsohn modification (178) of the original Paul-Bunnell heterophile antibody test (179). During recovery from EBV infectious mononucleosis (IM), IgM antibodies are produced that react with sheep red cells and cause agglutination. The Davidsohn modification demonstrates differential absorption: IM-associated heterophile antibodies are adsorbed by beef red cells but not guinea pig kidney cells. More recent modifications of the heterophile antibody test employ horse or beef red blood cell antigens, which are more sensitive and specific (103) and are attached to latex or other particles and constructed in a slide or card test format (180,181).

Heterophile antibodies occur in approximately 85% of patients with the IM syndrome (182,183) and usually disappear in a few weeks to a few months. Responses may be delayed in some individuals (182), thus repeat testing may be required to confirm the diagnosis. Importantly, the heterophile test is negative in 70% to 80% of EBV infections in children younger than four years of age (184,185), therefore EBV-specific antibody tests are required for accurate diagnosis in this age group. Also important is the fact that sheep or horse red blood cell heterophile antibodies may persist for more than a year after acute illness in 20% to 70% of patients (103). This persistence of heterophile positivity can result in the erroneous diagnosis of recurrent or chronic IM. In cases of heterophile-negative mononucleosis, CMV is the cause in 70% and

EBV (proven by EBV-specific serology) in 16%, whereas rubella virus, *Toxoplasma gondii*, and adenovirus are rare causes (186).

In several situations, EBV-specific serologic tests may be necessary: (a) when heterophile antibody testing is negative in patients with clinical IM and EBV disease is strongly suspected, (b) in children younger than 5 years of age with IM in whom the heterophile test is insensitive, and (c) where the clinical picture of IM is atypical (182,187).

The following EBV-specific antibody tests are commercially available: IgG and IgM against viral capsid antigens (VCA), anti-EA, and anti-Epstein-Barr nuclear antigen (EBNA). The kinetics of the response to these antigens are shown in Fig. 59.2. The most useful test in the diagnosis of acute IM is VCA IgM; it appears soon after the onset of symptoms, disappears after a few weeks to months, and is 91% to 98% sensitive and 99% specific in the diagnosis (103,182). VCA IgG antibody titers are already elevated when patients become ill and persist for life and therefore are less useful for diagnosis of acute infection. Anti-EA antibodies rise early and disappear in a few months, whereas anti-EBNA appears late and persists for life in individuals who recover. Therefore, in the individual who recovers from IM and is several months post acute illness, there would be moderate antibody titers to VCA IgG and EBNA but low to absent titers against VCA IgM and EA (188).

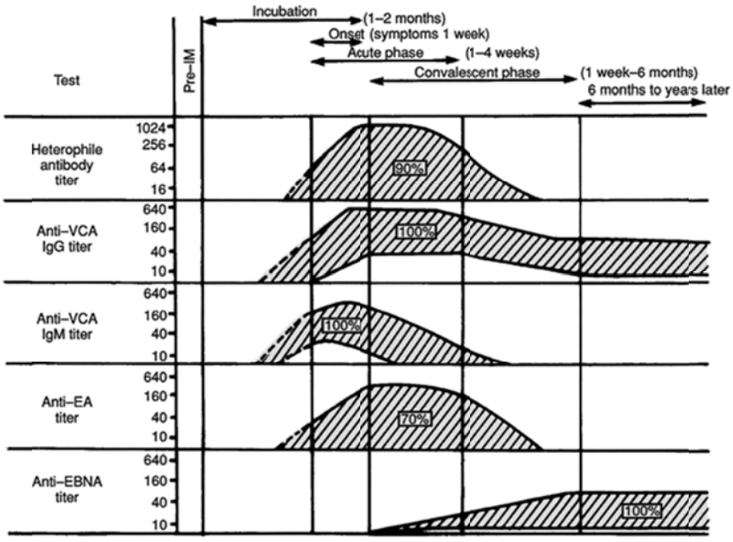


FIGURE 59.2. Antibody responses during EBV-induced infectious mononucleosis. (Modified from Sullivan JL. Epstein-Barr virus and the x-linked lymphoproliferative syndrome. *Adv Pediatr* 1983;30:365.)

In 1985, two papers described what was called the chronic mononucleosis syndrome (189,190). EBV was implicated as a cause based on (a) very high VCA IgG titers, (b) absent to low VCA IgM, (c) high anti-EA, and (d) absent to low anti-EBNA. Subsequent studies demonstrated that EBV plays little or no etiologic role in this chronic fatigue syndrome (CFS) (191,192,193,194 and 195). Many otherwise normal individuals have EBV antibody titers similar to those originally described in chronic mononucleosis (191), and most individuals with CFS do not have the classic chronic EBV titers (192). The cause(s) for the CFS remain(s) elusive (194,195).

Direct tests for EBV, such as cultivation in cord blood leukocytes (196) or detection of the genome by DNA probe (197), are performed in research laboratories but are not yet available in a commercial kit.

EBV may rarely cause or contribute to acute fulminant disease: the X-linked lymphoproliferative syndrome (198) and the virus-associated hemophagocytic syndrome (199). Heterophile and EBV antibodies may be absent, therefore diagnosis depends on demonstration of the virus by culture or DNA probe. *In situ* hybridization against EBV early RNA is the most sensitive method for detection of EBV in pathologic tissue (200).

Molecular techniques such as PCR are not widely available but may be useful in the diagnosis of EBV-associated encephalitis (201). Some investigators demonstrated an increase in EBV levels by quantitative PCR in immunocompromised patients weeks before the development of post-transplantation lymphoproliferative disorder (202).

Varicella-Zoster Virus

The diagnosis of chickenpox or herpes zoster can usually be made clinically. In the selected instances in which laboratory diagnosis is important, virus can be cultured from vesicle fluid and viral antigen can be demonstrated in cells scraped from the base of lesions (49). Infectious virus is quite labile. As much clear vesicle fluid should be aspirated as possible, mixed with a small volume of cell culture media (0.5 to 1.0 mL) and transported to the virology laboratory quickly. Isolation of the virus requires approximately 4 to 6 days. The DFA with MAB to VZV antigen is 92% sensitive (49).

VZV serologic tests are used usually to determine susceptibility or immunity, e.g., in a health care worker who has no history of varicella and is exposed to a patient with chickenpox or who wants to receive the varicella vaccine (203). Interestingly, 90% of such history-negative health care workers have VZV antibodies and are therefore immune (204,205). It is important to note that the CF test is an insensitive method for detecting VZV antibody, as false-negative results occur in an appreciable proportion of cases (108). Several other sensitive methods exist (206). The gold standard fluorescent antibody against membrane antigen test (108) is the most sensitive assay but is not widely available. Latex agglutination or ELISA tests are adequate substitutions (206,207), but at least one paper reports the failure of ELISA assays to detect antibody in patients immunized with the live attenuated vaccine (207). Therefore, an alternative method should be used in persons who have been immunized (207).

Human Herpesvirus Types 6 and 7

Human herpesvirus types 6 (HHV-6) and 7 (HHV-7) are considered emerging viral pathogens (208). Infections with HHV-6 occur in several clinical stages. Primary infection, roseola or exanthem subitum, a febrile illness with rash in infants and young children, was definitely linked to HHV-6 by Yamanishi et al. (209). The second stage of infection occurs in healthy children and adults who secrete virus in the salivary glands but are usually asymptomatic. Rarely, lymphadenopathy or a mononucleosis syndrome has been described (208). The most serious disease associated with HHV-6 occurs in the immunosuppressed patient, especially the transplant recipient (208,210). Reactivation has been associated with pneumonitis, hepatitis, encephalitis, and graft rejection. Synergistic infection with other viruses such as HHV-7 and CMV has resulted in increased severity of disease (208). The literature on the association between HHV-6 and CFS and multiple sclerosis is controversial (208). Like HHV-6, HHV-7 has also been associated with roseola and infections in immunocompromised patients (208,211).

HHV-6 infections can be diagnosed by viral isolation, serologic studies, and nucleic acid amplification techniques in research laboratories (210). Peripheral blood mononuclear cells appear to be the best specimens for virus isolation. Isolation using conventional cell culture is slow and labor intensive. A rapid shell vial assay using MRC-5 cells followed by staining with antiserum specific for the major immediate EA has been described (210).

Diagnosis of both HHV-6 and HHV-7 can also be made by the demonstration of IgG seroconversion or IgM antibody in serum (197,210,211,212 and 213). PCR assays are in development but must be capable of distinguishing actively infected cells from latent virus (210).

Human Herpesvirus Type 8

HHV-8 is the etiologic agent of endemic and AIDS-related Kaposi's sarcoma (214). HHV-8 has also been detected in body-cavity B-cell non-Hodgkin's lymphoma (215) and likely has a role in the pathogenesis of Castleman's disease (216). Diagnosis is typically made based on histopathologic evaluation of biopsied lesions. HHV-8 is difficult to cultivate. No commercial amplification or serologic assays are currently available.

Respiratory Syncytial Virus

RSV causes 50% to 90% of the episodes of bronchiolitis and 50% of the cases of pneumonitis in infants (217). Virtually all infants or toddlers are infected by two years of age, and approximately 1% of infants are hospitalized during the first year of life because of RSV-related disease (217). Infants with cyanotic congenital heart disease and infants with chronic lung disease such as bronchopulmonary dysplasia, particularly if they are oxygen dependent, are at much greater risk for mortality and serious morbidity from RSV bronchiolitis or pneumonitis (218). Diagnosis is usually made based on the characteristic epidemiology and clinical presentation (217,219). However, rapid viral diagnosis may be useful in facilitating appropriate isolation procedures and decisions regarding antiviral treatment and the need for additional laboratory studies, especially in the hospitalized patient (219).

Because culture for RSV requires a mean of 5 to 7 days to positive results, attention has focused on rapid detection of viral antigen in respiratory secretions obtained by NP aspirate or nasal-wash specimens. Two major approaches are available commercially: ELISA and DFA. For the most recently developed or modified ELISA kits, sensitivity is 71% to 98% and specificity 80% to 100%, whereas for DFA, the figures are 83% to 94% and 95% to 98%, respectively (28,50,51 and 52,220). There are membrane filter ELISA kits with individual tests for processing small numbers of specimens that have reported sensitivities of 83% to 100% and specificities of 74% to 96% (220,221). Although serologic tests are available, there is the problem of delay in diagnosis, and the antibody response may be poor or absent in very young infants, immunocompromised patients, and the elderly (219,222).

Influenza Viruses

Influenza is the most common cause for hospitalization for infectious lower respiratory tract disease in adults. It is the major contributor to pneumonia-related deaths in adults, particularly the elderly and individuals with underlying cardiopulmonary disease (19). However, more cases of influenza occur in children than in adults, and children are usually the first cases to occur in communities and families (223). Like RSV, influenza viruses can cause severe disease in children with congenital heart disease and chronic pulmonary disease (224). The availability of the antivirals

amantadine and rimantadine for influenza A and the new class of neuraminidase inhibitors with activity against influenza A and B (12) provide a strong stimulus for rapid viral diagnosis. Therapy must begin within 24 to 48 hours of the onset of symptoms to be maximally effective (22).

As with RSV, attention has turned toward antigen detection for rapid diagnosis of influenza because culture requires a mean time of 3 to 5 days for positive results. The SVC/VAS method has shortened the time for detection of influenza viruses to 48 hours. In early evaluations of this method, sensitivities ranged from 37% to 60% compared with culture (53,225). More recent reports evaluating SVC/VAS using Madin-Darby canine kidney cells demonstrated improved sensitivities of influenza A isolation approaching 100% (226,227).

Initial evaluation of a rapid membrane-bound ELISA for detection of influenza A indicated a sensitivity of 100% and a specificity of 92% compared with culture and DFA (228). More recent results indicated sensitivity of only 50% to 84% but similar specificity (228, 229). Most studies of DFA using commercial reagents for influenza antigen in respiratory secretion cells demonstrated low sensitivity (9% to 59%) but specificity of 95% or greater (54,225,227,230,231). New assays that detect both influenza A and B, the Zstat Flu (ZymeTX, Inc., Oklahoma City, OK) and the BioStar OIA Flu Test (BioStar, Inc., Boulder, CO) do not have appreciably improved sensitivity (60% to 76%) (232,233). Again, serology is available but results in a delay in diagnosis (19).

Other Respiratory Viruses

There are less urgent needs to make a rapid diagnosis of parainfluenza or adenovirus infections because of the lack of proven effective antiviral therapy. The time required for positive cell culture results is 4 to 6 days (55). DFA is available, but the sensitivity is low: 50% to 60% for the parainfluenza viruses and only 10% to 20% for the adenoviruses (54,0230).

A commercial multiplex reverse transcription PCR (RT-PCR)–enzyme hybridization assay (Hexaplex, Prodesse, Inc., Milwaukee, WI) that detects influenza A and B, RSV A and B, and parainfluenza virus types 1 to 3 is available for research use only (234). The cost and length of the assay (6 hours) likely preclude its routine use, but it may be helpful in the diagnosis of these pathogens in the immunocompromised patient with suspected viral pneumonia in whom rapid institution or ribavirin, for example, may be life saving.

Hepatitis Viruses

There are now six known hepatitis viruses—A, B, C, D, E, and G—a veritable alphabet soup of agents encompassing both DNA and RNA viruses. Only the first four are known to cause disease in the United States. Hepatitis viruses can be transmitted by the fecal-oral or percutaneous routes. HAV and hepatitis E virus (HEV) are transmitted primarily by the fecal-oral route. HBV, HCV, HDV, and HVG are blood-borne infections. The annual number of cases of hepatitis reported to the Centers for Disease Control and Prevention (CDC) in 1989 were HAV, 28,500; HBV, 23,200; and non-HAV, non-HBV (mostly C) hepatitis, 2,600 (235). The actual number of cases may be five- to 10-fold higher.

HAV is a member of the Picornaviridae family, genus *Hepatovirus* and is closely related to poliovirus and other enteroviruses (236). Transmission is typically by fecal-oral spread, and outbreaks associated with contaminated food or water (237) or closed settings such as day care centers (238) occur. Most infections occur in young children, but these are usually asymptomatic. HEV, also known as epidemic or enteric-transmitted non-HAV, non-HBV hepatitis, has been classified as a calicivirus (239). HEV is not endemic in the United States. Epidemics in developing countries are primarily related to contaminated water (240).

For HBV and HCV, the major risk factors are similar: intravenous drug abuse, homosexuality, sexual partner of a known chronic carrier, multiple sexual partners, recipient of blood or blood products, and health care workers exposed to blood (235). HBV may be transmitted perinatally (241), but the frequency of maternal fetal/neonatal transmission with HCV is not yet defined clearly (242). Before implementation of universal screening of blood donors for HCV, this virus was the cause of 80% of posttransfusion hepatitis cases (243). Because HDV is known to infect only individuals with acute or chronic HBV infection, the risk factors for infection with HDV are very similar to those for HBV (244).

Chronic hepatitis is known to occur with HBV, HCV, and HDV, but not HAV (235). Cirrhosis is a known sequela of chronic hepatitis with all three viruses (235), but so far only HBV and HCV have been associated with hepatocellular carcinoma (245,246). HGV was discovered in 1996 and is a member of the family Flaviviridae (247). It is acquired primarily by blood transfusions and does not seem to cause symptomatic hepatitis in most patients (247).

Table 59.4 lists the tests to diagnose acute hepatitis and to determine immunity for each of the four hepatitis viruses seen in the United States, as well as HEV because it may occur in returning travelers. Two serologic markers are most often used to diagnose HAV: IgM anti-HAV and total anti-HAV. IgM anti-HAV is a marker of acute disease, usually present just before the onset of symptoms and peaking 1 to 2 weeks later (248). Tests for total anti-HAV detect both IgG primarily, and IgM in both acute and convalescent sera. Samples that are IgM negative but total anti-HAV positive are indicative of past infection or recent vaccination. Total anti-HAV and IgM anti-HAV are usually performed by EIA (248). The reader is referred to reference 248 for a complete description of the various methods available. HEV can likewise be diagnosed by anti-HEV EIA; however, these commercial assays are not yet licensed in the United States (248).

Numerous tests exist for the diagnosis and evaluation of patients with HBV. These include HBsAg, antibody to surface antigen (anti-HBs), total antibody to core antigen (anti-HBc), IgM antibody to core antigen (anti-HBc IgM), HBV e antigen (HBeAg), and antibody to e antigen (anti-HBe). More recently, amplification assays, both qualitative and quantitative, have been introduced to characterize disease status further. In acute and chronic HBV, HBsAg and total anti-HBc are both present. Anti-HBc IgM is generally present in acute hepatitis but absent in chronic disease. Presence of HBeAg and absence of anti-HBe antibody

are markers of greater infectivity (increased likelihood of perinatal, sexual, and blood transfusion transmission) and a worse prognosis (greater risk of progression to chronic hepatitis, cirrhosis, and hepatocellular carcinoma) (249,250 and 251). A typical response when the patient recovers is loss of HBsAg, HBeAg, and HBV DNA and persistence of anti-HBs. With the advent of available molecular techniques, the best predictor of infectivity and viral replication is now detection of HBV DNA. Both molecular hybridization techniques and PCR assays are available (78,97). Quantitative assays are particularly valuable in assessing candidates appropriate for antiviral therapy and in monitoring response to treatment (97).

The remarkable effort to isolate and clone the gene of the major non-HAV, non-HBV (252) and to express the nonstructural protein c100-3 in a bacterial vector (253) led to the identification and characterization of HCV and to the demonstration that this virus was the causative agent in most cases of posttransfusion hepatitis (254). The c100-3 protein was used to develop a first-generation anti-HCV antibody assay. As this assay became more widely used, several limitations became apparent, including a delay in antibody detection for months and as long as a year after the initial infection (243) and false-positive results in individuals with active autoimmune chronic liver disease (255). A second-generation ELISA that incorporates HCV structural proteins appears to be more sensitive and specific, and antibody is detectable sooner after initial infection (256). Development of the confirmatory recombinant immunoblot assay for HCV antibody significantly reduced the problem of false-positive test results (255). Finally, RT-PCR for HCV RNA in serum appears to be the most sensitive and specific test, and several commercial assays are available and widely used in clinical practice (87,88,257). Most qualitative RT-PCR assays detect 100 to 1,000 viral genome copies per milliliter (257) and can be used to diagnose acute and chronic HCV. Quantitative tests are used to assist with predicting the likelihood of response to therapy (257). At least six different genotypes (1 to 6) and more than 90 subtypes of HCV exist. Seventy percent of cases in the United States are associated with genotype 1, with 1a occurring more frequently than 1b (257). Various non-FDA–approved genotyping methods are available in reference laboratories. Genotyping may be beneficial when stratifying treatment regimens because differences do exist with respect to response to therapy among the various genotypes (258).

Gastroenteritis Viruses

Viruses are a major cause of diarrheal disease worldwide. Four viral groups have been identified as established etiologic agents of gastroenteritis. In the United States, rotavirus causes disease primarily in the infant/toddler age group during the winter months (259,260) and accounts for 30% to 50% of pediatric cases of gastroenteritis that require hospitalization (61). Enteric noncultivatable adenoviruses (types 40 and 41) are responsible for 5% to 15% of hospitalized cases in the same age group year round (60,61,259,261). Human astroviruses are found worldwide and in the United States occur primarily in winter, associated with day care center outbreaks among young children (262). *Calicivirus*, Norwalk virus, and Norwalk-like viruses (small round structured viruses) are associated with food and water-borne, nonseasonal outbreaks of gastroenteritis among all age groups (61,63). In the United States, 30% to 50% of outbreaks are believed to be related to this group of viruses (61).

None of the gastroenteritis viruses grows in conventional cell cultures, but all can be detected by EM (60,61,63,66). Commercial ELISA tests with greater than 95% sensitivity and specificity (57,58,59,60 and 61) are available to detect both rotavirus and enteric adenovirus.

Enteroviruses

Nonpolio enteroviruses (Coxsackie A and B, echo) are the most common cause of hospitalization for febrile illness in infants younger than 2 months of age (264) and the most common cause of aseptic meningitis in older infants, children and adults (18,265). The ability to differentiate between bacterial sepsis/meningitis and enteroviral disease may allow the discontinuation of unnecessary diagnostic tests and antibiotic therapy and shortening of the duration of hospitalization (18,266). In addition, antiviral agents for enteroviral infections are currently undergoing clinical trials, increasing the urgency for rapid viral diagnosis (13).

Intensive efforts have been made to develop DNA probes to a conserved portion of the enteroviral genome (267). However, such methods for rapid diagnosis have not achieved sufficient sensitivity, specificity, and practicality for the usual diagnostic virology laboratory. The major means for laboratory diagnosis of enteroviral infections remains isolation of virus in cell culture by examining for viral CPE. Optimal specimens include throat swabs, CSF, and blood samples. There are two major problems with stool and rectal samples. First, enteroviruses may be shed asymptomatically in the stool for weeks after an illness. Second, live oral polio vaccine can be shed in the stool for days to weeks after vaccine administration resulting in a positive enterovirus culture. Therefore, stools should be avoided as clinical specimens for enterovirus culture in infants around the time of oral polio vaccine administration (268). By using four different cell lines for isolation, presumptive diagnosis of Coxsackie A versus Coxsackie B versus echovirus can be made more quickly (269,270 and 271). The Coxsackie B viruses grow best in buffalo green monkey cells and Coxsackie A and echo, in human rhabdomyosarcoma cells (271). The usual time for presumptive diagnosis by CPE is 4 to 7 days. SVC/VAS has shortened this to 72 hours for most cultures (272).

Serologic tests are not usually performed for the diagnosis of enteroviral disease unless there is a need to prove an etiologic role with an isolate from the stool of a patient. In this instance, one must demonstrate a fourfold or greater antibody titer rise in acute and convalescent sera that bracket the illness. At our diagnostic virology laboratory, however, we do offer two sets of neutralizing antibody titers against the enteroviruses: (a) the Coxsackie B group (1 to 6) for use when there are illnesses typically caused by these viruses (e.g., myocarditis, pericarditis, pleurodynia) and (b) the four to five most frequent echovirus serotypes currently causing disease in the community. Again, acute and convalescent sera bracketing the illness are necessary for optimal diagnosis.

RT-PCR is expected to become the diagnostic test of choice for the detection of enteroviruses in the CSF. The sensitivity and specificity of RT-PCR have been well demonstrated in the literature. In contrast to culture, which typically has a sensitivity of 65% to 70%, most amplification assays have sensitivities ranging from 92% to 100% (273,274 and 275). Investigators are applying this technology to blood samples to differentiate enteroviral sepsis from bacterial sepsis in febrile neonates, reducing the need for costly antibiotics and hospital stays in the case of viral disease (276).

Measles, Mumps, and Rubella

Although each of these viruses can be cultured in conventional cell lines, culture requires 7 to 10 days for measles and mumps and more than 3 weeks for rubella (277,278 and 279). With measles virus, the use of the SVC/VAS method with commercially available mouse MAB has resulted in a sensitivity compared with routine culture of 78% at 1 to 2 days and 100% at 5 days (280). Direct staining for measles virus antigen with this same MAB was 100% sensitive compared with culture for NP swab specimens but only 67% sensitive for throat swab and 85% for urine specimens (280).

The usual means for laboratory diagnosis of acute measles, mumps, and rubella is serologic: fourfold or greater rise in IgG antibody in sera bracketing the illness or presence of virus-specific IgM antibody in a single early convalescent serum. Although the traditional serologic test is HI for IgG antibody, several IFA and ELISA IgG and IgM kits are commercially available (281,282,283,284 and 285). Care must be used in interpreting positive IgM tests. The IgM antibody response may be delayed, so a negative test result does not rule out recent infection. If the diagnosis is strongly suspected, the test should be repeated 1 to 2 weeks later. False-positive results may also occur. Mumps IgM antibody has persisted for months after acute illness (284). IgM antibodies that cross-react with rubella have been found in patients with infectious mononucleosis (285), parvovirus B19 infection (286), and CMV infection (287). False-positive rubella IgM tests can be a particular concern in pregnant women (288). It may be wise, therefore, to confirm critical IgM positive results with (a) an IgM assay from another manufacturer or (b) a fourfold rise in IgG antibodies (287,288).

Human Immunodeficiency Virus

HIV is the cause of AIDS. HIV-1 is found worldwide and causes disease characterized by progressive decline in immune function resulting in death. HIV-2 is found primarily in West Africa and has a more benign clinical course. The major diagnostic tests for AIDS are both serologic and virologic: ELISA and Western immunoblot for HIV antibody, ELISA for p24 antigen, culture of peripheral blood mononuclear cells (PBMC) for infectious virus, use of PCR to detect HIV proviral DNA in PBMC, and qualitative and quantitative RT-PCR to detect viral RNA (89,90 and 91).

In most cases, infection is diagnosed by demonstrating HIV antibodies in serum. The adult or older child who presents with features of AIDS has been infected with HIV for weeks to months and already has antibody to the virus in serum. The traditional approach in this situation is to order a screening ELISA, with a repeat ELISA in those with positive results. Current generations of ELISA assays have sensitivities 99% or greater (289) and reported specificities of 99.6% to 100% (290,291 and 292). However, even a false-positive rate of 0.4% in a low prevalence population means that an unacceptable number of patients will be misdiagnosed. Therefore, all positive screening ELISA assays are confirmed by Western blot or IFA. The Western immunoblot measures the antibody response to nine HIV proteins or glycoproteins: gag proteins (p17, p24, p55), pol (p66, p51), and env (gp41, gp120, gp160). Various criteria have been proposed for interpretation of Western blot assays, but those defined by the CDC are the most widely accepted (293). The presence of any two of p24, gp 46, or gp120/160 is a positive test (293). If there is a response to some but not all the HIV antigen bands necessary to satisfy the CDC criteria, the result is said to be indeterminate, and repeat testing over the next 6 months is recommended. If a patient's Western blot results remain indeterminate over a 6-month period, he or she should be considered not infected with HIV-1 (293). In low-risk populations, patients with screening ELISA positive and indeterminate Western blot results are rarely, if ever, infected with HIV on follow-up serum testing (294).

If the results of serologic testing are not definitive for a suspected HIV-positive patient for whom it is important to make a diagnosis quickly (e.g., to institute anti-HIV therapy), there are several additional approaches. Tests may be undertaken for p24 antigen, culture of PBMC for the virus, and nucleic acid detection. Tests for p24 antigen are most useful in the period just before seroconversion in the acutely infected individual and in early diagnosis of the infected neonate (295,296). Culture of virus is available only in research or referral commercial laboratories and takes several days to a few weeks to perform, therefore, results are not readily available for patient management. Finally, nucleic acid amplification assays are available for both HIV-1 detection and quantitation. Proviral HIV DNA can be detected by PCR using primers within conserved sequences in the *gag* or *pol* genes (89,295,296). The reported sensitivity of HIV-1 DNA PCR is 96% to 99% (297).

HIV-1 DNA detection is most useful in early diagnosis in neonates (295,296,298). Several viral load assays that quantify HIV-1 RNA are commercially available. Quantification of HIV-1 RNA is most useful in assessing responses to antiretroviral therapy (90,91). Other applications include assessing prognosis and monitoring the course of disease (90,91).

In the infant with suspected AIDS whose mother is HIV seropositive, the presence of maternal transplacentally transmitted antibody confounds the interpretation of the screening ELISA or the Western immunoblot up to 18 months of age (299). In a child younger than 18 months of age, a definitive diagnosis of HIV disease can be made if the child meets clinical criteria for a diagnosis of AIDS or if the child is positive on two separate occasions utilizing one or more HIV-specific tests: culture, PCR, or p24 antigen (299).

Arbovirus Encephalitis

During 1996 to 1997, the major arboviruses causing encephalitis in the United States were LaCrosse virus, accounting for 88%

of total cases; St. Louis encephalitis virus, 5%; western equine encephalitis, less than 1%; and eastern equine encephalitis virus, 7% (300). Most cases of arbovirus encephalitis occur in the states with the largest mosquito populations: the coastal states, the Great Lakes regions, and the states bordering the Mississippi and Missouri Rivers (301). As the names of the viruses suggest, there is great regional variation in the number of cases and the particular viruses causing disease.

Because viremia has usually cleared by the time patients come for medical attention and because brain biopsy is rarely indicated or performed, the virus or viral product is not available for direct detection. The diagnosis is made by demonstrating seroconversion on acute and convalescent sera bracketing the illness. Available assays include traditional CF tests and ELISA kits (302,303).

Colorado Tick Fever Virus

Colorado tick fever is limited to the Rocky Mountain states and, as the name suggests, is transmitted by the wood tick *Dermacentor andersoni*. Typical features include a biphasic flulike illness without respiratory symptoms and an associated leukopenia. Viremia is prolonged and the virus is associated with red blood cells (304). The most sensitive diagnostic test is mouse inoculation of disrupted blood clot material (305). Demonstration of viral antigen by direct immunofluorescence staining of blood smears and detection of antibody in paired acute and convalescent sera are less sensitive alternatives to mouse inoculation (305).

Parvovirus B19

This virus causes erythema infectiosum (or fifth disease) (306), aplastic crisis in patients with hemolytic anemia syndromes (307), persistent anemia in immunosuppressed patients (308), and abortion and fetal hydrops in pregnant women (309). The virus is not cultivatable with routine methods. Diagnosis of acute infection in immunocompetent patients is best made by demonstration of IgG seroconversion or IgM antibody in serum (310,311). Past infection or immunity is determined by the presence of IgG antibody. Chronic infection in the immunocompromised host is made by detection of parvovirus DNA in serum or peripheral white blood cells (312,313).

Hantaviruses

Among the emerging viruses in the past decade are the new variant hantaviruses. Unlike the Old World or classic hantaviruses that cause primarily hemorrhagic fever with renal syndrome, these new variants are responsible for respiratory illness characterized by adult respiratory distress syndrome (ARDS) (314). The prototypic agent Sin Nombre virus was identified in 1993 after a cluster of unexplained deaths owing to severe respiratory illness in otherwise healthy young adults was described in the Four Corners area of the southwestern United States (315). Since then, several additional viral agents have evolved in the Western Hemisphere, and the associated disease is called hantavirus pulmonary syndrome. The source of infection is a rodent host, and the disease has four phases: (a) a prodrome indistinguishable from several other diseases, (b) pulmonary edema/shock (ARDS), (c) a spontaneous diuretic phase in survivors, and (d) the convalescent phase (316). Hallmarks of the disease during the ARDS phase include thrombocytopenia, hemoconcentration, and leukocytosis (316,317 and 318). Diagnosis is made clinically and usually confirmed by serology. Most state health laboratories have IgM and IgG assays available. Some university centers have RT-PCR available, and immunohistochemical staining can be performed on formalin fixed tissues (317,318).

Perinatal Viral Infections

The major viruses infecting the fetus and newborn infant are CMV, rubella, HSV, enteroviruses, HIV, and HBV. HIV was discussed previously, and HBV rarely results in acute disease in young infants. If we consider the four remaining viruses and the situations in which infection results in disease in the fetus or newborn, CMV and rubella are usually acquired congenitally by transplacental transmission with symptoms present at birth. HSV and enteroviruses are usually acquired during the birth process or postnatally with onset of symptoms several days to a few weeks after birth. Clinicians faced with a suspected perinatal viral infection in a neonate often request TORCH titers (*t*oxoplasma, *o*ther, *r*ubella, *C*MV, *H*SV) on the infant and mother. However, this is not an effective means to establish a rapid diagnosis. Routine TORCH titers measure IgG antibody that is passed transplacentally from mother to fetus. Because this is an active transport mechanism, the titer of virus-specific IgG antibody in the neonate may be twofold and occasionally fourfold higher than in the mother. Antibody present in neonatal serum, therefore, could represent an infection in the mother during this pregnancy or months to years earlier. Thus, the mere presence of antiviral IgG antibody in the serum of a neonate or the mother after birth does not prove current active infection. Only when IgG antiviral antibody persists beyond 4 to 6 months of age in the infant (the usual time for maternal transplacental antibody to disappear) can one assume active infection in the fetus or neonate. The IgG TORCH titer approach, therefore, cannot make a diagnosis for 4 to 6 months.

The virologic approach or culture of virus is the most direct and rapid method for CMV, HSV, and enteroviruses. Suggested specimens include (a) for CMV, urine and buffy coat obtained in the first several days of life; (b) for HSV, swabs of mucocutaneous vesicles or ulcers or conjunctival lesions, CSF, and buffy coat; and (c) for enteroviruses, throat swab, CSF, buffy coat, and stool. Although rubella can be recovered from throat swabs and occasionally CSF (319), the virus isolation process is tedious and can require 3 to 4 weeks for confirmation.

An alternative approach to the diagnosis of congenital viral infections is the demonstration of IgM antiviral antibody in neonatal serum. IgM is not normally passed transplacentally, and its presence indicates current or recent infection in the neonate. As mentioned previously, commercial tests for rubella, CMV, and HSV IgM are available. Because culture of rubella virus is delayed, the demonstration of rubella IgM in a neonate with features consistent with congenital rubella (e.g., small for

gestational age, hepatosplenomegaly, petechiae, jaundice, cataracts, abnormal heart sound) would confirm the diagnosis. However, the anti-CMV IgM is positive in only 50% to 70% of neonates with congenital infection proven by isolation of virus from urine obtained within the first few days of life (105,320). The use of IgM serology for rapid diagnosis of neonatal HSV infections is inappropriate because it may be 2 to 3 weeks after infection before a response is detected (321).

SUMMARY

Part of "59 - Diagnostic Virology"

Rapid diagnosis of infections with all the viruses listed in Table 59.4 requires the use of a variety of virologic and serologic techniques. In general, virologic methods are used when the clinical specimens readily available from the patient contain the virus. Serologic methods are used when virologic methods do not provide timely results or are not available and when the incubation period is long so that the patient already has antibody when presenting with illness. Among the virologic methods, antigen detection tests are preferred because results can be available within hours, but sensitivity and specificity must be sufficiently high to allow accurate conclusions. Finally, isolation of infectious virus is 100% specific and enables further biological characterization of the agent, and, if shell vial culture/viral antigen stain methods are available, can provide results within 1 to 3 days. Nucleic acid detection methods provide a sensitive, rapid, and in many cases cost-effective approach to viral diagnosis. At present, few FDA-approved assays are available except for detection and quantification of HBV, HCV, and HIV. An expanded menu of assays is likely to be available in the near future. It is also important for the clinician to understand the advantages and limitations of the methods used so that results can be properly interpreted.

REFERENCES

1. Roberts FJ, Geere IW, Colman A. A three-year study of positive blood cultures with emphasis on prognosis. *Rev Infect Dis* 1991;13:34–46.

2. Dennehy PH. Rapid diagnosis of pediatric viral infections. *Contemp Pediatr* 1990;7:80–96.

3. Drew WL. Controversies in viral diagnosis. *Rev Infect Dis* 1986; 8:814–824.

4. Smith TF, Wold AD, Espy MJ, et al. New developments in diagnosis of viral diseases. *Infect Dis Clin North Am* 1993;7:183–202.

5. Storch GA. The diagnosis of viral infections (review). *Infect Dis Clin Pract* 1993;2:1–20.

6. Overall JC Jr. Is it bacterial or viral? Laboratory differentiation. *Pediatr Rev* 1993;14:251–261.

7. Lakeman F. Diagnosis of viral infections. In: Richman DD, Whitley RJ, Hayden FG, eds. *Clinical virology*. New York: Churchill-Livingstone, 1997:251–270.

8. Miller MJ, Cherry JD. Use of the diagnostic virology lab. In: Feigin RD, Cherry JD, eds. *Textbook of pediatric infectious diseases*, 4th ed. Philadelphia: Saunders, 1998:2873–2892.

9. Murray PR, Baron EJ, Pfaller MA, et al. *Manual of clinical microbiology*, 7th ed. Washington, DC. American Society for Microbiology, 1999.

10. Lennette EH, Smith TF. *Laboratory diagnosis of viral infections*, 3rd ed. New York. Marcel Dekker, 1999.

11. Lennette EH, Lennette DA, Lennette ET, eds. *Diagnostic procedures for viral, rickettsial and chlamydial infections*. Washington, DC: American Public Health Association, 1995.

12. Calfee DP, Hayden FG. New approaches to influenza chemotherapy. Neuraminidase inhibitors. *Drugs* 1998;56:537–553.

13. Rotbart HA. Enteroviruses: overview of treatments, current and future. *Infect Med* 1999;16(Suppl D):27–32.

14. Hammer SM, Inouye RT. Antiviral agents. In: Richman DD, Whitley RJ, Hayden FG, eds. *Clinical virology*. New York: Churchill-Livingstone, 1997:251–270.

15. Bryson YJ. Antiviral agents. In: Feigin RD, Cherry JD, eds. *Textbook of pediatric infectious diseases*. Philadelphia: Saunders, 1998:2660–2696.

16. Balfour HH. Antiviral drugs. *N Engl J Med* 1999;340:1255–1268.

17. Chonmaitree T, Menegus MA, Powell KR. The clinical relevance of “CSF viral culture”: a two-year experience with aseptic meningitis in Rochester, NY. *JAMA* 1982;247:184–187.

18. Wildin S, Chonmaitree T. The importance of the virology laboratory in the diagnosis and management of viral meningitis. *Am J Dis Child* 1987;141:454–457.

19. Glezen WP, Couch RB. Influenza viruses. In: Evans AS, ed. *Viral infections of humans: epidemiology and control*. New York: Plenum, 1989:419–449.

20. Hall CB, Douglas RG Jr. Modes of transmission of respiratory syncytial virus. *J Pediatr* 1981;99:100–103.

21. Leclair JM, Freeman J, Sullivan BF, et al. Prevention of nosocomial respiratory syncytial virus infections through compliance with glove and gown isolation precautions. *N Engl J Med* 1987;317:329–334.

22. Tominack RL, Hayden FG. Rimantadine hydrochloride and amantadine hydrochloride use in influenza A virus infections. *Infect Dis Clin North Am* 1987;1:459–478.

23. Adcock PM, Stout GG, Hauck MA, et al. Effect of rapid viral diagnosis on the management of children hospitalized with lower respiratory tract infection. *Pediatr Infect Dis J* 1997:16:842–846.

24. Woo PCY, Chiu SS, Seto WH, et al. Cost-effectiveness of rapid diagnosis of viral respiratory tract infections in pediatric patients. *J Clin Microbiol* 1997;35:1579–1581.

25. Johnson FB. Transport of viral specimens. *Clin Microbiol Rev* 1990;3:120–131.

26. Josephson SL. An update on the collection and transport of specimens for viral culture. *Clin Microbiol Newslett* 1997;19:57–61.

27. Bettoli EJ, Brewer PM, Oxtoby MJ, et al. The role of temperature and swab materials in the recovery of herpes simplex virus from lesions. *J Infect Dis* 1982;145:399.

28. Ahluwalia G, Embree J, McNicol P, et al. Comparison of nasopharyngeal aspirate and nasopharyngeal swab specimens for respiratory syncytial virus diagnosis by cell culture, indirect immunofluorescence assay, and enzyme-linked immunosorbent assay. *J Clin Microbiol* 1987;25:763–767.

29. Masters HB, Weber KO, Groothuis JR, et al. Comparison of nasopharyngeal washings and swab specimens for diagnosis of respiratory syncytial virus by EIA, FAT, and cell culture. *Diagn Microbiol Infect Dis* 1987;8:101–105.

30. Treuhaft NIW, Soukup JM, Sullivan BJ. Practical recommendations for the detection of pediatric respiratory syncytial virus infections. *J Clin Microbiol* 1985;22:270–273.

31. McSharry JJ, Costantino R, Robbiano E, et al. Detection and quantitation of human immunodeficiency virus-infected peripheral blood mononuclear cells by flow cytometry. *J Clin Microbiol* 1990;28:724–733.

32. Dankner WM, McCutchan JA, Richman DD, et al. Localization of human cytomegalovirus in peripheral blood leukocytes by in situ hybridization. *J Infect Dis* 1990;161:31–36.

33. Prather SL, Dagan R, Jenista JA, et al. The isolation of enteroviruses from blood: a comparison of four processing methods. *J Med Virol* 1984;14:221–227.

34. Paya CV, Wold AD, Smith TF. Detection of cytomegalovirus from blood leukocytes separated by Sepracell-MN and Ficoll-Paque/Macrodex methods. *J Clin Microbiol* 1988;26:2031–2033.

35. Bayliss GJ, Jesson WJ, Mortimer PP, et al. Cultivation of human immunodeficiency virus from whole blood: effect of anticoagulant and inoculum size on virus growth. *J Med Virol* 1990;31:161–164.

36. Vaheri A, Cantell K. The effect of heparin on herpes simplex virus. *Virology* 1963;21:661–662.

37. Jensen C, Johnson FB. Comparison of various viral transport media for viability maintenance of herpes simplex virus, respiratory syncytial virus and adenovirus. *Diagn Microbiol Infect Dis* 1994;19:137–142.

38. Ogburn JR, Hoffpauir JT, Cole E, et al. Evaluation of new transport medium for detection of herpes simplex virus by culture and direct enzyme-linked immunosorbent assay. *J Clin Microbiol* 1994;32;3082–3084.

39. Yeager AS, Morris JE, Prober CG. Storage and transport of cultures for herpes simplex, virus type 2. *Am J Clin Pathol* 1979;72:977–979.

40. Fung JC, Stanley J, Tilton RC. Comparison of the detection of herpes simplex virus in direct clinical specimens with herpes simplex virus-specific DNA probes and monoclonal antibodies. *J Clin Microbiol* 1985;22:748–753.

41. Johnson FB, Leavitt RW, Richards DF. Comparison of the Scott Selecticult-HSV kit with conventional culture and direct immunoperoxidase staining for detection of herpes simplex virus in cultures of clinical specimens. *J Clin Microbiol* 1985;21:438–441.

42. Lipson SM, Schutzbank TE, Szabo K. Evaluation of three immunofluorescence assays for culture confirmation and typing of herpes simplex virus. *J Clin Microbiol* 1987;25:391–394.

43. Peterson EM, Hughes BL, Aarnaes SL, et al. Comparison of primary rabbit kidney and MRC-5 cells and two stain procedures for herpes simplex virus detection by a shell vial centrifugation method. *J Clin Microbiol* 1988;26:222–224.

44. Salmon VC, Turner RB, Speranza MJ, et al. Rapid detection of herpes simplex virus in clinical specimens by centrifugation and immunoperoxidase staining. *J Clin Microbiol* 1986;23:683–686.

45. Ashley R, Peterson E, Abbo H, et al. Comparison of monoclonal antibodies for rapid detection of cytomegalovirus in spin-amplified plate cultures. *J Clin Microbiol* 1989;27:2858–2860.

46. Gleaves CA, Lee CF, Kirsch L, et al. Evaluation of a direct fluorescein-conjugated monoclonal antibody for detection of cytomegalovirus in centrifugation culture. *J Clin Microbiol* 1987;25:1548–1550.

47. Gleaves CA, Smith TF, Shuster EA, et al. Rapid detection of cytomegalovirus in MRC-5 cells inoculated with urine specimens by using low-speed centrifugation and monoclonal antibody. *J Clin Microbiol* 1984;19:917–919.

48. Randazzo DN, Michalski FJ. Comparison of antibodies for rapid detection of cytomegalovirus. *J Clin Microbiol* 1988;26:369–370.

49. Gleaves CA, Lee CF, Bustamente CI, et al. Use of murine monoclonal antibodies for laboratory diagnosis of varicella-zoster virus infection. *J Clin Microbiol* 1988;261623–1625.

50. Freymuth F, Quibriac M, Petitjean J, et al. Comparison of new tests for rapid diagnosis of respiratory syncytial virus infections by enzyme-linked immunosorbent assay and immunofluorescence techniques. *J Clin Microbiol* 1986;24:1013–1016.

51. Halstead DC, Todd S, Fritch G. Evaluation of five methods for respiratory syncytial virus detection. *J Clin Microbiol* 1990;28:1021–1025.

52. Johnston SLG, Siegel CS. Evaluation of direct immunofluorescence, enzyme immunoassay, centrifugation culture, and conventional culture for the detection of respiratory syncytial virus. *J Clin Microbiol* 1990;28:2394–2397.

53. Espy MJ, Smith TF, Harmon MW, et al. Rapid detection of influenza virus by shell vial assay with monoclonal antibodies. *J Clin Microbiol* 1986;24:677–679.

54. Stout C, Murphy MD, Lawrence S, et al. Evaluation of a monoclonal antibody pool for rapid diagnosis of respiratory viral infections. *J Clin Microbiol* 1989;27:448–452.

55. Espy MJ, Hierholzer JC, Smith TF. The effect of centrifugation on the rapid detection of adenovirus in shell vials. *Am J Clin Pathol* 1987;88:358–360.

56. Lim KA, Benyesh-Melnick M. Typing of enteroviruses by combination of antiserum pools. Application to typing of enteroviruses (Coxsackie and echo). *J Immunol* 1960;84:309–317.

57. Dennehy PH, Gauntlett DR, Tente WE. Comparison of nine commercial immunoassays for the detection of rotavirus in fecal specimens. *J Clin Microbiol* 1988;26:1630–1634.

58. Mathewson JJ, Winsor DK, DuPont HL, et al. Evaluation of assay systems for the detection of rotavirus in stool specimens. *Diagn Microbiol Infect Dis* 1989;12:139–141.

59. Herrmann JE, Perron-Henry DM, Blacklow NR. Antigen detection with monoclonal antibodies for the diagnosis of adenovirus gastroenteritis. *J Infect Dis* 1987;155:1167–1171.

60. Christensen ML. Human viral gastroenteritis. *Clin Microbiol Rev* 1989;2:51–89.

61. Fairchild PG, Blacklow NR. Viral diarrhea. *Infect Clin North Am* 1988;2:677–684.

62. Pappo O, Demetris AJ, Raikow RB, et al. Human polyoma virus infection of renal allografts: histopathologic diagnosis, clinical significance and literature review. *Mod Pathol* 1996;9:105–109.

63. Miller SE. Electron microscopy of viral infections. In: Lennette EH, Smith TF, eds. *Laboratory diagnosis of viral infections*, 3rd ed. New York: Marcel Dekker, 1999:45–70.

64. Kimura A, Tosaka K, Nakao T. An electron microscopic study of varicella skin lesions. *Arch Gesamte Virusforsch* 1972:36:1–12.

65. Nahmias AJ, Whitley RJ, Visintine AN, et al. Collaborative Antiviral Study Group. Herpes simplex virus encephalitis: laboratory evaluations and their diagnostic significance. *J Infect Dis* 1982;145:829–836.

66. Lewis D, Ando T, Humphrey CD, et al. Use of solid-phase immune electron microscopy for classification of Norwalk-like viruses into six antigenic groups from 10 outbreaks in the United States. *J Clin Microbiol* 1995;33:501–504.

67. Engleberg NC. Nucleic acid probe tests for clinical diagnosis—where do we stand? *ASM News* 1991;57:183–186.

68. Wolcott MJ. Advances in nucleic-acid based detection methods. *Clin Microbiol Rev* 1992;5:370–386.

69. Persing DH, Smith TF, Tenover FC, et al., eds. *Diagnostic molecular microbiology, principles and applications*. Washington, DC: American Society for Microbiology, 1993.

70. Tenover FC. Diagnostic deoxyribonucleic acid probes for infectious diseases. *Clin Microbiol Rev* 1988;1:82–101.

71. Warford AL, Levy RA. Use of commercial DNA probes. *Clin Lab Sci* 1989;2:105–108.

72. Matthews JA, Batki A, Hynds C, et al. Enhanced chemiluminescent method for the detection of DNA dot hybridization assay. *Ann Biochem* 1985;151:205.

73. Arnold LJ, Hammond PW, Wiese WA, et al. Assay formats involving acridinium-ester-labeled DNA probes. *Clin Chem* 1989;35:1588–1594.

74. LaRocco MT, Wanger A, Ocera H, et al. Evaluation of a commercial rRNA amplification assay for direct detection of *Mycobacterium tuberculosis* in processed sputum. *Eur J Clin Microbiol Infect Dis* 1994;13:726–731.

75. Lorincz A. Hybrid Capture™ method for detection of human papillomavirus DNA in clinical specimens. *Papillomavirus Rep* 1996;7:1–2.

76. Schiffman MH, Kiviat NB, Burk BD, et al. Accuracy and interlaboratory reliability of human papillomavirus DNA by hybrid capture. *J Clin Microbiol* 1995;33:545–550.

77. Mazulli T, Wood S, Chua R, et al. Evaluation of the Digene Hybrid Capture™ system for the detection and quantitation of human cytomegalovirus viremia in human immunodeficiency virus-infected patients. *J Clin Microbiol* 1996;34:2959–2962.

78. Aspinall S, Steele AD, Peenze I, et al. Detection and quantitation of hepatitis B virus DNA: comparison of two commercial hybridization assays with polymerase chain reaction. *J Viral Hepatol* 1995;2:107–111.

79. Langenberg A, Smith D, Brakel CI, et al. Detection of herpes simplex virus DNA from genital lesions by *in situ* hybridization. *J Clin Microbiol* 1988; 26:933–937.

80. Eisenstein BI. The polymerase chain reaction: a new method of using molecular genetics for medical diagnosis. *N Engl J Med* 1990;322:178–181.

81. Mullis KB, Faloona FA. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods Enzymol* 1987;155:335–350.

82. Oste C. Polymerase chain reaction. *Biotechniques* 1988;6:162–167.

83. Schochetman G, Ou C-Y, Jones WK. Polymerase chain reaction. *J Infect Dis* 1988;158:1154–1157.

84. Mitchell PS, Persing DH. Current trends in molecular microbiology. *Lab Med* 1999;30:263–270.

85. Whelen AC, Persing DH. The role of nucleic acid amplification and detection in the clinical microbiology laboratory. *Annu Rev Microbiol* 1996;50:349–373.

86. Tang YW, Procop GW, Persing DH. Molecular diagnostics of infectious diseases. *Clin Chem* 1997;43:2021–2038.

87. Lunel F, Mariotti M, Cresta P, et al. Comparative study of conventional and novel strategies for detection of hepatitis C RNA in serum: Amplicor, branched-DNA, NASBA and in-house PCR. *J Virol Methods* 1995;54:159–171.

88. Jacob SJ, Baudy D, Jones E, et al. Comparison of quantitative HCV RNA assays in chronic hepatitis C. *Am J Clin Pathol* 1997;107:362–367.

89. Khadir A, Coutlee F, Saint-Antoine P, et al. Clinical evaluation of Amplicor HIV-1 test for detection of human immunodeficiency virus type 1 proviral DNA in peripheral blood mononuclear cells. *J Acquir Immune Defic Syndr Hum Retrovirol* 1995;9:257–263.

90. Mulder J, McKinney N, Christopherson C, et al. Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 in plasma: application to acute retroviral infection. *J Clin Microbiol* 1994;32:292–300.

91. Coste J, Montes B, Reynes J, et al. Comparative evaluation of three assays of the quantitation of human immunodeficiency virus type 1 RNA in plasma. *J Med Virol* 1996;50:293–302.

92. National Committee for Clinical Laboratory Standards. *Molecular diagnostic methods for infectious diseases: Approved guideline MM3-A*. NCCLS. Wayne, PA: 1995.

93. Association for Molecular Pathology. Recommendations for in-house development and operation of molecular diagnostic tests. *Am J Clin Pathol* 1999;111:449–463.

94. Wiedbrauk DL, Hodinka RL, Cunningham W, et al. PCR proficiency testing for HSV: a multicenter study. PASCV Molecular Virology Workshop 1997:25–32.

95. Lakeman FD, Whitley RJ, and the NIAID Collaborative Antiviral Study Group. Diagnosis of herpes simplex encephalitis: application of polymerase chain reaction to cerebrospinal fluid from brain biopsied patients and correlation with disease. *J Infect Dis* 1995;171:857.

96. Lina B, Pozzetto B, Andreoletti L, et al. Multicenter evaluation of a commercially available PCR assay for enterovirus infection in a panel of cerebrospinal fluid specimens. *J Clin Microbiol* 1996;34:3002–3006.

97. Ranki M, Schatzl HM, Zachoval R, et al. Quantification of hepatitis B virus DNA over a wide range from serum for studying viral replicative activity in response to treatment and in recurrent infection. *Hepatology* 1995;21:1492–1499.

98. Bowen EF, Wilson P, Cope A, et al. Cytomegalovirus retinitis in AIDS patients: influence of cytomegaloviral load on response to ganciclovir, time to recurrence and survival. *AIDS* 1996;10:1515–1520.

99. Storch GA, Buller RS, Bailey TC, et al. Comparison of PCR and pp65 antigenemia assay with quantitative shell vial culture for detection of cytomegalovirus in blood leukocytes from solid-organ transplant recipients. *J Clin Microbiol* 1994;32:997–1003.

100. Nielsen CM, Hansen K, Andersen HMK, et al. An enzyme immunoassay for detection of cytomegalovirus IgM antibodies in human serum: specific and nonspecific reaction. *J Med Virol* 1987;7:111–113.

101. Champsaur H, Fattal-German M, Arranhado R. Sensitivity and specificity of viral immunoglobulin M determination by indirect enzyme-linked immunosorbent assay. *J Clin Microbiol* 1988;26:328–332.

102. Salonen EM, Vaheri A, Suni J, et al. Rheumatoid factor in acute viral infections: interference with determination of IgM, IgG and IgA antibodies in an enzyme immunoassay. *J Infect Dis* 1980;142:250–255.

103. Evans AS, Niederman JC, Cenabre LC, et al. A prospective evaluation of heterophile and Epstein-Barr virus-specific IgM antibodies in clinical and subclinical infectious mononucleosis: specificity and sensitivity of the tests and persistence of antibody. *J Infect Dis* 1975;132:546–554.

104. Kuhn JE, Dunkler G, Munk K, et al. Analysis of the IgM and IgG antibody response against herpes simplex virus type 1 (HSV-1): structural and nonstructural proteins. *J Med Virol* 1987;25:135–150.

105. Stagno S, Tinker MK, Elrod C, et al. Immunoglobulin M antibodies detected by enzyme-linked immunosorbent assay and radioimmunoassay in the diagnosis of cytomegalovirus infections in pregnant women and newborn infants. *J Clin Microbiol* 1985;21:930–935.

106. Frickhofen N, Abkowitz JL, Safford M, et al. Persistent B19 parvovirus infection in patients infected with human immunodeficiency virus type 1 (HIV-1): a treatable cause of anemia in AIDS. *Ann Intern Med* 1990;113:926–933.

107. Kahlon J, Whitley RJ. Antibody response of the newborn after herpes simplex virus infection. *J Infect Dis* 1988;158:925–933.

108. Arvin AM. Varicella-zoster virus. In: Fields BN, Knipe DM, Howley PM, et al. eds. *Fields virology*. Philadelphia: Lippincott-Raven Publishers, 1996:2547–2585.

109. Schomogyi M, Wald A, Corey L. Herpes simplex virus-2. An emerging disease? *Infect Dis Clin North Am* 1998;12:47–61.

110. Salmon VC, Stanberry LR, Overall JC Jr. More rapid isolation of herpes simplex virus in a continuous line of mink lung cells than in vero or human fibroblast cells. *Diagn Microbiol Infect Dis* 1984;2:317–324.

111. Stabell EC, O'Rourke SR, Storch GA, et al. Evaluation of a genetically engineered cell line and a histochemical ß-galactosidase assay to detect herpes simplex virus in clinical specimens. *J Clin Microbiol* 1993;31:2796–2798.

112. Lafferty WE, Krofft S, Remington M, et al. Diagnosis of herpes simplex virus by direct immunofluorescence and viral isolation from samples of external genital lesions in a high-prevalence population. *J Clin Microbiol* 1987;25:323–326.

113. Dorian KJ, Beatty E, Atterbury KE. Detection of herpes simplex virus by the Kodak SureCell herpes test. *J Clin Microbiol* 1990;28:2117–2119.

114. Verano L, Michalski FJ. Herpes simplex virus antigen direct detection in standard virus transport medium by DuPont Herpchek enzyme-linked immunosorbent assay. *J Clin Microbiol* 1990;28:2555–2558.

115. Whitley RJ, Alford CA, Hirsch MS, et al. Vidarabine versus acyclovir therapy in herpes simplex encephalitis. *N Engl J Med* 1986;314:144–149.

116. Whitley RJ, Soong SJ, Dolin R, et al. Adenine arabinoside therapy of biopsy-proved herpes simplex encephalitis. *N Engl J Med* 1977;297:289–294.

117. Whitley RJ, Soong SJ, Hirsch MS, et al. Herpes simplex encephalitis: vidarabine therapy and diagnostic problems. *N Engl J Med* 1981;304:313–318.

118. Whitley RJ, Cobbs G, Alford CA, et al. Diseases that mimic herpes simplex encephalitis. *JAMA* 1989;262:234–239.

119. Whitley RJ, Soong SJ, Linneman C Jr, et al. Herpes simplex encephalitis: clinical assessment. *JAMA* 1982;247:317–320.

120. Klapper PE, Laing I, Longson M. Rapid non-invasive diagnosis of herpes encephalitis. *Lancet* 1981;2:607–608.

121. Koskiniemi ML, Vaheri A. Diagnostic value of cerebrospinal fluid antibodies in herpes simplex virus encephalitis. *J Neurol Neurosurg Psychiatry* 1982;45:239–242.

122. Koskiniemi M, Vaheri A, Taskinen E. Cerebrospinal fluid alternations in herpes simplex virus encephalitis. *Rev Infect Dis* 1984;6:608.

123. Levine DP, Lauter CB, Lerner M. Simultaneous serum and CSF antibodies in herpes simplex virus encephalitis. *JAMA* 1978;240:356–360.

124. Grimaldi LME, Roos RP, Manservigi R, et al. An isoelectric focusing study in herpes simplex virus encephalitis. *Ann Neurol* 1988;24:227–232.

125. Kahlon J, Chatterjee S, Lakeman FD, et al. Detection of antibodies to herpes simplex virus in the cerebrospinal fluid of patients with herpes simplex encephalitis. *J Infect Dis* 1987;155:38–44.

126. Lakeman FD, Koga J, Whitley RJ. Detection of antigen to herpes simplex virus in cerebrospinal fluid from patients with herpes simplex encephalitis. *J Infect Dis* 1987;155:1172–1178.

127. Rowley AH, Whitley RJ, Lakeman FD, et al. Rapid detection of herpes-simplex-virus DNA in cerebrospinal fluid of patients with herpes simplex encephalitis. *Lancet* 1990;335:440–441.

128. Aurelius E, Johansson B, Sköldenberg B, et al. Rapid diagnosis of herpes simplex encephalitis by nested polymerase chain reaction assay of cerebrospinal fluid. *Lancet* 1991;337:189–192.

129. Rozenberg F, Lebon P. Amplification and characterization of herpes virus DNA in cerebrospinal fluid from patients with acute encephalitis. *J Clin Microbiol* 1991;29:2412–2417.

130. Aurelius E, Johansson B, Sköldenberg B, et al. Encephalitis in immunocompetent patients due to herpes simplex virus 1 or 2 as determined by type-specific polymerase chain reaction and antibody assays of cerebrospinal fluid. *J Med Virol* 1993;39:179–186.

131. Mitchell PS, Espy MJ, Smith TF, et al. Laboratory diagnosis of central nervous system infections with herpes simplex virus by PCR performed with cerebrospinal fluid specimens. *J Clin Microbiol* 1997;35:2873–2877.

132. Tebas P, Nease RF, Storch GA. Use of the polymerase chain reaction in the diagnosis of herpes simplex encephalitis: A decision analysis model. *Am J Med* 1998;105:287–295.

133. Domingues RB, Lakeman FD, Mayo MS, et al. Application of competitive PCR to cerebrospinal fluid samples from patients with herpes simplex encephalitis. *J Clin Microbiol* 1998;36:2229–2234.

134. Fodor PA, Levin MJ, Weinberg A, et al. Atypical herpes simplex virus encephalitis diagnosed by PCR amplification of viral DNA from CSF. *Neurology* 1998;51:554–559.

135. Cinque P, Cleator GM, Weber T, et al. The role of laboratory investigation in the diagnosis and management of patients with suspected herpes simplex encephalitis: a consensus report. *J Neurol Neurosurg Psychiatry* 1996;61:339–345.

136. Overall JC. Dermatologic viral diseases. In: Galasso GJ, Buchanan RA, Merigan TC, eds. *Antiviral agents and viral diseases of man*. New York: Raven Press, 1984:247–300.

137. Johnson RE, Nahmias AJ, Madger LS, et al. A seroepidemiologic survey of the prevalence of herpes simplex virus type 2 infection in the United States. *N Engl J Med* 1989;321:7–12.

138. Ashley RL, Militoni J, Lee F, et al. Comparison of Western blot (immunoblot) and glycoprotein G-specific immunodot enzyme assay for detecting antibodies to herpes simplex virus types 1 and 2 in human sera. *J Clin Microbiol* 1988;26:662–667.

139. Ashley R, Cent A, Maggs V, et al. Inability of enzyme immunoassays to discriminate between infections with herpes simplex virus types 1 and 2. *Ann Intern Med* 1991;115:520–526.

140. Demmler GJ. Infectious Diseases Society of America and Centers for Disease Control: summary of a workshop on surveillance for congenital cytomegalovirus disease. *Rev Infect Dis* 1991;13:315–329.

141. Jacobsen MA, Mills J. Serious cytomegalovirus disease in the acquired immunodeficiency syndrome (AIDS). *Ann Intern Med* 1988;108:585–594.

142. Dummer JS, White LT, Ho M, et al. Morbidity of cytomegalovirus infection in recipients of heart or heart-lung transplants who received cyclosporine. *J Infect Dis* 1985;152:1182–1191.

143. Drew WL. Diagnosis of cytomegalovirus infection. *Rev Infect Dis* 1988;10(Suppl 3):S476–S486.

144. Emanuel D, Peppard J, Stover D, et al. Rapid immunodiagnosis of cytomegalovirus pneumonia by bronchoalveolar lavage using human and murine monoclonal antibodies. *Ann Intern Med* 1986;104:476–481.

145. Collin BA, Ramphal R. Pneumonia in the compromised host including cancer patients and transplant patients. *Infect Dis Clin North Am* 1998;12:781–805.

146. Shelhammer JH, Gill VJ, Quinn TC, et al. The laboratory evaluation of opportunistic pulmonary infections. *Ann Intern Med* 1996;124:585–599.

147. Emanuel D, Cunningham I, Jules-Elysee K, et al. Cytomegalovirus pneumonia after bone marrow transplantation successfully treated with the combination of ganciclovir and high-dose intravenous immune globulin. *Ann Intern Med* 1988;109:777–782.

148. Reed EC, Bowden RA, Dandliker PS, et al. Treatment of cytomegalovirus pneumonia with ganciclovir and intravenous cytomegalovirus immunoglobulin in patients with bone marrow transplants. *Ann Intern Med* 1988;109:783–788.

149. Snydman DR. Cytomegalovirus immunoglobulins in the prevention and treatment of cytomegalovirus disease. *Rev Infect Dis* 1990;12(Suppl 7):S839–S848.

150. Ljungman P. Cytomegalovirus pneumonia: presentation, diagnosis and treatment. *Semin Respir Infect* 1995;10:209–215.

151. Buhles WC, Mastre BJ, Tinker AJ, et al. Syntex Collaborative Ganciclovir Treatment Study Group Ganciclovir treatment of life- or sight-threatening cytomegalovirus infection: experience in 314 immunocompromised patients. *Rev Infect Dis* 1988;10(Suppl 3):S495–S506.

152. Crawford SW, Bowden RA, Hackman RC, et al. Rapid detection of cytomegalovirus pulmonary infection by bronchoalveolar lavage and centrifugation culture. *Ann Intern Med* 1988;108:180–185.

153. Eriksson BM, Byrtting M, Zweyberg WB, et al. Diagnosis of cytomegalovirus in bronchoalveolar lavage by polymerase chain reaction, in comparison with virus isolation and detection of viral antigen. *Scand J Infect Dis* 1993;25:421–427.

154. Storch GA, Ettinger NA, Ockner D, et al. Quantitative cultures of the cell fraction and supernatant of bronchoalveolar lavage fluid for the diagnosis of cytomegalovirus pneumonitis in lung transplant recipients. *J Infect Dis* 1993;168:1502–1506.

155. Slavin MA, Gleaves CA, Schoch HG, et al. Quantification of cytomegalovirus in bronchoalveolar lavage fluid after allogeneic marrow transplantation by centrifugation culture. *J Clin Microbiol* 1992;30:2776–2779.

156. Meyers JD, Reed EC, Shepp DH, et al. Acyclovir for prevention of cytomegalovirus infection and disease after allogeneic marrow transplantation. *N Engl J Med* 1988;318:70–75.

157. Pass RF, Whitley RJ, Diethelm AG, et al. Cytomegalovirus infection in patients with renal transplants: potentiation by antithymocyte globulin and an incompatible graft. *J Infect Dis* 1980;142:9–17.

158. Boeckh M, Boivin G. Quantitation of cytomegalovirus: methodologic aspects and clinical applications. *Clin Microbiol Rev* 1998;11:533–554.

159. Gerna G, Furione M, Baldanti F, et al. Comparative quantitation of human cytomegalovirus DNA in blood leukocytes and plasma of transplant and AIDS patients. *J Clin Microbiol* 1994;32:2709–2717.

160. Landry ML, Ferguson D. Comparison of quantitative cytomegalovirus antigenemia assay with culture methods and correlation with clinical disease. *J Clin Microbiol* 1993;31:2851–2856.

161. Fox JC, Kidd IM, Griffiths PD, et al. Longitudinal analysis of cytomegalovirus load in renal transplant recipients using a quantitative polymerase chain reaction: correlation with disease. *J Gen Virol* 1995;76:309-319.

162. Gor D, Sabin C, Prentice HG, et al. Longitudinal fluctuation in cytomegalovirus load in bone marrow transplant recipients: relationship between peak virus load, donor/recipient serostatus, acute GVHD and CMV disease. *Bone Marrow Transplant* 1998;21:597–605.

163. Erice A, Holm MA, Gill PC, et al. Cytomegalovirus (CMV) antigenemia assay is more sensitive than shell vial cultures for rapid detection of CMV in polymorphonuclear blood leukocytes. *J Clin Microbiol* 1992;30:2822–2825.

164. Grefte JM, van der Gun BT, Schmolke S, et. al. The lower matrix protein pp65 is the principal viral antigen present in peripheral blood leukocytes during an active cytomegalovirus infection. *J Gen Virol* 1992;73;2923–2932.

165. Boeckh M, Woogerd PM, Stevens-Ayers T, et al. Factors influencing detection of quantitative cytomegalovirus antigenemia. *J Clin Microbiol* 1994;32:832–834.

166. Spector SA, Merrill R, Wolf D, et al. Detection of human cytomegalovirus in plasma of AIDS patients during acute visceral disease by DNA amplification. *J Clin Microbiol* 1992;30:2359–2365.

167. Reynes J, Montes B, Atoui N, et al. Significance of cytomegalovirus (CMV) pp65 antigenemia in the diagnosis of CMV disease in human immunodeficiency virus-infected patients. *J Med Virol* 1996;49:195–198.

168. van den Berg AP, van der Bij W, van Son WJ, et al. Cytomegalovirus antigenemia as a useful marker of symptomatic cytomegalovirus infection after renal transplantation-a report of 130 consecutive patients. *Transplantation* 1989;48:991–995.

169. van der Bij W, Schirm J, Torensma R, et al. Comparison between viremia and antigenemia for detection of cytomegalovirus in blood. *J Clin Microbiol* 1988;26:2531–2535.

170. Balcarek KB, Bagley R, Cloud GA, et al. Cytomegalovirus infection among employees of a children's hospital. *JAMA* 1990;263:840–844.

171. Pass RF, Hutto C. Group day care and cytomegalovirus infections of mothers and children. *Rev Infect Dis* 1986;8:599–605.

172. Adler S. Transfusion-associated cytomegalovirus infections. *Rev Infect Dis* 1983;5:977–993.

173. Bowden RA, Sayers M, Flournoy N, et al. Cytomegalovirus immune globulin and seronegative blood products to prevent primary cytomegalovirus infection after marrow transplantation. *N Engl J Med* 1986;314:1006–1010.

174. Gleaves CA, Wendt SF, Dobbs DR, et al. Evaluation of the CMV-CUBE assay for detection of cytomegalovirus serologic status in marrow transplant patients and marrow donors. *J Clin Microbiol* 1990;28:841–842.

175. Hursh DA, Abbot AD, Sun R, et al. Evaluation of a latex particle agglutination assay for the detection of cytomegalovirus antibody in patient serum. *J Clin Microbiol* 1989;27:2878–2879.

176. Leland DS, Barth KA, Cunningham EB, et al. Evaluation of four methods for cytomegalovirus antibody detection for use by a bone marrow transplantation service. *J Clin Microbiol* 1989;27:176–178.

177. Schaefer L, Cesario A, Demmler G, et al. Evaluation of Abbott CMV-M enzyme immunoassay for detection of cytomegalovirus immunoglobulin M antibody. *J Clin Microbiol* 1988;26:2041–2043.

178. Davidsohn I. Serologic diagnosis of infectious mononucleosis. *JAMA* 1937;108:289–295.

179. Paul JR, Bunnell WW. The presence of heterophile antibodies in infectious mononucleosis. *Am J Med Sci* 1932;183:90–104.

180. Kim M, Wadke M. Comparative evaluation of two test methods (enzyme immunoassay and latex fixation) for the detection of heterophile antibodies in infectious mononucleosis. *J Clin Microbiol* 1990;28:2511–2513.

181. Tilton RC, Dias F, Ryan RW. Comparative evaluation of three commercial tests for detection of heterophile antibody in patients with infectious mononucleosis. *J Clin Microbiol* 1988;26:275–278.

182. Fleisher GR, Collins M, Fager S. Limitations of available tests for diagnosis of infectious mononucleosis. *J Clin Microbiol* 1983;17:619–624.

183. Henle W, Henle GE, Horwitz CA. Epstein-Barr virus specific diagnostic tests in infectious mononucleosis. *Hum Pathol* 1974;5:551–565.

184. Sumaya CV, Ench Y. Epstein-Barr virus infectious mononucleosis in children: I. Clinical and general laboratory findings. *Pediatrics* 1985;75:1003–1010.

185. Sumaya CV, Ench Y. Epstein-Barr virus infectious mononucleosis in children. II. Heterophile antibody and viral-specific responses. *Pediatrics* 1985;75:1011–1019.

186. Horwitz CA, Henle W, Henle G, et al. Heterophile-negative infectious mononucleosis and mononucleosis-like illnesses. *Am J Med* 1977;63:947–951.

187. Sumaya CV. Epstein-Barr virus serologic testing: diagnostic indications and interpretations. *Pediatr Infect Dis* 1986;5:337–342.

188. Schleupner CJ, Overall JC Jr. Infectious mononucleosis and Epstein-Barr virus. *Postgrad Med* 1979;65:83–89, 95–105.

189. Jones JF, Ray CG, Minnich LL, et al. Evidence for active Epstein-Barr virus infection in patients with persistent, unexplained illnesses: elevated anti-early antigen antibodies. *Ann Intern Med* 1985;102:1–7.

190. Straus SE, Tosato G, Armstrong G, et al. Persisting illness and fatigue in adults with evidence of Epstein-Barr virus infection. *Ann Intern Med* 1985;102:7–16.

191. Buchwald D, Sullivan JL, Komaroff AL. Frequency of “chronic active Epstein-Barr virus infection” in a general medical practice. *JAMA* 1987;257:2303–2307.

192. Hellinger WC, Smith TF, Van Scoy RE, et al. Chronic fatigue syndrome and the diagnostic utility of antibody to Epstein-Barr virus early antigen. *JAMA* 1988;260:971–973.

193. Holmes GP, Kaplan JE, Stewart JA, et al. A cluster of patients with a chronic mononucleosis-like syndrome; is Epstein-Barr virus the cause? *JAMA* 1987;257:2297–2302.

194. Straus SE. The chronic mononucleosis syndrome. *J Infect Dis* 1988;157:405–412.

195. Tobi M, Straus SE. Chronic Epstein-Barr virus disease: a workshop held by the National Institute of Allergy and Infectious Diseases. *Ann Intern Med* 1985;103:951–953.

196. Niederman JC, Miller G, Pearson HA, et al. Infectious mononucleosis: Epstein-Barr virus shedding in saliva and the oropharynx. *N Engl J Med* 1976;294:1355–1359.

197. Miller G, Katz BZ, Niederman JC. Some recent developments in the molecular epidemiology of Epstein-Barr virus infections. *Yale J Biol Med* 1987;60:307–316.

198. Grierson H, Purtillo DT. Epstein-Barr virus infections in males with X-linked lymphoproliferative syndrome. *Ann Intern Med* 1987;106:538–545.

199. McClain K, Gehrz R, Grierson H, et al. Virus-associated histiocytic proliferations in children: frequent association with Epstein-Barr virus and congenital or acquired immunodeficiencies. *Am J Pediatr Hematol Oncol* 1988;10:196–205.

200. Ambinder RF, Mann RB. Detection and characterization of Epstein-Barr virus in clinical specimens. *Am J Clin Pathol* 1994;145:230–252.

201. Tselis A, Duman R, Storch GA, et al. Epstein-Barr virus encephalomyelitis diagnosed by polymerase chain reaction: detection of the genome in CSF. *Neurology* 1997;48:1351–1355.

202. Rowe DT, Qu L, Reyes J, et al. Use of quantitative competitive PCR to measure Epstein-Barr virus genome load in the peripheral blood of pediatric transplant patients with lymphoproliferative disorders. *J Clin Microbiol* 1997;35:1612–1615.

203. Centers for Disease Control and Prevention. Prevention of varicella. Update recommendations of the Advisory Committee on immunization practices (ACIP). *MMWR Morb Mortal Wkly Rep* 1999;48:1–5.

204. Krasinski K, Holzman RS, LaCouture R, et al. Hospital experience with varicella-zoster virus. *Infect Control* 1986;7:312–316.

205. McKinney WP, Horowitz MM, Battiola RJ. Susceptibility of hospital-based health care personnel to varicella-zoster virus infections. *Am J Infect Control* 1989;17:26–30.

206. LaRussa P, Steinberg S, Waithe E, et al. Comparison of five assays for antibody to varicella-zoster virus and the fluorescent-antibody-to-membrane-antigen test. *J Clin Microbiol* 1987;25:2059–2062.

207. Steinberg S, Gershon A. Measurement of antibodies to varicella-zoster virus by using a latex agglutination test. *J Clin Microbiol* 1991;29:1527–1529.

208. Campadelli-Fume G, Mirandola P, Menotti L. Human herpesvirus 6: an emerging pathogen. *Emerg Infect Dis* 1999;5:353–366.

209. Yamanishi K, Shiraki K, Kondo T, et al. Identification of human herpesvirus-6 as a causal agent for exanthem subitum. *Lancet* 1988;1:1065–1067.

210. Singh N, Carrigan DR. Human herpesvirus-6 in transplantation: an emerging pathogen. *Ann Intern Med* 1996;124:1065–1071.

211. Tanaka K, Kondo T, Torigoe S, et al. Human herpesvirus 7: another causal agent for roseola (exanthem subitum). *J Pediatr* 1994;125:1–5.

212. Niederman JC, Kaplan MH, Liu CR, et al. Clinical and serological features of human herpesvirus-6 infection in three adults. *Lancet* 1988;2:817–819.

213. Knowles WA, Gardner SD. High prevalence of antibody to human herpesvirus-6 and seroconversion associated with rash in two infants. *Lancet* 1988;2:912–913.

214. Chang Y, Cesarman E, Pessin MS, et al. Identification of herpes-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 1994;266:1865–1869.

215. Cesarman E, Chang Y, Moore PS, et. al. Kaposi's sarcoma-associated herpes-virus like DNA sequences in AIDS-related body cavity-based lymphomas. *N Engl J Med* 1995;332:1186–1191.

216. Soulier J, Grollet L, Oskenhendler E, et al. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood* 1995;86:1276–1280.

217. Hall CB. Respiratory syncytial virus. In: Feigin RD, Cherry JD, eds. *Textbook of pediatric infectious diseases*. Philadelphia: Saunders, 1998:2084–2111.

218. MacDonald NE, Hall CB, Suffin SC, et al. Respiratory syncytial viral infection in infants with congenital heart disease. *N Engl J Med* 1982;307:397–400.

219. McCarthy CA, Hall CB. Recent approaches to the management and prevention of respiratory syncytial virus infection. *Curr Clin Topics Infect Dis* 1998;18:1–18.

220. Kellogg JA. Culture vs direct antigen assays for detection of microbial pathogens from lower respiratory tract specimens suspected of containing the respiratory syncytial virus. *Arch Pathol Lab Med* 1991;115:451–458.

221. Thomas EE, Book LE. Comparison of two rapid methods for detection of respiratory syncytial virus (RSV) (TestPack RSV and Ortho RSV ELISA) with direct immunofluorescence and virus isolation for the diagnosis of pediatric RSV infection. *J Clin Microbiol* 1991;29:632–635.

222. Welliver RC. Detection, pathogenesis, and therapy of respiratory syncytial virus infections. *Clin Microbiol Rev* 1988;1:27–39.

223. Glezen WP, Decker M, Joseph SW, et al. Acute respiratory disease associated with influenza epidemics in Houston, 1981–1983. *J Infect Dis* 1987;155:1119–1126.

224. Glezen WP, Decker M, Perrotta DM. Survey of underlying conditions of persons hospitalized with acute respiratory disease during influenza epidemics in Houston, 1978–1981. *Am Rev Respir Dis* 1987;136:550–555.

225. Johnston SLG, Siegel CS. A comparison of direct immunofluorescence, shell vial culture, and conventional cell culture for the rapid detection of influenza A and B. *Diagn Microbiol Infect Dis* 1991;14:131–134.

226. Reina J, Fernandez-Baca V, Blanco I, et al. Comparison of Madin-Darby canine kidney cells (MDCK) with a green monkey continuous cell line (Vero) and Human lung embryonated cells (MRC-5) in the isolation of influenza A virus from nasopharyngeal aspirates by shell vial culture. *J Clin Microbiol* 1997;35:1900–1901.

227. Reina J, Munar M, Blanco I. Evaluation of a direct immunofluorescence assay, dot-blot enzyme immunoassay and shell vial culture in the diagnosis of lower respiratory tract infections caused by influenza A virus. *Diagn Microbiol Infect Dis* 1996;25:143–145.

228. Waner JL, Todd SJ, Shalaby H, et al. Comparison of Directigen FLU-A with viral isolation and direct immunofluorescence for the rapid detection and identification of influenza A virus. *J Clin Microbiol* 1991;29:479–482.

229. Ryan-Poirier KA, Katz JM, Webster RG, et al. Application of Directigen FLU-A for the detection of influenza A virus in human and nonhuman specimens. *J Clin Microbiol* 1992;30:1072–1075.

230. Salomon HE, Grandien M, Avila MM, et al. Comparison of three techniques for detection of respiratory viruses in nasopharyngeal aspirates from children with lower acute respiratory infections. *J Med Virol* 1989;28:159–162.

231. Takimoto S, Grandien M, Ishida MA, et al. Comparison of enzyme-linked immunosorbent assay, indirect immunofluorescence assay, and virus isolation for detection of respiratory viruses in nasopharyngeal secretions. *J Clin Microbiol* 1991;29:470–474.

232. Zstat Flu Product Package Insert. ZymeTx, Inc. Oklahoma City, OK, 1998.

233. Gander RM, Reddy ST, Patel SK. Clinical evaluation of the BioStar [RM] AB FLU OIA [RM] Test for direct detection of Influenza types A and B in respiratory secretions. Abstracts of the 38th Annual ICAAC, 1998;H-59:332.

234. Fan J, Henrickson J, Savatski LL. Rapid simultaneous diagnosis of infections with respiratory syncytial viruses A and B, and human parainfluenza virus types 1,2, and 3 by multiplex quantitative reverse transcription-polymerase chain reaction-enzyme hybridization assay (Hexaplex). *Clin Infect Dis* 1998;26:1397–1402.

235. Centers for Disease Control. Protection against viral hepatitis. *MMWR Morb Mortal Wkly Rep* 1990;39:1–26.

236. Weitz M, Siegl G. Hepatitis A: structure and molecular virology. In: Zuckerman A, Thomas H, eds. *Viral hepatitis*. Edinburgh: Churchill-Livingstone 1993:21–33.

237. Kosatsky T, Middaugh JP. Linked outbreaks of hepatitis A in homosexual men and in food service patrons and employees. *West J Med* 1986;144:307–310.

238. Hadler SC, McFarland L. Hepatitis in day care centers: epidemiology and prevention. *Rev Infect Dis* 1986;8:548–557.

239. Cubitt D, Bradley D, Carter M, et al. Caliciviridae, virus taxonomy. Arch Virol 1995;10(Suppl):359–363.

240. Bile K, Isse A, Mohamud O, et al. Contrasting roles of rivers and wells as sources of drinking water on attack and fatality rates in a hepatitis E epidemic in Somalia. *Am J Trop Med Hyg* 1994;51:466–474.

241. Schalm SW, Mazel A, de Gast GC, et al. Prevention of hepatitis B infection in newborns through mass screening and delayed vaccination of all infants of mothers with hepatitis B surface antigen. *Pediatrics* 1989;83:1041–1047.

242. Centers for Disease Control. Public Health Service inter-agency guidelines for screening donors of blood, plasma, organs, tissues, and semen for evidence of hepatitis B and hepatitis C. *MMWR Mort Morbid Wkly Rep* 1991;40:1–17.

243. Alter HJ, Purcell RH, Shih JW, et al. Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. *N Engl J Med* 1989;321:1494–1500.

244. Chatzinoff M, Friedman LS. Delta agent hepatitis. *Infect Dis Clin North Am* 1987;1:529–545.

245. Alward WLM, McMahon BJ, Hall DB, et al. The long-term serological course of asymptomatic hepatitis B virus carriers and the development of primary hepatocellular carcinoma. *J Infect Dis* 1985;151:604–609.

246. Kaklamani E, Trichopoulos D, Tzonou A, et al. Hepatitis B and C viruses and their interaction in the origin of hepatocellular carcinoma. *JAMA* 1991;265:1974–1976.

247. Linne J, Wages J Jr, Zhang-Keck ZY, et al. Molecular cloning and disease association of hepatitis G virus: a transfusion-transmissible agent. *Science* 1996;71:505–508.

248. Lambert SB, Fields HA. The enterically transmitted hepatitis viruses. In: Lennette EH, Smith TF, eds. *Laboratory diagnosis of viral infections*, 3rd ed. New York: Marcel Dekker, 1997:467–487.

249. Cappel R, De Cuyper F, Van Beers D. e Antigen and antibody, DNA polymerase, and inhibitors of DNA polymerase in acute and chronic hepatitis. *J Infect Dis* 1977;136:617–622.

250. Fay O, Tanno H, Roncoroni M, et al. Prognostic implications of the e antigen of hepatitis B virus. *JAMA* 1977;238:2501–2503.

251. Okada K, Kamiyama I, Inomata M, et al. e Antigen and anti-e in the serum of asymptomatic carrier mothers as indicators of positive and negative transmission of hepatitis B virus to their infants. *N Engl J Med* 1976;294:746–749.

252. Choo QL, Kuo G, Weiner AJ, et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989;244:359–362.

253. Kuo G, Choo QL, Alter HJ, et al. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* 1989;244:362–364.

254. Esteban JI, Gonzalez A, Hernandez JM, et al. Evaluation of antibodies to hepatitis C virus in a study of transfusion-associated hepatitis. *N Engl J Med* 1990;323:1107–1112.

255. Martin P. Hepatitis C: from laboratory to bedside. *Mayo Clin Proc* 1990;65:1372–1376.

256. Aach RD, Stevens CE, Hollinger FB, et al. Hepatitis C virus infection in post-transfusion hepatitis. *N Engl J Med* 1991;325:1325–1329.

257. Centers for Disease Control and Prevention. Recommendations for prevention and control of hepatitis C virus (HCV). 1998;47:1–39.

258. Davis GL, Esteban-Mur R, Rustgi V, et al. Interferon alfa-2b alone or in combination with ribavirin for the treatment of relapse of chronic hepatitis C. International Hepatitis Interventional Therapy Group. *N Engl J Med* 1998;339:1493–1499.

259. Lew JF, Glass RI, Petric M, et al. Six-year retrospective surveillance of gastroenteritis viruses identified at ten electron microscopy centers in the United States and Canada. *Pediatr Infect Dis J* 1990;9:709–714.

260. LeBaron CW, Lew J, Glass RI, et al. Annual rotavirus epidemic patterns in North America. *JAMA* 1990;264:983–988.

261. Kotloff KL, Losonsky GA, Morris JG Jr, et al. Enteric adenovirus infection and childhood diarrhea: an epidemiologic study in three clinical settings. *Pediatrics* 1989;84:219.

262. Matson DO. Astroviruses. In: Long SS, Pickering LK, Prober CG, eds. *Principles and practice of pediatric infectious diseases*. New York: Churchill- Livingstone, 1997:1306–1308.

263. Caul EO, Ashley C, Pether JVS. "Norwalk"-like particles in epidemic gastroenteritis in the U.K. *Lancet* 1979;2:1292.

264. Dagan R. Nonpolio enteroviruses and the febrile young infant: epidemiologic, clinical and diagnostic aspects. *Pediatr Infect Dis J* 1996;15:67–71.

265. Rotbart HA. Enteroviral infections of the central nervous system. *Clin Infect Dis* 1995;20:971–87.

266. Dagan R, Hall CB, Powell KR, Menegus MA. Epidemiology and laboratory diagnosis of infection with viral and bacterial pathogens in infants hospitalized for suspected sepsis. *J Pediatr* 1989;115:351–356.

267. Rotbart HA. Nucleic acid detection systems for enteroviruses. *Clin Microbiol Rev* 1991;4:156–168.

268. American Academy of Pediatrics. Poliovirus infections. In: Peter G, Hall CB, Halsey NA, et al., eds. *Red book: report of the committee on infectious diseases*. 24th ed. Elk Grove Village, IL: American Academy of Pediatrics, 1997:425.

269. Chonmaitree T, Ford C, Sanders C, et al. Comparison of cell cultures for rapid isolation of enteroviruses. *J Clin Microbiol* 1988;26:2576–2580.

270. Dagan R, Menegus MA. A combination of four cell types for rapid detection of enteroviruses in clinical specimens. *J Med Virol* 1986;19:219–228.

271. Menegus MA, Hollick GE. Increased efficiency of group B Coxsackievirus isolation from clinical specimens by use of BGM cells. *J Clin Microbiol* 1982;15:945–948.

272. Klespies SL, Cebula DE, Kelley CL, et al. Detection of enteroviruses from clinical specimens by spin amplification shell vial culture and monoclonal antibody assay. *J Clin Microbiol* 1996;34:1465–1467.

273. Sawyer MH, Holland D, Aintablan N, et al. Diagnosis of enteroviral central nervous system infection by polymerase chain reaction during a large community outbreak. *Pediatr Infect Dis J* 1994;13:177–182.

274. Kessler HH, Santner B, Rabenau H, et al. Rapid diagnosis of enterovirus infection by a new one-step reverse transcription-PCR assay. *J Clin Microbiol* 1997;35:976–977.

275. Pozo F, Casas I, Tenorio A, et al. Evaluation of a commercially available reverse transcription-PCR assay for diagnosis of enteroviral infection in archival and prospectively collected cerebrospinal fluid specimens. *J Clin Microbiol* 1998;36:1741–1745.

276. Byington CL, Taggart B, Carroll K, et al. The use of PCR for the identification of enteroviral infection in febrile infants under 90 days of age. *Pediatrics* 1999;103:27.

277. Fuccillo DA, Sever JL. Measles, mumps, and rubella. In: Specter S, Lancz GJ, eds. *Clinical virology manual*. New York: Elsevier, 1986:437–449.

278. Germann D, Matter L. Rubella virus. In: Lennette EH, Smith FF, eds. *Laboratory diagnosis of viral infections*, 3rd ed. New York: Marcel Dekker, 1999:803–816.

279. Leland DS, Kleiman MB. Mumps virus. In: Lennette EH, Smith FF, eds. *Laboratory diagnosis of viral infections*, 3rd ed. New York: Marcel Dekker, 1999:623–634.

280. Minnich LL, Goodenough F, Ray CG. Use of immunofluorescence to identify measles virus infections. *J Clin Microbiol* 1991;29:1148–1150.

281. Field PR, Ho DWT, Cunningham AL. Evaluation of rubella immune status by three commercial enzyme-linked immunosorbent assays. *J Clin Microbiol* 1988;26:990–994.

282. Linde GA, Granstrom M, Orvell C. Immunoglobulin class and immunoglobulin G subclass enzyme-linked immunosorbent assays compared with microneutralization assay for serodiagnosis of mumps infection and determination of immunity. *J Clin Microbiol* 1987;25:1653–1658.

283. Neumann PW, Weber JM, Jessamine AG, et al. Comparison of measles antihemolysin test, enzyme linked immunosorbent assay, and hemagglutination inhibition test with neutralization test for determination of immune status. *J Clin Microbiol* 1985;22:296–298.

284. Benito RJ, Larrad L, Lasierra MP, et al. Persistence of specific IgM antibodies after natural mumps infection. *J Infect Dis* 1987;155:156–157.

285. Morgan-Capner P, Tedder RS, Mace JE. Rubella specific IgM reactivity in sera from cases of infectious mononucleosis. *J Hyg (London)* 1983;90:407–413.

286. Kurtz JB, Anderson MJ. Cross-reactions in rubella and parvovirus specific IgM tests. *Lancet* 1985;2:1356.

287. Morgan-Capner P. Diagnosing rubella. *BMJ* 1989;299:338–339.

288. Bellin E, Safyer S, Braslow C. False positive IgM-rubella enzyme-linked immunoassay in three first trimester pregnant patients. *Pediatr Infect Dis J* 1990;9:671–672.

289. Constantine NT, van der Groen, Belsey EM, et al. Sensitivity of HIV-antibody assays determined by seroconversion panels. *AIDS* 1994;8:1715–1720.

290. Reesink HW, Lelie PN, Huisman JG, et al. Evaluation of six enzyme immunoassays for antibody against human immunodeficiency virus. *Lancet* 1986;2:483–486.

291. Jackson JB, Balfour HH. Practical diagnostic testing for human immunodeficiency virus. *Clin Microbiol Rev* 1988;1:124.

292. Ozanne G, Fauvel M. Performance and reliability of five commercial enzyme-linked immunosorbent assay kits in screening for anti-human immunodeficiency virus antibody in high-risk subjects. *J Clin Microbiol* 1988;26:1496–1500.

293. Association of State and Territorial Public Health Laboratory Directors and AIDS Program, Centers for Disease Control. Interpretation and use of the Western blot assay for serodiagnosis of human immunodeficiency virus type 1 infections. *MMWR Morb Mortal Wkly Rep* 1989;38(Suppl S-7):1–7.

294. Jackson JB, MacDonald KL, Cadwell J, et al. Absence of HIV infection in blood donors with indeterminate Western blot tests for antibody to HIV-1. *N Engl J Med* 1990;322:217–222.

295. Kovacs A, Xu J, Rasheed S, et al. Comparison of a rapid nonisotopic polymerase chain reaction assay with four commonly used methods for the early diagnosis of human immunodeficiency virus type 1 infection in neonates and children. *Pediatr Infect Dis J* 1995;14:948–954.

296. Nesheim S, Lee F, Kalish MR, et al. Diagnosis of perinatal human immunodeficiency virus infection by polymerase chain reaction and p24 antigen detection after immune complex dissociation in an urban community hospital. *J Infect Dis* 1997;175:1333–1336.

297. Whetsell AJ, Drew JB, Milman G, et al. Comparison of three nonradioisotopic polymerase chain reaction-based methods for detection of human immunodeficiency virus type 1. *J Clin Microbiol* 1992;30:845–853.

298. Rogers MF, Ou Cy, Rayfield M, et al. Use of the polymerase chain reaction for early detection of the proviral sequences of human immunodeficiency virus in infants born to seropositive mothers. *N Engl J Med* 1989;320:1649–1654.

299. Centers for Disease Control and Prevention. 1994 revised classification system for human immunodeficiency virus infection in children less than 13 years of age. *MMWR Morb Mortal Wkly Rep* 1994;43(Suppl RR-12):1–19.

300. Centers for Disease Control. Arboviral infections of the central nervous system, United States, 1996–1997. *Morb Mortal Wkly Rep* 1998;47:517–522.

301. Whitley RJ. Viral encephalitis. *N Engl J Med* 1990;323:242–250.

302. Calisher CH, Berardi VP, Muth DJ, et al. Specificity of immunoglobulin M and G antibody responses in humans infected with eastern and western equine encephalitis viruses: application to rapid serodiagnosis. *J Clin Microbiol* 1986;23:369–372.

303. Calisher CH, Pretzman CI Muth DJ, et al. Serodiagnosis of Lacrosse virus infections in humans by detection of immunoglobulin M class antibodies. *J Clin Microbiol* 1986;23:667–671.

304. Goodpasture HC, Poland JD, Francy DB, et al. Colorado tick fever: clinical, epidemiologic, and laboratory aspects of 228 cases in Colorado in 1973–1974. *Ann Intern Med* 1978;88:303–310.

305. Wong JK. Colorado tick fever virus and other arthropod-borne reoviridae. In: Richman DD, Whitley RJ, Hayden FG, eds. *Clinical virology*. New York: Churchill-Livingstone, 1997:755–764.

306. Plummer FA, Hammond GW, Forward K, et al. An erythema infectiosum-like illness caused by human parvovirus infection. *N Engl J Med* 1985;313:74–79.

307. Serjeant GR, Mason K, Topley JM, et al. Outbreak of aplastic crises in sickle cell anaemia associated with parvovirus like agent. *Lancet* 1981;2:595–597.

308. Kurtzman GJ, Ozawa K, Hanson G, et al. Chronic bone marrow failure due to persistent B19 parvovirus infection. *N Engl J Med* 1987;317:287–294.

309. Anand A, Gray ES, Brown T, et al. Human parvovirus infection in pregnancy and hydrops fetalis. *N Engl J Med* 1987;316:183–186.

310. Anderson LJ, Tsou C, Parker RA, et al. Detection of antibodies and antigens of human parvovirus B19 by enzyme-linked immunosorbent assay. *J Clin Microbiol* 1986;24:522–526.

311. Erdman DD, Usher MJ, Tsou C, et al. Human parvovirus B19 specific IgG, IgA, and IgM antibodies and DNA in serum specimens from persons with erythema infectiosum. *J Med Virol* 1991;35:110–115.

312. Koch WC, Adler SP. Detection of human parvovirus B19 DNA by using the polymerase chain reaction. *J Clin Microbiol* 1990;28:65–69.

313. Mori J, Field AM, Clewley JP, et al. Dot blot hybridization assay of B19 virus DNA in clinical specimens. *J Clin Microbiol* 1989;27:459–464.

314. Butler JC, Peters CJ. Hantaviruses and hantavirus pulmonary syndrome. *Clin Infect Dis* 1994;19:387–395.

315. Centers for Disease Control and Prevention. Outbreak of acute illness-southwestern United States. *MMWR Morb Mortal Wkly Rep* 1993;42:412.

316. Doyle TJ, Bryan RT, Peters CJ. Viral hemorrhagic fevers and Hantavirus infections in the Americas. *Infect Dis Clin N Am* 1998;12:95–110.

317. Jenison S, Hjelle B, Simpson S, et al. Hantavirus pulmonary syndrome: clinical, diagnostic and virologic aspects. *Semin Respir Infect* 1995;10:259–269.

318. Zaki SR, Greer PW, Coffield LM, et al. Hantavirus pulmonary syndrome: pathogenesis of an emerging infectious disease. *Am J Pathol* 1995;146:552–579.

319. Korones SB, Ainger LE, Monif GRG, et al. Congenital rubella syndrome: new clinical aspects with recovery of virus from affected infants. *J Pediatr* 1965;67:166–181.

320. Thomson RB Jr, Benedict ML, Blaise LC. Evaluation of a commercially available cytomegalovirus-specific IgM test for the diagnosis of cytomegalovirus infection in the symptomatic newborn. *Diagn Microbiol Infect Dis* 1987;7:211–214.

321. Kohl S. The neonatal human's immune response to herpes simplex virus infection: a critical review. *Pediatr Infect Dis J* 1989;8:67–74.

60

Parasitology

Michael A. Pfaller

Medical parasitology is the study of invertebrate animals capable of causing disease in humans. In the context of this chapter, the term parasite refers to organisms belonging to three major subdivisions or taxonomic groups. These include the Protozoa (amebas, flagellates, ciliates, sporozoans, coccidia, microsporidians), the Platyhelminthes or flatworms (cestodes, trematodes), and the Nematodes or roundworms (Table 60.1). The Protozoa are rather simple microscopic unicellular eukaryotic organisms, whereas the Platyhelminthes and Nematoda are highly complex macroscopic multicellular worms possessing differentiated tissues and complex organ systems.

TABLE 60.1. CLASSIFICATION OF HUMAN PARASITES

Protozoa	**Nematodes**
Amebas	*Ascaris lumbricoides*
Entamoeba histolytica	*Enterobius vermicularis*
Entamoeba dispar	*Ancylostoma duodenale*
Entamoeba coli	*Necator americanus*
Entamoeba polecki	*Strongyloides stercoralis*
Entamoeba hartmanni	*Trichuris trichiura*
Endolimax nana	*Trichinella spiralis*
Blastocystis hominis	*Toxocara canis*
Iodamoeba bütschlii	*Ancylostoma braziliense*
Naegleria fowleri	*Wuchereria bancrofti*
Acanthamoeba culbertsoni	*Brugia malayi*
Flagellates	*Loa loa*
Giardia lamblia	*Onchocerca volvulus*
Chilomastix mesnili	*Mansonella ozzardi*
Dientamoeba fragilis	*Mansonella streptocerca*
Trichomonas hominis	*Mansonella perstans*
Trichomonas vaginalis	*Dirofilaria spp.*
Leishmania tropica	**Trematodes**
Leishmania mexicana	*Fasciolopsis buski*
Leishmania braziliensis	*Heterophyes heterophyes*
Leishmania donovani	*Metagonimus vokogawai*
Trypanosoma gambiense	*Opisthorchis sinensis*
Trypanosoma rhodesiense	*Opisthorchis viverrini*
Trypanosoma cruzi	*Fasciola hepatica*
Ciliates	*Paragonimus westermani*
Balantidium coli	*Schistosoma mansoni*
Sporozoa	*Schistosoma japonicum*
Plasmodium vivax	*Schistosoma haematobium*
Plasmodium malariae	**Cestodes**
Plasmodium ovale	*Diphyllobothrium latum*
Plasmodium falciparum	*Diphylidium caninum*
Babesia microti	*Hyrnenolepsis nana*
Coccidia	*Hymenolepsis diminuta*
Cryptosporidium spp.	*Taenia solium*
Cyclospora cayetanensis	*Taenia saginata*
Isospora belli	*Echinococcus granulosus*
Sarcocystis spp.	*Echinococcus multilocularis*
Toxoplasma gondii	*Multiceps multiceps*
Microsporidia	
Encephalitozoon	
Nosema	
Pleistophora	
Enterocytozoon	
Trachipleistophora	

TABLE 60.2. ESTIMATED WORLDWIDE PREVALENCE OF PARASITIC INFECTIONS

Infection	**Number Infected**	**Annual Deaths**
Amebiasis	10% of world population	40,000–110,000
Malaria	400–490 million	2.2–2.5 million
African trypanosomiasis	100,000 new cases/year	5,000
American trypanosomiasis	24 million	60,000
Leishmaniasis	1.2 million	
Schistosomiasis	>200 million	0.5–1 million
Opisthorchiasis	19 million	
Paragonimiasis	3.2 million	
Fasciolopsiasis	10 million	
Filariasis	85–100 million	
Onchocerciasis	>30 million	
Dracunculiasis	10 million	
Ascariasis	1 billion	1,550 (intestinal obstruction)
Hookworm	900 million	
Trichuriasis	500–800 million	
Strongyloidiasis	35 million	
Cestodiasis	65 million	

Although parasitic diseases are frequently considered tropical and thus of little relevance to physicians practicing in the more temperate, developed countries of the world, it is clear that the world has become a very small place and that it is essential for physicians to have knowledge of parasitic diseases. By any estimation, the numbers of parasitic infections and parasite-associated deaths are staggering and must be of concern to all health care workers (Table 60.2). Increasingly, tourists, missionaries, Peace Corps volunteers, and others are visiting and working for extended periods in exotic, remote parts of the world and thus are at risk for parasitic and other infections that are rare in the United States. Another source of infected patients is the ever-increasing population of refugees from developing countries. Finally, the profound immunosuppression accompanying advances in medical therapy as well as that associated with infection with the human immunodeficiency virus (HIV) places an increasing number of individuals at risk for developing infections attributable to particular parasites. Given these considerations, clinicians and laboratory workers should certainly be aware of the possibility of parasitic disease and should be trained in the ordering, performance, and interpretation of the appropriate laboratory tests to aid in diagnosis and therapy.

Although the mainstay of diagnostic clinical microbiology is the isolation of the causative pathogen in culture, diagnosis of parasitic diseases is most commonly accomplished by morphologic (usually microscopic) demonstration of parasites in clinical material. Occasionally, serodiagnosis is helpful in establishing the diagnosis. Increasingly, antigen detection and molecular diagnostic methods are proving useful in diagnosing parasitic infections (1, 2, 3, 4, 5 and 6). Proper diagnosis requires that (a) the physician consider the possibility of a parasitic infection, (b) appropriate specimens are obtained and transported to the laboratory in a timely fashion, (c) the laboratory competently performs the appropriate procedures for recovery and identification of the etiologic agent, (d) the laboratory results are effectively communicated to the physician, and (e) the results are correctly interpreted by the physician and applied to the care of the patient. For most parasitic diseases, appropriate test selection and interpretation are based on an understanding of the life cycle of the parasite as well as the pathogenesis of the disease process in humans. For example, although Ascaris is usually considered an intestinal nematode, early and profound clinical symptoms may be caused by larval migration through tissues weeks before eggs are present in feces.

This chapter provides an overview of the basic parasitology, epidemiology, clinical disease, and pathology of the major protozoan and helminthic parasites infecting humans. Individual organisms and disease processes are discussed under the major headings of blood and tissue protozoa, intestinal and urogenital protozoa, nematodes, trematodes, and cestodes. The reader is referred to several excellent reference texts for more detailed information (7, 8, 9 and 10).

BLOOD AND TISSUE PROTOZOA

Part of "60 - Parasitology"

The protozoa of blood and tissues include the sporozoan parasites *Plasmodium*, *Babesia microti*, and *Toxoplasma gondii*; the hemoflagellates include *Leishmania* and *Trypanosoma*; and the free-living amebae include *Naegleria*, *Acanthamoeba*, and *Balamuthia*. The major clinical manifestations of the protozoa causing bloodstream infection (malaria and babesiosis) are secondary to the destruction of red blood cells or sludging of infected red blood cells in the microvasculature of the brain and other organs. The protozoa causing tissue infection may cause significant damage to specific organs such as the eyes (toxoplasmosis, malaria, *Acanthamoeba* keratitis), the brain (toxoplasmosis, amebic meningoencephalitis, African sleeping sickness), the heart (toxoplasmosis, Chagas' disease), or the gastrointestinal tract (Chagas' disease) (Table 60.3).

TABLE 60.3. MAJOR BLOOD AND TISSUE PROTOZOA

Organism	Disease	Vector	Location	Laboratory Diagnosis
Plasmodium	Malaria	*Anopheles* mosquito	Blood	Giemsa-stained blood film
Babesia	Babesiosis	*Ixodes* tick	Blood	Giemsa-stained blood film
Toxoplasma gondii	Toxoplasmosis	None; food-borne, fecal-oral, transplacental	Reticuloendothelial system	Parasites in biopsy, serology
Leishmania	Leishmaniasis	*Phlebotomus* sandfly	Skin Mucosa Reticuloendothelial system	Parasites in Giemsastained smears, biopsy
Trypanosoma cruzi	Chagas' disease American trypanosomiasis	Reduviid bug	Blood Reticuloendothelial system	Trypanosomes in blood, biopsy
Trypanosoma brucei, gambiense or rhodesiense	Sleeping sickness African trypanosomiasis	Tsetse fly	Blood Lymphatics CNS	Trypanosomes in blood, CSF

CNS, central nervous system; CSF, cerebrospinal fluid.

Malaria

Malaria is the most important of all protozoan diseases; annually, it infects more than 250 million individuals and is a leading cause of illness and death in the developing world (Table 60.2). Efforts at eradication have failed because of the ability of both

the anopheline mosquito vector and the parasite to develop resistance to various eradication and treatment options. In the United States, malaria is well controlled, although it remains the most common acute febrile illness imported into the country.

Classification

Malaria is caused by blood-borne pathogens of the genus *Plasmodium*. There are four species that can infect humans and that are associated with somewhat different clinical presentations, laboratory characteristics, and geographic distribution (Table 60.4). These four species include *P. vivax*, *P. ovale*, *P. malariae*, and *P. falciparum*. *P. vivax* is the most common species worldwide, being found in India, Pakistan, Southeast Asia, South and Central America, and rarely in Africa. *P. ovale* is limited in distribution primarily to Africa. *P. malariae* is found in all countries with malaria, although it is much less prevalent than either *P. vivax* or *P. falciparum*. *P. falciparum* ranks second in prevalence to *P. vivax* and is primarily a tropical species. It is found most often in individuals from Africa, Haiti, and New Guinea but is also seen in Southeast Asia, South America, and the Pacific Islands. Clinically, *P. falciparum* infections are the most serious among the four species of plasmodia. Figure 60.1 depicts the distribution of malaria transmission, including that of drug-resistant strains.

TABLE 60.4. SELECTED CLINICAL CHARACTERISTICS OF FOUR TYPES OF MALARIA

Characteristic	*Plasmodium falciparum*	*Plasmodium vivax*	*Plasmodium ovale*	*Plasmodium malariae*
Usual incubation period (d)	8–11	10–17 or longer	10–17 or longer	18–40 or longer
Severity of primary attack	Severe in nonimmune	Mild to severe	Mild	Mild
Periodicity (h)	None	48	48	72
Duration of untreated primary attack (wk)	2–3	3–8	2–3	3–24
Duration of untreated infection	6–17 mo	5–7 yr	12 mo	20+ yr
Average parasitemia (per mm^2)	≥20,000	10,000	9,000	6,000
Anemia	Frequent and severe	Mild	Mild	Mild
CNS involvement	Yes, severe	Rare	Rare	Rare
Nephrotic syndrome	Rare	Rare	No	Frequent

CNS, central nervous system

FIGURE 60.1. World distribution of malaria. Black areas indicate distribution of chloroquine-susceptible *Plasmodium falciparum* malaria and gray areas indicate distribution of chloroquine-resistant *P. falciparum* malaria. (From Barat LM, Bloland PB. Drug resistance among malaria and other parasites. *Infect Dis Clin North Am* 1997;11:969–987.)

Clinical Manifestations

Malaria is characterized by acute cyclical periods (paroxysms) of high fever and shaking chills. After several replicative cycles of the parasite within the patient, a classic pattern of illness occurs. Patients develop a dull headache and generalized malaise followed by severe chills, peripheral vasoconstriction, and polyuria. Within a few hours, patients have high fever and flushing and, after 2 to 6 hours, a drenching sweat that resolves the episode but leaves the patient exhausted. Complete cycles occur in a tertian (48-hour) pattern in *P. vivax* and *P. ovale* and in a quartan (72-hour) pattern in *P. malariae* infections. A definite pattern frequently cannot be identified in *P. falciparum* infections; however, sometimes pronounced synchronization may be seen with a 36- to 48-hour periodicity. In addition to fever and chills, patients may also have flulike symptoms with myalgias, arthralgias, nausea, vomiting, abdominal pain, and headache. On physical examination, patients may have splenomegaly, hepatomegaly, or jaundice, although none of these is seen often. The most common laboratory abnormality associated with malaria is hemolytic anemia and hyperbilirubinemia. Elevated liver function tests, hyponatremia, and other nonspecific laboratory abnormalities may also occur.

The density of parasitemia may be extremely high in *P. falciparum* malaria, with as many as 50% of the red blood cells being parasitized. This degree of parasitemia can lead to severe hemolysis with hemoglobinuria and severe anemia. Sequestration of *P.*

falciparum infected erythrocytes in the microvasculature of the body may lead to occlusion of these vessels, causing symptoms related to capillary obstruction. With severe disease, virtually always caused by *P. falciparum*, evidence of cerebral infection with altered consciousness, encephalopathy, seizures, and focal neurologic deficits or intravascular hemolysis with renal failure (blackwater fever) may occur. This syndrome is considered a medical emergency and requires immediate and aggressive therapy.

There are several genetic factors that may alter the susceptibility to, and the clinical presentation of, malaria. The development of *P. falciparum* is suppressed by the presence of fetal hemoglobin, hemoglobin S, and perhaps other abnormal hemoglobins as well (11, 12). Persons with sickle cell trait have less severe infections with *P. falciparum* and do not have lethal sickle cell disease. The high frequency of hemoglobin S in certain parts of Africa is owing to the selective advantage of this balanced polymorphism. Likewise, persons with the Duffy-negative blood type (Fy^a and Fy^b negative) are resistant to infection with *P. vivax*. The Duffy blood group antigen is the specific receptor for invasion of red cells by *P. vivax* merozoites (11, 12). The low incidence of *P. vivax* in Africa is explained by the fact that most African blacks are Duffy negative.

The course of untreated malaria depends greatly on the infecting species. As noted previously, most fatal cases of malaria are caused by *P. falciparum*. In nonfatal cases of malaria, the febrile paroxysms become less severe over time and the disease gradually subsides. Clinical and parasitologic relapses may occur after months or years in individuals with *P. vivax* or *P. ovale* infection. These relapses are attributable to the emergence of latent exoerythrocytic forms present in the liver (hypnozoite; see Life Cycle section). *P. malariae* frequently causes an extremely low-grade parasitemia, and individuals with *P. malariae* infection may be asymptomatic with recrudescences occurring sporadically. It is thought that relapses or recrudescences may be associated with changes in the host's defense mechanisms or possibly with antigenic variation in the infecting organism.

Life Cycle

Malaria is transmitted to humans by the bite of the female Anopheles mosquito. Between 65 and 70 species of naturally infected *Anopheles* mosquitoes have been found in various parts of the world. Most malaria seen in the United States is acquired by visitors or residents from countries with endemic disease and is considered “imported” malaria. However, because the appropriate vector Anopheles mosquito is found in the United States, domestic transmission of the disease has been observed (introduced malaria) (11, 12). In addition to transmission by mosquitoes, malaria can be acquired by blood transfusions from an infected donor or by sharing of needles and syringes by intravenous drug users. Congenital transmission of malaria from an infected mother to the fetus is also described.

The four species of *Plasmodium* infecting humans share a common life cycle (Fig. 60.2) involving an asexual phase (schizogony) in humans and a sexual phase (sporogony) in the mosquito. The sexual phase requires a developmental period of

8 to 12 days in the mosquito and results in the production of infectious sporozoites.

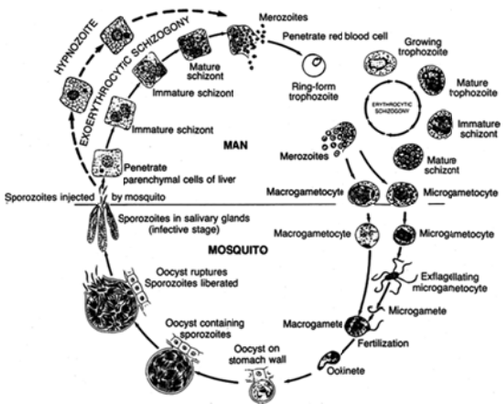


FIGURE 60.2. Life cycle of malaria. (From Strickland GT. *Tropical medicine*, 7th ed. Philadelphia: Saunders.)

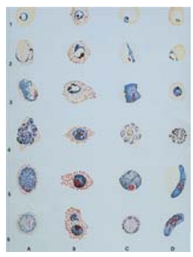


FIGURE 60.3. Developmental forms of four species causing human malaria. Columns: **A**, *Plasmodium vivax*; **B**, *P. ovale*; **C**, *P. malariae*; **D**, *P. falciparum*. Rows: 1, young trophozoites; 2, growing trophozoites; 3, mature trophozoites; 4, mature schizonts; 5, macrogametocytes; 6, microgametocytes. (From Wilcox A. *Manual for the microscopical diagnosis of malaria in man*. Bulletin no. 180. Bethesda, MD: National Institutes of Health, 1942.)

The infective sporozoites are injected into the subcutaneous capillaries of the vertebrate host as the mosquito feeds. The sporozoites penetrate the parenchymal cells of the liver, where they multiply asexually (exoerythrocytic schizogony) to produce thousands of uninucleate merozoites and form a tissue schizont. This growth of parasites in the liver is called the exoerythrocytic phase.

After a period of development (8 to 25 days depending on the species), the merozoites rupture from the hepatic cell and are released into the circulation, where they invade red blood cells. In infection caused by *P. falciparum* and *P. malariae*, the tissue schizonts all rupture simultaneously, and none persists in the liver. In contrast, *P. vivax* and *P. ovale* produce two types of exoerythrocytic forms: a primary type that develops and ruptures within 6 to 9 days and a secondary type called a hypnozoite that may remain dormant in the liver for weeks, months, or years. The hypnozoite form is responsible for relapses of erythrocytic infection observed with *P. vivax* and *P. ovale*.

Although relapses of *P. vivax* and *P. ovale* malaria may occur owing to the persistent exoerythrocytic (hypnozoite) phase, this is not the case with *P. falciparum* and *P. malariae*, which lack a dormant hepatic stage. Recurrent parasitemia with these species is caused by proliferation of persistent erythrocytic forms and is known as recrudescence. The prolonged delayed recrudescences occasionally seen with *P. malariae* are caused by erythrocytic parasites that have persisted in the tissue microcirculation.

The merozoites released from tissue schizonts invade circulating red blood cells and initiate the erythrocytic phase of infection. The erythrocytic phase is responsible for the clinical manifestations of malaria and on which laboratory diagnosis is based (Fig. 60.3). The different species of plasmodia selectively infect red blood cells of different ages. *P. vivax* and *P. ovale* preferentially infect young red blood cells, *P. malariae* infects only older circulating red cells, and *P. falciparum* can infect red cells of any age. The merozoites infecting red cells become small, rounded trophozoites or ring forms. Over time, the parasites develop into irregular ameboid trophozoites that utilize hemoglobin and produce an iron-containing pigment known as hematin or hemozoin. The mature trophozoite undergoes nuclear division and segmentation to form a mature schizont containing as many as 24 merozoites in a process known as erythrocytic schizogony. The erythrocytes containing the mature schizont rupture, liberating merozoites, which then go on to invade new red blood cells. This erythrocytic cycle is repeated with a periodicity that differs according to the species of *Plasmodium*.

Subpopulations of merozoites may also differentiate into sexual forms or gametocytes. The female form is called the macrogametocyte and the male form is called the microgametocyte. The gametocytes are ingested in the blood meal of a mosquito, sexual reproduction takes place, and the life cycle of the parasite is completed within the gastrointestinal tract of the mosquito with the production of new sporozoites.

Diagnosis

Malaria should be considered in any patient with an acute febrile illness and recent travel to or migration from an endemic area. The diagnosis is generally made by demonstration of the parasite within red blood cells. Organisms are easily seen in capillary blood obtained by fingerstick, although venous blood may also be used. Blood does not have to be obtained during the febrile period of infection; parasites may be present in blood even during the asymptomatic periods. The level of parasitemia may be low enough, however, that multiple attempts may be necessary to detect organisms within red cells.

Two types of blood films are prepared and examined for the laboratory diagnosis of malarial infections: thin films, with the blood spread over the slide in a thin layer, and thick films, with the blood concentrated in a small area. In the thick-film preparation, the red blood cells are lysed, and only white blood cells, platelets, and parasites (if present) are visible. The thick film is preferred for diagnosis because it permits the examination of a relatively large amount of blood and may be a more sensitive means of detecting low-grade parasitemias that may be missed with thin films. However, morphologic features of parasites are more distinct and typical in thin films, and thus both types of blood films should be prepared.

The preferred stain for malarial parasites in both thick and thin films is Giemsa buffered to pH 7.0. Although Wright's stain is commonly used to stain peripheral smears for hematologic examination, neither Wright's nor other Romanowsky stain produces the critical differentiation of morphologic features obtained with a Giemsa stain.

Identification of malarial parasites to the species level is critical. Patients with falciparum malaria are likely to have more severe disease, including severe hemolysis and cerebral involvement that can result in death in a matter of hours if not appropriately treated. Patients with infections caused by *P. vivax* or *P. ovale* require treatment for both erythrocytic and exoerythrocytic forms to avoid relapsing infection.

Identification to species level is accomplished by a combination of epidemiologic, clinical, and morphologic criteria (Table 60.4 and Table 60.5). As noted previously, some malarial species are seen more frequently in various geographic locations. The periodicity of clinical signs and symptoms is helpful, and patients with high levels of parasitemia or profound symptoms are more likely to have *P. falciparum* infection. When examining blood smears, there are three major factors to be considered (Table 60.5 and Fig. 60.3): appearance of erythrocytes, appearance of parasites, and developmental stages found. For falciparum malaria (Fig. 60.3, Fig. 60.4 and Fig. 60.5), infected red blood cells are normal in size, multiple parasites may be present per cell, the trophozoite stage is a small ring form with double chromatin dots sometimes peripherally located against the wall of the cell in an “appliqué” pattern, schizonts or other mature forms are seldom seen, and the gametocyte is banana shaped. In infections caused by *P. vivax* and *P. ovale* (Fig. 60.3 and Fig. 60.5, Fig. 60.6 and Fig. 60.7), infected red blood cells are usually enlarged and contain numerous pink granules or Schüffner's dots, the trophozoite is ring shaped but ameboid in appearance, more mature trophozoites and erythrocytic schizonts containing as many as 24 merozoites are present, and the gametocytes are round. The erythrocytes infected with *P. ovale* are often oval or fimbriated (more than 30% of infected red blood cells). Infections caused by *P. malariae* (Fig. 60.3, Fig. 60.5, Fig. 60.7, and Fig. 60.8) are characterized by parasitized red cells that are normal to small in size, trophozoites with dense cytoplasm and occasional band forms, schizonts with

six to 12 merozoites occasionally arranged in a rosette pattern around a pigment clump, and round gametocytes.

TABLE 60.5. COMPARATIVE MORPHOLOGIC FEATURES OF MALARIAL PARASITES IN STAINED THIN BLOOD FILMS

Characteristics	*Plasmodium falciparum*	*Plasmodium vivax*	*Plasmodium ovale*	*Plasmodium malariae*
Infected erythrocyte				
Enlarged[a]	–	+	±	–
Oval, fimbriated	±	±	+	-
Schüffner's dots[a]	–	+	+	-
Maurer's dots	+	-	–	–
Parasite				
Multiple in single erythrocyte	+	±	–	–
Multiple forms	–	+	+	+
Only ring forms[a]	+	-	–	–
Large ameboid rings	–	+	+	+
Double chromatin dots	+	±	–	–
Peripheral location in erythrocyte	+	±	–	–
Band forms	–	–	–	+
Banana-shaped gametocytes[a]	+	-	–	–
Number of merozoites	8–24	12–24	8–12	6–12

[a] Reliable criteria.

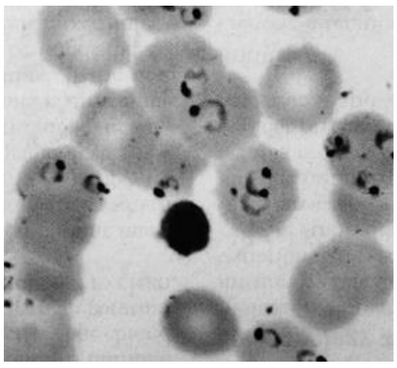


FIGURE 60.4. *Plasmodium falciparum*: Ring forms in film of peripheral blood. (From Garcia LS, Salzer AJ, Healy GR, et al. Blood and tissue protozoa. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 6th ed. Washington, DC: American Society for Microbiology, 1995:1171–1195.)

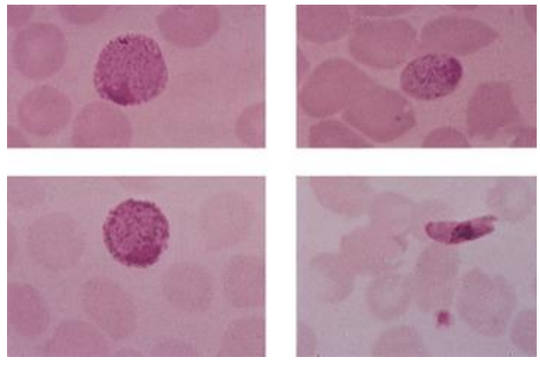


FIGURE 60.5. Macrogametocytes of *Plasmodium vivax* **(top left)**, *P. malaria* **(top right)**, *P. ovale* **(bottom left)**, and *P. falciparum* **(bottom right)**. (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

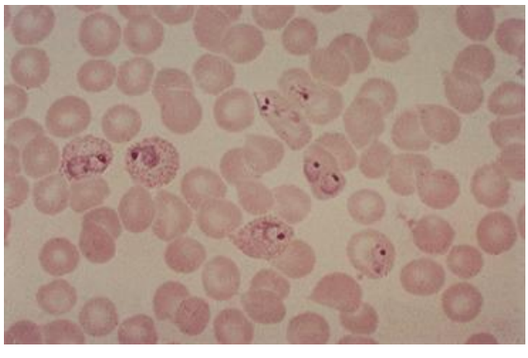


FIGURE 60.6. *Plasmodium vivax*: young growing trophozoites and ring forms (original magnification, ×1,000). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

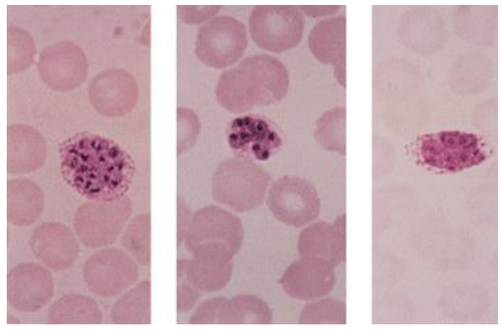


FIGURE 60.7. Mature schizonts of *Plasmodium vivax* **(left)**, *P. malariae* **(center)**, and *P. ovale* **(right)** (original magnification, ×1,000). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

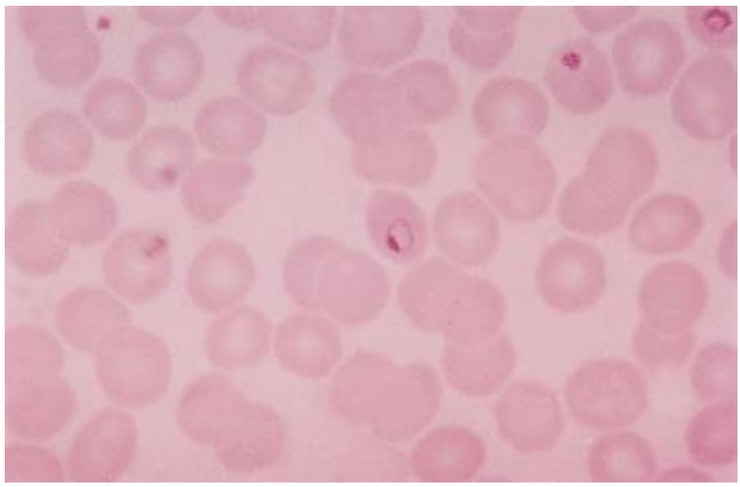


FIGURE 60.8. *Plasmodium malariae*: ring forms in thin film of peripheral blood. (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Although mixed infections with more than one species of *Plasmodium* can occur, documentation is rare and caution should be used in making such a diagnosis unless there is definitive evidence of two species. The most common mixed infection is attributable to a combination of *P. falciparum* and *P. vivax*. The presence of characteristic *P. falciparum* gametocytes in a patient otherwise infected with *P. vivax* (enlarged red blood cells, multiple developmental forms present) is considered diagnostic. If only ring forms are present, the diagnosis of dual infection must be considered tentative at best.

Several serodiagnostic tests have been developed for malaria; however, they are primarily used for seroepidemiologic studies and are rarely necessary to diagnose clinical infections (3). Newer methods include microhematocrit centrifugation and staining with acridine orange, antigen detection by enzyme-linked immunosorbent assay (ELISA) or radioimmunoassay, nucleic acid

detection by polymerase chain reaction (PCR), and detection of sporozoite- and merozoite-specific antibodies by ELISA or indirect immunofluorescence (5, 11, 13, 14, 15 and 16). None of these methods has yet supplemented routine microscopic diagnosis.

Treatment

Chloroquine is the drug of choice for both prophylaxis and treatment of infection owing to *P. ovale*, *P. malariae*, and susceptible strains of *P. falciparum* and *P. vivax* (11, 12, 17, 18, 19, 20, 21 and 22). Chloroquine-resistant strains of *P. falciparum* exist in virtually all malarious areas of the world except Central America northwest of the Panama Canal and limited parts of the Middle East, and their existence clearly complicates the prophylaxis and treatment of malaria. Updates on the prevalence of chloroquine-resistant *P. falciparum* and recommendations for antimalarial prophylaxis are published annually by the U.S. Centers for Disease Control and Prevention and by the World Health Organization. Resistance of *P. vivax* to chloroquine has been reported from Papua New Guinea, the Solomon Islands, Indonesia, and Brazil (17, 18, 19, 20, 21 and 22).

Patients infected with chloroquine-resistant *P. falciparum* (or *P. vivax*) may be treated with other agents, including mefloquine, quinine, quinidine, Fansidar (pyrimethamine-sulfadoxine), or doxycycline. Because quinine and Fansidar are potentially toxic, they are used more often for treatment rather than prophylaxis. Amodiaquine, an analogue of chloroquine, is effective against chloroquine-resistant *P. falciparum*; however, toxicity (agranulocytosis and hepatitis) limits its use. Newer agents with promising activity against multidrug-resistant strains of *P. falciparum* include halofantrine, a phenathrenenemethanol, and artemesinin, a sesquiterpene derivative of quinghaosu. They are not available in the United States.

Chloroquine and other blood schizonticides (quinine, quinidine, mefloquine, Fansidar, doxycycline) are not effective against the exoerythrocytic phase (hypnozoite) of *P. vivax* or *P. ovale*. To obtain a radical cure with these infections, primaquine is used to eradicate the hypnozoite and prevent its development to the tissue schizont stage and the subsequent release of infectious merozoites. Primaquine is not indicated for either *P. falciparum* or *P. malariae* infections because these parasites do not produce a hypnozoite stage in the liver. Use of primaquine may be dangerous in individuals with glucose-6-phosphate dehydrogenase deficiency and in whom the drug may cause hemolysis of older erythrocytes.

Prevention of malaria by vector control (insecticides, drainage of breeding sites) and personal protection measures against mosquito bites (mosquito netting, screening, insect repellents) are adjunctive measures to specific chemoprophylaxis and treatment regimens. The development of vaccines to protect individuals living in or traveling to endemic areas has been disappointing thus far.

Babesiosis

Babesiosis is a zoonosis infecting a variety of animals, including deer, cattle, and rodents. Humans are accidental hosts. Babesiosis is transmitted in nature by ixodid, or hard-bodied, ticks.

Classification

Babesiosis is caused by intraerythrocytic sporozoan parasites of the genus *Babesia*. Morphologically, *Babesia* resemble malaria parasites but can be distinguished from them by the absence of pigment within infected erythrocytes. More than 70 different species of *Babesia* have been described, including *B. bigemina*, *B. argentina*, and *B. divergens* in cattle; *B. caballi* and *B. equi* in horses; *B. canis* in dogs; and *B. microti* and *B. rodhaini* in rodents. Identification of the different species is based on morphology and the vertebrate host in which the parasite is found.

Babesiosis is a zoonotic disease occurring in Europe and the United States. The tick vector responsible for transmitting babesiosis, *Ixodes dammini*, is found throughout much of the United States. The natural reservoir hosts are field mice, voles, and other small rodents. Babesiosis is endemic in the United States, particularly on the East Coast. *B. microti*, the usual cause of babesiosis in the United States, generally causes a mild disease and is rarely fatal. *B. divergens*, which has been reported more frequently from Europe, causes severe, often fatal infections in individuals who have undergone splenectomies. Although most infections follow tick bites, transfusion-related infections have been reported.

Clinical Manifestations

Human babesiosis ranges in severity from asymptomatic to prolonged severe illness. For the most part, infection is self-limited and is characterized by general malaise, fever without periodicity, headache, chills, sweating, fatigue, and weakness (7, 23). The incubation period varies from 1 to 4 weeks. Mild to moderately severe hemolytic anemia and hepatosplenomegaly may develop secondary to red blood cell destruction by the parasite. In severe cases, renal failure and a picture of acute disseminated intravascular coagulation have been observed. The illness may last from a few weeks to several months. Low-grade parasitemia, with or without symptoms, may persist for as long as 4 months. Splenectomy or functional asplenia, as well as immunosuppression and advanced age, appear to increase individual susceptibility to infections, as well as more severe disease.

Life Cycle

Human babesiosis follows contact with an infected ixodid tick. The infectious forms are introduced into the bloodstream by the tick bite and infect erythrocytes. There is no evidence for an exoerythrocytic cycle such as that occurring in malaria. *Babesia* multiplies by budding within the red blood cell to form two to four daughter parasites. The red blood cells lyse, releasing merozoites, which can then infect other red blood cells and maintain the infection. The major clinical features of babesiosis, including hemolytic anemia, jaundice, and renal failure, are the result of multiplication within the red blood cells and their subsequent destruction.

The infected red blood cells can also be ingested by feeding ticks, in which additional replication can take place. Transovarial transmission from adult to larval tick also serves to maintain the infection within the tick population.

Diagnosis

As with malaria, the diagnosis of babesiosis requires the identification of characteristic intraerythrocytic parasites on Giemsa-stained thin or thick blood films (Fig. 60.9). Laboratory personnel must be experienced in differentiating *Babesia* and *Plasmodium* species. The morphologic characteristics of *Babesia* parasites are variable; however, *B. microti* usually appears as a small ring form indistinguishable from young trophozoites of *P. falciparum*. In contrast to *Plasmodium* species, no pigment is produced in erythrocytes infected with *Babesia*. Although characteristic, the tetrad form is rarely seen in human blood films. *Babesia* infections can also be diagnosed by animal (hamster) inoculation. Serologic tests are also available.

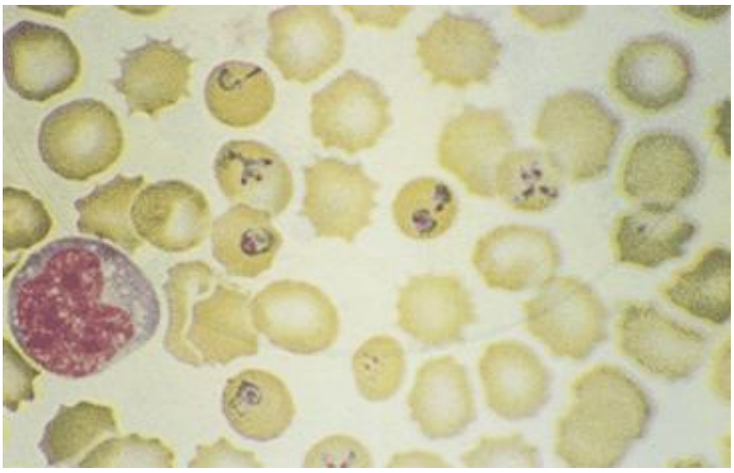


FIGURE 60.9. *Babesia* species ring and tetrad forms (Giemsa stain; original magnification, ×1,000). (From Garcia LS, Sulzer AJ, Healy GR, et al. Blood and tissue protozoa. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 6th ed. Washington, DC: American Society for Microbiology, 1995:1171–1195.)

Treatment

Human infections caused by *B. microti* are generally self-limited, and thus the effectiveness of treatment regimens has been difficult to evaluate. Symptomatic therapy is probably adequate for most patients. Various antiprotozoal regimens including chloroquine and pentamidine have been used with variable results. Severely ill (usually asplenic) patients have been treated successfully with quinine and clindamycin. Exchange blood transfusion has also been successful in splenectomized patients with severe infections caused by *B. microti* or *B. divergens* (7, 11, 23).

Preventive measures in endemic areas include wearing protective clothing and using insect repellents to minimize tick exposure. Because ticks must feed on humans for several hours before transmission occurs, prompt removal of ticks may also be protective.

Hemoflagellates

The blood and tissues of humans may be infected by one of several species of flagellate protozoa belonging to the family Trypanosomatidae. All belong to the *Leishmania* or *Trypanosoma* genus. All these organisms have developmental stages in blood-sucking arthropods (intermediate host) and in humans (definitive host), and many have a nonhuman mammalian reservoir host. *Leishmania* species are transmitted by sandflies (*Phlebotomus*, *Lutzomyia*), and the trypanosomes are transmitted by either the tsetse fly (*Glossina*) (African trypanosomes) or triatomid bugs (American trypanosomes).

The hemoflagellates may occur in a variety of stages in the human host and the insect vectors (Fig. 60.10). The amastigote form is an intracellular form that is small (2 to 5 μm) and round or oval in shape with a postcentral nucleus, anterior to which are the rodlike kinetoplast and an intracytoplasmic remnant of the

flagellum called the axoneme. The promastigote is an extracellular form present in culture and in the arthropod host. It is elongate and slender with a free flagellum. The kinetoplast is located at the anterior (flagellar) end of the parasite. The trypomastigote is the mature extracellular form that may be found in both the arthropod and the blood of the human host. It is long (17 to 30 mm) and slender with a centrally located nucleus and a subterminal kinetoplast. The axoneme forms the outer edge of the undulating membrane and extends anteriorly. The free flagellum projects from the anterior end of the parasite. The flagellum and the undulating membrane serve as a means of locomotion.

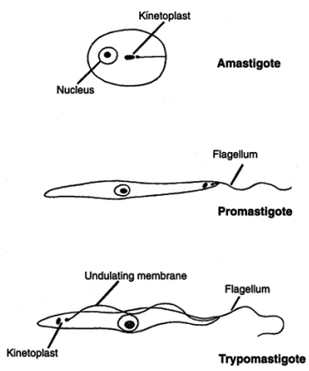


FIGURE 60.10. Morphologic types seen in various hemoflagellates of humans. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

The number of stages present in the life cycle of the parasite varies with the genus and species. The amastigote form is present in the reticuloendothelial cells of individuals infected with *Leishmania* species and in cardiac muscle and other tissues of patients with American trypanosomiasis (Chagas' disease). The trypomastigote form occurs in the bloodstream, lymphatics, and cerebrospinal fluid in both Gambian and Rhodesian forms of African trypanosomiasis and in the bloodstream in Chagas' disease. Nonflagellate, cryptic forms have also been observed in the choroid plexus of individuals with African trypanosomiasis. The promastigote form is the developmental stage present in the midgut and pharynx of the arthropod vector of leishmaniasis. An epimastigote form (not shown), having a free flagellum and partial undulating membrane, is found in the salivary glands of the tsetse fly vector of African trypanosomiasis. Both promastigote and epimastigote forms may be present in the triatomid vector of Chagas' disease.

Leishmania

Classification

Leishmaniasis is caused by a group of protozoan parasites transmitted by species of sandflies (*Phlebotomus*, *Lutzomyia*). There are three general forms of leishmaniasis: cutaneous, mucocutaneous, and visceral. Individual members of this group have unique geographic distributions and propensities to cause disease. The *L. tropica* group causes cutaneous disease in the Old World (around the Mediterranean Sea, North and East Africa, the Near East, Southern Russia, the Middle East, Afghanistan, India, and China); the *L. mexicana* and *L. braziliensis* groups cause cutaneous and mucosal disease in the New World (from Mexico south through Central and South America, especially Colombia, Ecuador, Venezuela, Peru, and Brazil); and the *L. donovani* group causes visceral disease in both the Old and New World (24).

Leishmania live as amastigotes within the reticuloendothelial cells of the skin, mucous membranes, and viscera. Multiplication is by binary fission. The different species of *Leishmania* cannot be distinguished morphologically. Separation into species and subspecies has been based on unique clinical syndromes, geographic distribution, specific animal reservoirs, and species of sandfly associated with transmission. Recently, the application of newer taxonomic and genetic techniques such as restriction analysis of kinetoplast DNA, isoenzyme patterns, and serologic testing has led to some changes in classification and will likely lead to additional changes in the future.

Clinical Manifestations: Visceral Leishmaniasis

Visceral leishmaniasis or kala azar is caused by *L. donovani* and is endemic in some parts of South America, Africa, China, India, and the Mediterranean area. Infection ranges from asymptomatic in the majority of individuals to an overwhelming systemic illness with death in some. The disease may begin with a skin lesion but primarily is a disease involving the reticuloendothelial system of the visceral organs. The incubation period varies, but clinical manifestations generally occur 3 to 8 months after exposure. The disease tends to be most pronounced in small children.

Although clinical symptoms may arise quickly with an acute illness, most patients have the insidious development of fever, lassitude, abdominal discomfort, and distention secondary to massive hepatosplenomegaly, and progressive emaciation. Fever may be high and intermittent, similar to malaria, or may be continuous and low grade with night sweats. Symptoms are generally present for weeks to months before patients seek medical attention. These long delays are responsible for the high complication and fatality rates (as high as 75%) that have been reported in many studies.

On physical examination, the major findings include pallor, jaundice, and hepatosplenomegaly. Both liver and spleen can become massively enlarged but remain soft. The skin becomes thin and dry and may acquire an earthy-gray color. Laboratory abnormalities include anemia, leukopenia, and thrombocytopenia. A polyclonal hypergammaglobulinemia is common, and circulating immune complexes may lead to glomerulonephritis.

The pathogenesis of the disease is a mixture of the direct effects of the organism on infected cells and the host response to the infectious process. *L. donovani* disseminates throughout the reticuloendothelial system after ingestion by macrophages. Multiplication of the parasites within the liver, spleen, and bone marrow leads to the clinical syndrome of kala azar.

Although most infected patients mount an immune response based on production of antibody, T cells appear to be necessary for activation of macrophages to phagocytize and kill the organism. Patients with the most severe disease, usually the young or malnourished, have an inadequate T-cell response to infection.

Visceral leishmaniasis is usually a sporadic disease occurring in endemic areas. Outbreaks have also been described. For example, in the State of Bihar, India, in 1977, 70,000 to 100,000 cases with 4,000 deaths were reported (25).

Clinical Manifestations: Cutaneous and Mucocutaneous Leishmaniasis

Cutaneous leishmaniasis is caused by several species of *Leishmania* and occurs in both the Old and New World. Cutaneous leishmaniasis of the Old World is an infection characterized by nodular and ulcerative skin lesions caused by *L. tropica*, *L. major*, and *L. aethiopica*. It is also known as Oriental sore, Baghdad boil, and Delhi boil, among other names.

Cutaneous leishmaniasis of the Old World usually begins as a small papule that develops between 2 weeks and 6 months after the bite of an infected sandfly. This papule slowly enlarges and ulcerates. The ulcer remains shallow and has a well-defined, slightly raised margin. The lesion may be dry or moist and may

be single or develop multiple satellite lesions. Regional adenopathy may be present. The ulcer is usually painless and heals spontaneously after several months unless it is secondarily infected with bacteria. Alternative presentations include papules that do not ulcerate but that disseminate to distant sites and remain active for years, or leishmaniasis recidivans, a relapsing disease with lesions that heal in the center but continue to advance at the periphery and may spread to involve mucous membranes, leading to local destruction of nasal tissue.

Cutaneous leishmaniasis of the New World is also characterized by ulcerative skin lesions. The organism complex causing infection in Central America and southern Mexico is referred to as the *L. mexicana* complex. *L. mexicana* causes a lesion known as chiclero ulcer. The characteristic lesions of the ear occur in 40% of patients and are chronic, lasting many years. *L. braziliensis* causes cutaneous and mucocutaneous infections in humans occurring mainly in Brazil but also seen in other parts of South America. Simple cutaneous lesions caused by *L. braziliensis* generally heal spontaneously within 6 to 18 months but may persist much longer. Persons infected with *L. braziliensis* may develop metastatic spread of the disease to the nasal, pharyngeal, and buccal mucosa. This mucocutaneous form of infection is known as espundia. The mucosal lesions may develop months to years after the onset of cutaneous lesions. Lesions of mucous membranes are slow healing and progressive and may result in massive destruction of the nasal septum, palate, lips, pharynx, and larynx. If untreated, this form of the disease may result in death caused by aspiration, secondary bacterial infection, or progressive malnutrition.

Life Cycle and Transmission

The life cycle of *Leishmania* is relatively simple and is generally shared by all species (Fig. 60.11). Sandflies feeding on infected individuals ingest free amastigotes or parasitized macrophages. On reaching the midgut of the insect, the amastigotes transform into flagellated promastigotes that multiply in the gut and migrate to the proboscis, where they are introduced into the skin of a vertebrate host during a blood meal. Promastigotes injected into the host invade the reticuloendothelial cells, transform into amastigotes, multiply within the phagolysosome, and destroy the cell. The liberated amastigotes then invade other cells, continuing the cycle.

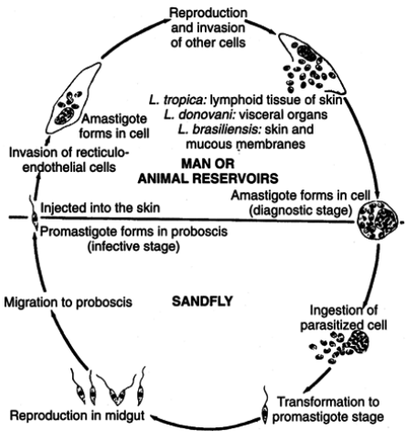


FIGURE 60.11. The life cycle of the *Leishmania*. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

Leishmania live in a variety of canine and rodent reservoirs throughout the world. These reservoir hosts may have mild or asymptomatic disease, cutaneous lesions only, or severe visceral disease. Humans may also serve as the reservoir host, particularly in epidemic situations. Rare cases of transmission by blood transfusion or parenteral contact in laboratory settings has been described.

Because there is no animal reservoir of infection in North America, cases of leishmaniasis that occur in the United States occur in individuals who have traveled to endemic areas and who have exposure to the insect vector. Visceral leishmaniasis is most commonly seen in infants and small children in South America and the Mediterranean region and in older children and adults elsewhere. The immunologic mechanism for the predisposition of certain age groups to infection is poorly understood.

Diagnosis

The diagnosis of leishmaniasis is usually made by demonstration of amastigotes in reticuloendothelial cells. Cutaneous and mucocutaneous infection may be diagnosed by examination of material scraped or aspirated from the margin of the ulcer. The material is smeared on a slide and stained with Wright's or Giemsa stain. Biopsies from the edge of the lesion may also be obtained and stained by standard histologic methods. Visceral leishmaniasis may be diagnosed by microscopic examination of Giemsa-stained aspirates of bone marrow or spleen. The amastigote stage of the parasite may be found within macrophages or spread out from ruptured cells (Fig. 60.12). The kinetoplast of the organism stains well with Giemsa stain and provides a means of differentiating *Leishmania* from fungal pathogens such as *Histoplasma*. Although culture of promastigotes on artificial media may also establish the diagnosis, it is not widely available. A skin test, the Montenegro test, is frequently positive in cutaneous or mucocutaneous disease; however, it is usually negative in visceral infection. Several serologic tests are available and may be useful in diagnosis or epidemiologic studies. DNA probes and other nucleic acid–based test have been developed for direct examination of clinical material (5, 6, 24, 26). There are no commercially available products for these tests, and careful studies to determine the accuracy of testing procedures have not yet been performed.

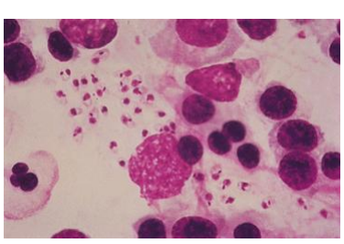


FIGURE 60.12. *Leishmania donovani*: Amastigote in human bone marrow smear (Giemsa stain; original magnification, ×1,280). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Treatment

Visceral leishmaniasis is treated with pentavalent antimonial compounds. Pentostam (stibogluconate sodium) is the agent most commonly employed and readily available. Therapy is not uniformly successful, and relapse rates of 2% to 8% are seen. Repeat

therapy with the same compound for extended periods may ultimately prove successful in difficult cases. Alternative approaches include the addition of allopurinol or treatment with pentamidine, amphotericin B, or one of several alternative oral agents (27,28). There is some toxicity associated with treatment with any of these compounds, and data on their effectiveness are limited.

Because simple cutaneous infection is usually self-limited, treatment is generally not administered. Treatment is considered for large lesions, for patients with potential or ongoing mucosal disfigurement, and occasionally for cosmetic reasons.

Trypanosoma

The trypanosomes infecting humans cause two distinctly different forms of disease: African trypanosomiasis (sleeping sickness) caused by *T. brucei gambiense* and *T. brucei rhodesiense* and American trypanosomiasis (Chagas' disease) caused by *T. cruzi*. All these species utilize blood-sucking arthropods as intermediate hosts and vectors for the transmission of infection.

African Trypanosomiasis

Classification. The African trypanosomes, *T. brucei gambiense* and *T. brucei rhodesiense*, are transmitted by species of *Glossina* (tsetse fly). Transmission is limited to the tsetse fly belt in central Africa. The two subspecies are indistinguishable morphologically but differ in the severity and rapidity of disease progression. In both *T. brucei gambiense* and *T. brucei rhodesiense*, the individual trypanosomes vary considerably in size and shape, ranging from delicate, spindle-shaped organisms with a free flagellum to broad, stumpy forms in which a free flagellum is not evident. Parasites may range from 14 to 22 μm in length and from 1.5 to 3.5 μm in breadth. The nucleus is usually centrally located or slightly posterior, and the posterior end of the organism is blunted rather than pointed.

Clinical Manifestations. Infection with the two subspecies results in different but overlapping clinical syndromes separated primarily by the rapidity with which the infection progresses. *T. brucei gambiense* is endemic in the Central African area, extending from the west coast through the central and equatorial zone as far east as Lake Victoria and Tanzania. It produces chronic disease, often ending fatally, with central nervous system involvement after several years duration. *T. brucei rhodesiense* is generally limited to the savanna areas of East Africa. Rhodesian sleeping sickness is characterized by a more fulminant course with the rapid onset of severe symptoms, frequently resulting in the death of the patient within a few months. Both varieties of African trypanosomiasis are associated with a mortality rate approaching 100% in untreated patients.

Infection with *T. brucei gambiense* and *T. brucei rhodesiense* occurs after a tsetse fly bite. At the site of the bite, a painless nodule develops known as the trypanosomal chancre. This lesion becomes firm and rubbery and resolves after approximately 2 weeks. Headache, fever, arthralgias, insomnia, edema, and lymphadenopathy accompany the early skin lesion as the parasite enters the bloodstream. Swelling of the posterior cervical lymph nodes is characteristic of Gambian disease and is called Winterbottom's sign. Eventually the organism enters the central nervous system, resulting in a variety of signs and symptoms, including irritability, tremors, ataxia, convulsions, meningoencephalitis, personality change, daytime somnolence, pronounced wasting, and eventually coma. Death is the result of central nervous system damage, combined with other infections such as malaria or pneumonia.

Infection with the more virulent *T. brucei rhodesiense* results in greater numbers of organisms in the blood and is generally without lymphadenopathy. Central nervous system invasion occurs early in the infection. The chronic stages of infection are not often seen with Rhodesian disease because, in addition to rapid central nervous system involvement, the organism produces kidney damage and myocarditis, leading to death. Persons with the Rhodesian form of sleeping sickness are usually dead within 9 to 12 months if untreated (24).

Life Cycle. The two African trypanosomes have similar life cycles (Fig. 60.13). Humans and other animals serve as hosts for the trypomastigote form, which lives and multiplies in the blood, lymph nodes, spleen, and cerebrospinal fluid. Tsetse flies become infected during meals on either humans or other infected animals, including several wild and domestic animals. Within the fly, the trypomastigotes multiply within the lumen of the gut, migrate to the salivary glands where they metamorphose into epimastigote forms, and eventually transform into metacyclic trypomastigotes that enter mammalian hosts during feeding. On entering the human host, the infective trypomastigotes remain in the tissues at the site of entry, often causing an interstitial inflammatory reaction. After several days, they gain entry into the bloodstream, where they multiply and then invade the lymphatics and central nervous system. Although they do not typically invade cells, they elicit an inflammatory response and hyperplasia of the vascular endothelium and produce injurious effects in every organ and tissue in the body.

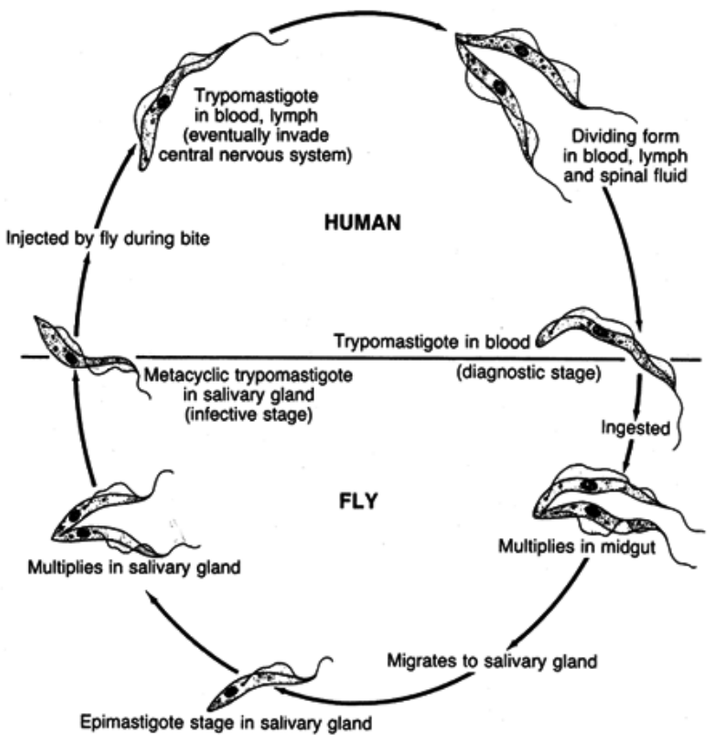


FIGURE 60.13. Life cycle of *Trypanosoma brucei gambiense* and *T. brucei rhodesiense*. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

Diagnosis. Definitive diagnosis of African trypanosomiasis is made by demonstration of the parasite in Giemsa-stained blood films during febrile periods and in lymph node aspirates at other

times. Examination of cerebrospinal fluid may be useful in individuals infected with *T. brucei gambiense*; however, patients with *T. brucei rhodesiense* infection rarely survive long enough for the parasites to invade the central nervous system in large numbers. Parasites may be seen on wet preparations of blood and other fluids but are most easily detected in thick films of blood stained with Giemsa (24,29). Ideally, both thick and thin films should be examined because the characteristic morphology of the trypomastigote is best seen in thin smears (Fig. 60.14). Methods for concentrating parasites in blood may also be helpful. Approaches include centrifugation of heparinized samples or DEAE (*O*-diethylaminoethyl)-cellulose anion-exchange chromatography. Levels of parasitemia vary widely, and several attempts to detect the organism over several days may be necessary. Because trypomastigotes disintegrate rapidly on removal from tissues, preparations should be fixed and stained almost immediately. *In vitro* culture methods or inoculation of blood into suckling mice with subsequent detection of the parasite in tissue are available but not generally used in the diagnosis of African trypanosomiasis.

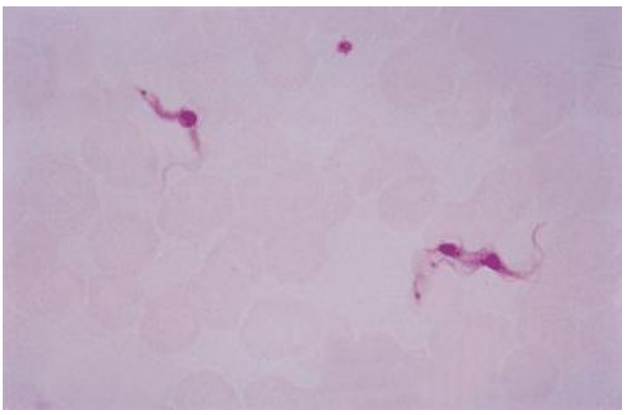


FIGURE 60.14. *Trypansoma rhodesiense*: trypomastigotes in thin blood film (Giemsa stain; original magnification, ×1,280). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Serologic tests are also useful diagnostic techniques. Immunofluorescence, ELISA, precipitin, and direct and indirect agglutination methods have been used. Most reagents for such tests are not available commercially. One difficulty in performing and applying serologic tests is the marked variability of the surface antigens of trypanosomes. Such antigens change over time for parasites infecting individual patients, thus antibody reactions may change over time unless antibody reagents identify more constant regions of the parasite surface.

Treatment. Therapy has generally been with suramin or pentamidine, although these drugs do not penetrate the central nervous system well. Arsenicals like melarsoprol have been used to treat more advanced central nervous system disease, although these agents are quite toxic. DL-α-Difluoromethylornithine (eflornithine) holds promise for treatment of all stages of disease, especially in Gambian trypanosomiasis. Alternative approaches to the control of trypanosomiasis have been to administer pentamidine as prophylaxis every 6 months in endemic areas or to attempt to eradicate fly vectors. Neither of these approaches has met with great success.

American Trypanosomiasis

Classification. American trypanosomiasis or Chagas' disease, although caused by a similar-appearing parasite, is different clinically and epidemiologically from African trypanosomiasis. Chagas' disease is caused by *T. cruzi*, an organism that differs from other trypanosomes infecting humans in that it has an intracellular amastigote stage in cardiac muscle and other tissues, as well as trypomastigote forms circulating in the blood. Unlike other trypanosomes, *T. cruzi* does not multiply in the blood. The amastigote stage replicates intracellularly within the tissues. *T. cruzi* is found in most of Central and South America and is transmitted to humans by the bite of reduviid bugs.

Clinical Manifestations. Chagas' disease may present clinically as an asymptomatic, acute, or chronic disease process. Some individuals may manifest all three stages of the disease. The disease is most severe in children younger than 5 years of age and frequently presents as an acute process with central nervous system involvement. The disease is usually more subacute or chronic in older children and adults. After the bite of the vector insect, the organisms proliferate locally in the tissues, producing an erythematous indurated area called a chagoma. These lesions frequently appear on the face and if present around the eye are known as Romaña's sign.

Acute infection is characterized by fever, chills, malaise, myalgia, and fatigue. Parasites may be present in the blood during the acute phase; however, they are sparse in patients older than 1

year. An acute attack may terminate in a few weeks in death or recovery, or the patient may enter the chronic stage of infection. Infected young children may develop central nervous system involvement and die within a few days or weeks.

In the chronic phase of disease, organisms proliferate as amastigotes in cardiac, skeletal, smooth muscle, and neurologic cells as well as the cells of the reticuloendothelial system. Damage to the autonomic nervous system of the heart and digestive tract as well as to the musculature of these organs produces the characteristic cardiac conduction and esophageal and colonic motility defects seen in chronic Chagas' disease. The chronic phase of disease commonly presents with myocarditis, megacardium, and electrocardiographic changes. Less commonly patients may have impaired esophageal and gastrointestinal motility and massive enlargement of the digestive tract, particularly the esophagus and colon (megaesophagus and megacolon). Hepatosplenomegaly is also apparent. Involvement of the central nervous system may produce a meningoencephalitis. Death from chronic Chagas' disease results from heart failure, cardiac arrhythmias, and other complications secondary to tissue destruction in the many areas invaded by the organisms.

Life Cycle. In the vertebrate host, *T. cruzi* occurs in the blood as a nondividing trypomastigote form. The life cycle of *T. cruzi* (Fig. 60.15) differs from that of the other *Trypanosoma* species with the development of an intracellular amastigote form within the tissues. The amastigote multiplies within the parasitized cells by binary fission, thus causing disruption of the cell. The released amastigotes then invade other cells and multiply within them. After a period of growth within the tissues, the amastigotes transform into trypomastigotes and are released into the circulating blood, where they may be ingested by new insect vectors. Within the reduviid bug, the organism transforms into the epimastigote stage and then back into the infective trypomastigote. The infective trypomastigotes pass out of the bug in its feces, are deposited on the skin during feeding, and then may be rubbed into the bite or a skin abrasion. The infective stages are almost immediately ingested by macrophages and transformed into amastigotes.

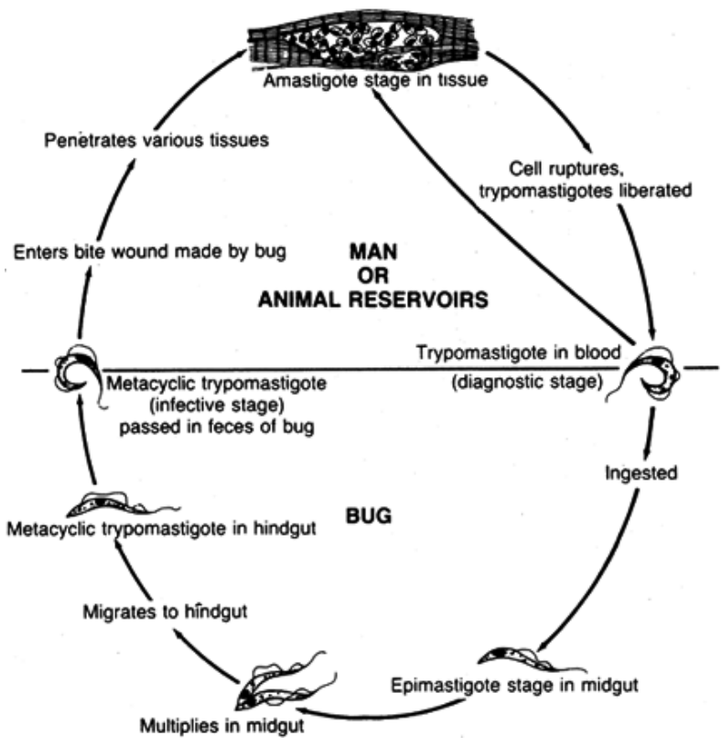


FIGURE 60.15. Life cycle of *Trypanosoma cruzi*. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

Diagnosis. The diagnosis of Chagas' disease is based on the visualization of trypomastigotes in blood during the initial acute phase or during febrile periods in the chronic stage of the disease. The highest yield occurs during the early acute phase; latent or chronic disease is seldom associated with enough parasitemia to make direct detection possible. Amastigote forms may be found in biopsies of affected tissue (Fig. 60.16) or in aspirates from spleen, liver, lymph node, or bone marrow. Culture of blood or inoculation into laboratory animals may be useful when the parasitemia is low. In endemic areas, xenodiagnosis is widely used. Trypanosome-free reduviid bugs are allowed to feed on the patient suspected of having Chagas' disease, and the feces of the bug are examined for the appearance of parasites over a period of approximately 10 days. PCR amplification of nuclear or kinetoplast DNA has been shown to be very sensitive; however, it is not routinely available except in specialized centers (1, 5, 24).

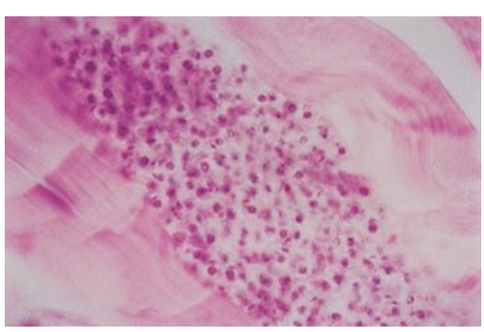


FIGURE 60.16. *Trypanosoma cruzi*: intracellular amastigotes in skeletal muscle (hematoxylin and eosin stain; original magnification ×1,280). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Diagnosis of chronic disease is largely based on serologic testing. Several serologic tests are available, including direct hemagglutination, complement fixation, immunofluorescence assays, and ELISA. No single method is thought to be adequate, and most patients are evaluated by multiple assays before making a final diagnosis (3, 4, 24).

Treatment. Treatment of Chagas' disease is limited by the lack of reliable agents. The current drug of choice is nifurtimox, a nitrofuran derivative. Although it has some activity against the acute phase of disease, it has little activity against tissue amastigotes and has several side effects. Alternative agents include allopurinol and benzimidazole, an imidazole compound. Another imidazole, the antifungal agent ketoconazole, has been shown to

inhibit the intracellular multiplication of *T. cruzi* amastigotes and to afford protection to mice infected with *T. cruzi* (30).

Prevention of infection is also important. Vector control through the use of insecticide treatment of reduviid bug infected dwellings, eradication of nests, and construction of homes to prevent nesting of bugs is essential. Screening of blood by serologic means or excluding blood donors from endemic areas prevents some infections that would otherwise be associated with transfusion therapy.

Toxoplasma gondii

Classification

Toxoplasma gondii is a typical coccidian parasite related to *Plasmodium*, *Babesia*, *Isospora*, and other members of the phylum Apicomplexa. *T. gondii* is an intracellular parasite and is found in a wide variety of animals, including birds and humans. There is only one species, and there appears to be little strain-to-strain variation. Human infection with *T. gondii* is ubiquitous; however, it is increasingly apparent that certain immunocompromised patients are more likely to have severe manifestations. The essential reservoir host of *T. gondii* is the common house cat and other felines.

Clinical Manifestations

Multiple clinical syndromes have been described for infection with *T. gondii*. Most common, however, is asymptomatic infection in normal hosts (31). When symptomatic disease occurs, the infection is characterized by intracellular replication of the organisms, cell destruction, and eventual cyst formation.

A small percentage of otherwise normal individuals on first exposure to the organism will develop an acute illness with cervical or generalized adenopathy. The adenopathy is typically nontender and nonfluctuant; however, occasionally the nodes can be quite tender. In some individuals, the acute infection may resemble infectious mononucleosis with fever, malaise, headache, night sweats, myalgias, sore throat, maculopapular rash, and hepatosplenomegaly (31). Atypical lymphocytes may be present on peripheral smear. A small percentage of these patients can have major complications, including retroperitoneal adenopathy with severe abdominal pain, chorioretinitis, myocarditis, pneumonia, or encephalitis. Most individuals with acute infection, however, have a benign course with spontaneous recovery in days to weeks.

Congenital infection with *T. gondii* also occurs in infants born to mothers who are infected during pregnancy. The likelihood of transmission of the organism from mother to infant increases during the course of gestation, with infection occurring in approximately 25% of infants exposed in the first trimester, 54% in the second, and 65% in the third. If infection occurs in the first trimester, the result is spontaneous abortion, stillbirth, or severe disease. Manifestations of infection in the infant infected after the first trimester include epilepsy, encephalitis, microcephaly, intracranial calcifications, hydrocephalus, psychomotor or mental retardation, chorioretinitis, strabismus, blindness, anemia, jaundice, rash, pneumonia, diarrhea, and hypothermia (31).

Infants may be asymptomatic at birth only to develop disease months to years later. Most often, these children develop chorioretinitis with or without blindness, or other neurologic problems, including retardation, seizures, microcephaly, and hearing loss.

In immunocompromised older patients, a different spectrum of acquired disease is seen. Reactivation of latent toxoplasmosis is a special problem in these individuals. The presenting symptoms of *Toxoplasma* infection in immunocompromised patients are usually neurologic, most frequently consistent with diffuse encephalopathy, meningoencephalitis, or cerebral mass lesions. Reactivation of cerebral toxoplasmosis has emerged as a major cause of encephalitis in patients with acquired immunodeficiency syndrome (AIDS) (32). The disease is usually multicentric, with more than one mass lesion appearing in the brain at the same time.

Symptoms are related to the location of the lesions and may include hemiparesis, seizures, visual impairment, confusion, and lethargy. Other sites of possible infection include the eye, lung, and testes (31, 33, 34). Although disease is seen predominantly in patients with AIDS, it may also occur with similar manifestations in other immunocompromised patients, in particular in those undergoing organ transplantation.

Life Cycle and Epidemiology

The life cycle of *T. gondii* is shown in Fig. 60.17. The complete life cycle (both asexual and sexual reproduction) occurs, as far as it is known, only in the cat. Humans and other animals are considered to be intermediate hosts in which only asexual reproduction, characterized by proliferative forms and cysts, has been observed. Felines ingest oocysts from the environment and cysts (bradyzoites) from tissues of other species. Organisms develop within the intestinal cells of the cat to form oocysts, which are passed in the feces. Oocysts shed in the feces are immature, becoming infective in 1 to 5 days. Intermediate hosts become infected by ingesting mature oocysts or by ingesting cyst-containing

meat from other infected animals. When ingested by the intermediate host, the organism develops into the rapidly multiplying tachyzoite form that is responsible for the initial spread of infection. The tachyzoites disseminate via the bloodstream to many organs and tissues, where they undergo intracellular replication. Multiplication of the parasites within an infected cell usually leads to death and rupture of the cell, releasing the tachyzoites to infect new cells, or it may lead to formation of a cyst within the infected tissue. The cyst form predominates in chronic or latent toxoplasmosis. Individual cysts contain the slowly developing bradyzoites and measure as much as 0.1 mm in diameter. Disease manifestations in chronic toxoplasmosis are primarily the result of the mechanical presence of the enlarging cyst.

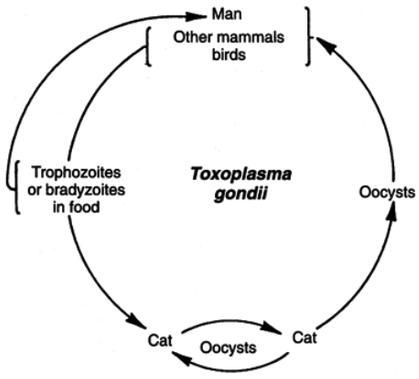


FIGURE 60.17. Life cycle of *Toxoplasma gondii*. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

Human infection occurs after ingestion of viable cysts or trophozoites in undercooked meat or infective oocysts from cat fecal contamination. Because oocysts require several days outside the feline host to mature, infection will generally occur only if there is exposure to soil previously contaminated or to cat feces in a litter box that has not been cleaned for several days. Transmission of infection from contaminated beef, pork, lamb, and goat's milk has been described. Fresh fruit and vegetables may also be a source of infection if the soil they are grown in is contaminated with oocysts. Aside from congenital infection, transmission of infection from person to person has been described only rarely via blood transfusion or solid organ transplantation.

Seroprevalence rates vary from place to place for reasons that remain obscure; however, studies indicate that populations with greater consumption of rare or uncooked meat are more likely to be seropositive. For example, the highest recorded rate (93%) has been reported in Parisian women who prefer undercooked or raw meat (8, 32). Rates may also vary by climate because oocysts do not survive well in colder regions.

Although the rate of seroconversion is similar for individuals within a geographic location, that of severe infection is dramatically affected by the immune status of the individual. Patients with defects in cell-mediated immunity, especially those with HIV infection or after organ transplantation (or immunosuppressive therapy), are most likely to have disseminated or central nervous system disease. Illness in this setting is generally thought to be caused by reactivation of previously latent infection rather than new exposure to the organism.

Diagnosis

There are several possible approaches to the diagnosis of toxoplasmosis. The diagnosis may be established by laboratory tests, including direct visualization of the organism in tissue, growth in culture, detection by molecular methods, or more commonly by serologic methods. In immunocompromised individuals, the diagnosis is often presumptive based on typical findings and clinical impression.

Direct examination of tissue suspected of infection is the definitive method of diagnosis (Fig. 60.18). Biopsy specimens from lymph nodes, brain, liver, myocardium, or other suspected tissue or from body fluids including cerebrospinal fluid, amniotic fluid, bronchoalveolar lavage fluid, or peritoneal fluid can be examined directly for the organisms. The most useful stains appear to be direct fluorescent antibody stains and peroxidase-antiperoxidase stains using antibodies to *T. gondii*. The fluorescent antibody stains available to date have not been as sensitive and specific as the immunoperoxidase stain. Newer monoclonal antibody–based fluorescent stains may prove to be better than the older polyclonal versions. In addition, Wright-Giemsa stains of tissue specimens can be utilized to see the organism.

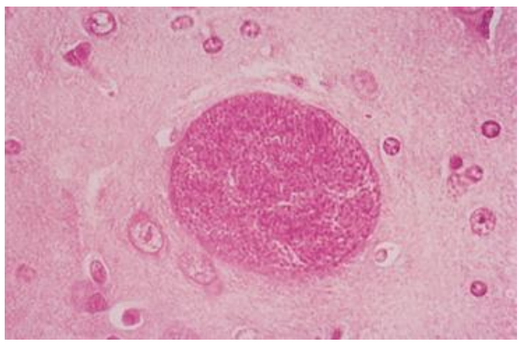


FIGURE 60.18. *Toxoplasma gondii*: cyst in brain section (hematoxylin and eosin stain; original magnification, ×560). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Culture methods for *T. gondii* are largely experimental and are not usually available in clinical laboratories. The two methods available are to inoculate potentially infected material into either mouse peritoneum or tissue culture. Tissue cultures are more readily available, may grow organisms more quickly, and are easier to process because identification simply requires microscopic examination of the culture vial. Mouse inoculation, however, may be more sensitive.

Application of molecular methods, such as PCR, has proven to be highly sensitive and specific for detecting *T. gondii* in cerebrospinal fluid, serum, and amniotic fluid (5, 31, 35). Unfortunately, commercial assays are not available, and these procedures are not widely available. Careful quality control is required to avoid false-positive and false-negative reactions.

Serologic testing is commonly employed for the diagnosis of toxoplasmosis. As with other infectious serologies, attention to increasing immunoglobulin (Ig) G titers in serially collected blood samples or a positive IgM titer is essential to differentiate acute, active infection from previous asymptomatic or chronic infection. Although the Sabin-Feldman dye test is considered the reference method for serologic testing, it is not widely employed in clinical laboratories because of the requirement for live organisms as the antigen source. Several alternative methods are available, including indirect fluorescent antibodies, indirect hemagglutination, complement fixation, and ELISA methods for both IgG and IgM antibodies (31, 36). Currently the ELISA tests for IgG and IgM are the most widely used because of their simplicity and rapidity in documenting acute infections. These tests are not generally satisfactory in immunocompromised individuals with latent or reactivated infections because these patients frequently fail to produce an IgM response or an increasing IgG titer.

Treatment

Therapy of toxoplasmosis depends on the nature of both the infectious process and the immunocompetence of the host. Most mononucleosis like infections in normal individuals resolve spontaneously without intervention. Conversely, disseminated or central nervous system infection in compromised patients must be treated. Before its association with HIV infection, immunocompromised patients with toxoplasmosis were treated for 4 to 6 weeks. In the setting of HIV infection, the rate of relapse when therapy is discontinued is approximately 25%. Such patients are now treated with an initial high-dose regimen of pyrimethamine and sulfadiazine and then continued on lower doses of both drugs indefinitely (37, 38). Although pyrimethamine and sulfadiazine are the agents of choice, side effects, in particular rash and bone marrow suppression, often require changes to alternative agents. Atovaquone and azithromycin (each alone or with pyrimethamine) also have some activity, although their efficacy and safety compared with those of clindamycin-pyrimethamine need to be assessed. Trimethoprim/sulfamethoxazole is not as acceptable as pyrimethamine-sulfadiazine for treatment of disseminated or central nervous system toxoplasmosis. Ocular infections generally respond well to pyrimethamine and sulfadiazine, with treatment periods lasting approximately 1 month.

Infection in the first trimester of pregnancy has been difficult to manage because pyrimethamine is teratogenic in laboratory animals. Both clindamycin and spiramycin have been substituted with apparent success. Spiramycin does not appear to be effective therapy for toxoplasmosis in immunocompromised individuals.

The use of corticosteroids in the treatment of toxoplasmosis remains controversial. It is definitely believed to be of benefit in ocular infections that involve or threaten the macula. The use of steroids for central nervous system infections, even when brain edema is evident on a computed tomographic scan, is not accepted as beneficial. Steroids are generally added in such cases only when progression of clinical symptoms continues despite appropriate antimicrobial therapy.

As more immunocompromised patients at risk for disseminated toxoplasmosis are identified, greater emphasis is being placed on preventive therapy. Routine serologic screening of patients is now being performed before organ transplantation and early in the course of HIV infection. Those with positive serologies are at much higher risk for development of disease and are now being considered for prophylaxis. Trimethoprim/sulfamethoxazole, which is also used as prophylaxis to prevent *Pneumocystis* infections, also appears to be effective at preventing or limiting infections with *Toxoplasma*. Additional preventive measures for pregnant women and immunocompromised hosts should include avoiding both consumption of raw or undercooked meat and exposure to cat feces.

Free-Living Amoebae

Classification

Free-living amebic infections are caused by organisms belonging to the genera *Naegleria*, *Acanthamoeba*, and *Balamuthia* (39, 40). These organisms are ubiquitous in nature and are free-living in soil or water. *Naegleria* species are classified as ameboflagellates and are able to exist temporarily in a flagellate form as well as in an aflagellar ameboid form. *Acanthamoeba* and *Balamuthia* organisms never produce flagella. All three organisms may undergo cyst formation outside the host, but only *Acanthamoeba* and *Balamuthia* form cysts in infected tissue (39).

Clinical Manifestations

Naegleria, *Acanthamoeba*, and *Balamuthia* organisms are opportunistic pathogens. Infections range from acute fulminating primary amebic meningoencephalitis to the more chronic granulomatous encephalitis and amebic keratitis. In most cases, it is thought that the infection was acquired from the environment either by swimming in contaminated water, inhalation of cysts in dust, or direct implantation on the surface of the cornea either by dust or contaminated contact lens material.

Acute primary amebic meningoencephalitis is most commonly caused by *N. fowleri* and occurs when amebae present in water invade the nasal mucosa and extend into the brain. Acute amebic meningoencephalitis may be indistinguishable from bacterial meningitis and is marked by fever, severe frontal headache, stiff neck, nausea, and vomiting. Abnormalities of taste and smell may be present and seizures are common. The process is rapid and may progress to coma within 48 hours. The cerebrospinal fluid is purulent and may contain many erythrocytes and motile amebae. Most cases are fatal.

In contrast to *Naegleria*, central nervous system involvement with *Acanthamoeba* and *Balamuthia* is generally a more subacute or chronic process with a granulomatous encephalitis and single or multiple brain abscesses. Headache and fever are present but are more insidious and low grade. A wide variety of neurologic signs eventually appears, depending on the area of the brain involved. The cerebrospinal fluid may contain both neutrophils and mononuclear cells, but amebae are rare. This form of amebic meningoencephalitis occurs more commonly in immunosuppressed individuals and is uniformly fatal.

Other forms of infection with free-living amebae are usually owing to *Acanthamoeba* and include ocular, cutaneous, pulmonary, and sinus involvement. The most common of these infections is amebic keratitis. The keratitis is usually associated with corneal trauma followed by contact with contaminated soil or water. In recent years, the most common form of amebic keratitis has been associated with improper use of contact lenses, particularly extended-wear lenses. Cases of apparent disseminated cutaneous and subcutaneous infection with *Acanthamoeba* and *Balamuthia* have been described in AIDS patients (39, 41, 42 and 43). These infections present with multiple soft-tissue nodules that on biopsy contain amebae. Central nervous system or deep-tissue involvement may also be present with this form of infection.

Life Cycle

The life cycle of the free-living amebae is relatively simple, involving both cyst and motile trophozoite forms. In addition, *Naegleria* has a flagellate stage that alternates with the ameboid trophozoite phase. Free-living forms of the amebae come into

contact with nasal or respiratory mucosa after exposure to contaminated water or dust. After multiplication locally, invasion of the central nervous system and other tissues occurs by hematogenous spread. In *Naegleria* infection, only the ameboid trophozoites are found within the tissue, whereas with *Acanthamoeba* and *Balamuthia* infection, both trophozoites and cysts are found in tissues. Amebic keratitis occurs by direct contact of the cornea with amebae, which may be introduced through mild corneal trauma or by exposure to contaminated water or contact lens material or solutions. Trophozoites and cysts of *Acanthamoeba* are found in infected corneal tissue.

Diagnosis

Diagnosis of free-living amebic infections is established by demonstrating the organisms in cerebrospinal fluid, nasal discharge, or tissue (corneal scraping, biopsy) by light or phase microscopy, culture, or animal inoculation (39). *Naegleria* and *Acanthamoeba* may be difficult to differentiate in tissue. Isolation in culture may be accomplished by growing the organism on plain agar plates seeded with gram-negative enteric bacilli. *Naegleria* measure 8 to 15 µm in diameter and have blunt pseudopodia, whereas *Acanthamoeba* are 10 to 15 µm and have spinelike pseudopodia. *Balamuthia* trophozoites are irregular in shape, produce broad pseudopodia, and measure from 12 to 60 µm in length. Additional differentiating features are the presence of a flagellate stage with *Naegleria* and the production of a cyst form in tissue with *Acanthamoeba* and *Balamuthia*. Morphologically, *Balamuthia* cannot be differentiated from *Acanthamoeba* by routine histology, although differences may be detected at the ultrastructural level. Immunofluorescent staining has been used to differentiate these organisms in tissue. *Balamuthia* do not grow on agar plates used for *Naegleria* or *Acanthamoeba* but have been recovered in tissue culture using mammalian cell lines.

Treatment

Treatment of free-living amebic infections is largely ineffective. Amebic meningoencephalitis caused by either *Naegleria*, *Balamuthia*, or *Acanthamoeba* is unresponsive to antibacterial and antiamebic drugs. Amphotericin B, a toxic antifungal agent, is the anti-*Naegleria* agent for which there is evidence of clinical effectiveness. Additional agents include miconazole, rifampin, and/or tetracycline, usually administered in combination with amphotericin B. *Acanthamoeba* and *Balamuthia* appear resistant to ketoconazole, miconazole, itraconazole, sulfadiazine, flucytosine, pentamidine, and polymyxin B. These agents may have some effect *in vitro* but have not been adequately tested in patients. Treatment of amebic keratitis may require repeated corneal transplantation. In some cases enucleation of the eye may be required.

INTESTINAL AND UROGENITAL PROTOZOA

Part of "60 - Parasitology"

Protozoa may colonize or infect the intestinal tract, oropharynx, and urogenital tract of humans. Most of these parasites belong to the amebae or flagellates; however, infection with ciliate, coccidian, or microsporidian parasites may also be encountered. In a review of stool specimens examined for intestinal parasites by the College of American Pathologists in the United States, the most common potential pathogen was *Giardia lamblia* followed in order by *Cryptosporidium* species, *Dientamoeba fragilis*, *Entamoeba histolytica*, *Isospora* species, and *Microsporidia* (44). These organisms are generally of worldwide distribution, and most are acquired by fecal-oral contamination.

Amoebae

The amoebae are simple, unicellular organisms. The amoebae that inhabit the intestinal tract of humans include members of four genera: *Entamoeba*, *Endolimax*, *Iodamoeba*, and *Blastocystis* (45). These organisms have a relatively simple life cycle involving an actively motile feeding trophozoite stage and a nonmotile, resistant, infective cyst stage. The cysts are ingested by the host and excyst in the small intestine. The motile trophozoites proliferate in the lumen of the colon. Motility of amebae is accomplished by the extrusion of cytoplasmic foot processes or pseudopods. The cyst form develops in response to environmental changes. Both trophozoites and cysts may be shed in the stool.

Of the amebae that occur in humans, most are considered nonpathogenic commensals, including *E. coli*, *E. hartmanni*, *E. dispar*, *E. gingivalis* (oral commensal), *I. bütschlii*, and *E. nana*. *E. histolytica*, the agent of amebiasis, is the major pathogen in this group. *E. polecki*, a parasite of pigs and monkeys, may occasionally cause diarrhea in humans. The pathogenicity of *B. hominis* is still controversial (45, 46 and 47). These organisms are differentiated one from other by morphologic characteristics as observed on microscopic examination of wet mounts and permanently stained preparations (45) (Table 60.6 and Fig. 60.19).

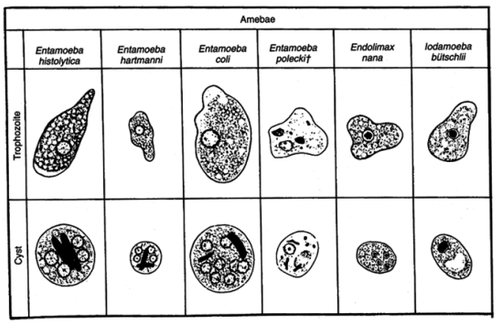


FIGURE 60.19. Amebae found in stool specimens of humans. *Entamoeba polecki* are rare, probably of animal origin. (From Smith JW, McQuay RM, Ash LR, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

Entamoeba histolytica

Classification

E. histolytica is characterized by an actively motile trophozoite form containing a single nucleus with evenly distributed peripheral chromatin granules and a small, centrally located karyosome. The trophozoite ranges in size from 10 to 60 µm in diameter and may contain phagocytized red blood cells (Fig. 60.20). The cyst stage may be as large as 20 µm in diameter and contains four nuclei in the mature form. The presence of rod-shaped chromatoid bodies with rounded ends in the immature cyst form may aid in distinguishing *E. histolytica* from other *Entamoeba* species (Fig. 60.21). It is now recognized that the ameba morphologically identified as *E. histolytica*, is actually two distinct species (45, 48). The pathogenic species is *E. histolytica* and the nonpathogenic species is *E. dispar*. *E. dispar* is differentiated from *E. histolytica* based on specific isoenzyme profiles (zymodemes). The zymodeme profiles represent stable genetic differences and do not interconvert (45, 48, 49 and 50). Additional biochemical, molecular, and immunologic evidence supports the existence of the two species (45,48–50).

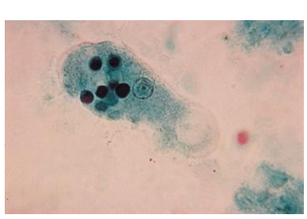


FIGURE 60.20. *Entamoeba histolytica* trophozoite with ingested erythrocytes in fecal smear (trichrome stain; original magnification, ×1,125). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

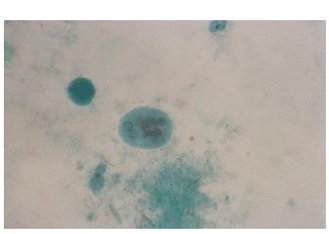


FIGURE 60.21. Mature cyst of *Entamoeba histolytica* containing chromatoidal bodies (trichrome stain of fecal smear; original magnification, ×1,200). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

Clinical Manifestations

The clinical manifestations of amebiasis owing to *E. histolytica* are varied and depend in large measure on the extent of tissue invasion

TABLE 60.6. MORPHOLOGY OF AMEBAE TROPHOZOITES AND CYSTS

	Entamoeba histolytica/ E. dispar[a]	*Entamoeba hartmanni*	*Entamoeba coli*	*Iodamoeba butschlii*	*Endolimax nana*	*Blastocystis hominis*[b]
Trophozoites						
Size (µm)	10–60	5–12	10–50	8–20	6–12	
Motility	Progressive	Nonprogressive	Nonprogressive	Nonprogressive	Nonprogressive	Very difficult to identify rarely seen
Pseudopodia	Fingerlike, explosive	Fingerlike	Blunt, sluggish	Short, blunt, sluggish	Blunt, sluggish	
Cytoplasm	Finely granular, red blood cell inclusions	Finely granular with bacterial inclusions	Rough, vacuolated with yeast and bacterial inclusions	Rough, vacuolated with bacterial inclusions	Rough, vacuolated with bacterial inclusions	
Nucleus	One, evenly distributed chromatin, small centrally located karyosome	One, evenly distributed chromatin small eccentric karyosome	One, unevenly distributed chromatin, large eccentric karyosome	One, no chromatin, large karyosome surrounded by a clear area, consisting of granules not apparent on stained slides	One, no chromatin, large karyosome surrounded by a clear area consisting of granules apparent on stained slides	
Cysts						
Size (µm)	10–20	5–10	10–35	5–20	5–10	6–40
Glycogen	Diffuse, present in young cysts; stains reddish brown with iodine	Diffuse, present in young cysts; stains reddish brown with iodine	Diffuse, well-defined mass in young cysts; stains reddish brown with iodine	Compact, stains dark brown with iodine	Diffuse, occasionally appears in young cysts; stains reddish brown with iodine	Not present
Nuclei	Up to 4	Up to 4	Usually up to eight; may have 16	1	4	Multiple, small surrounding large central body
Chromatoid bodies	Elongated with blunt, round ends	Elongated with blunt round ends	Oval or rodlike with splintered or pointed ends	Not present	Not present	Not present
Clinical manifestations	Amebic dysentery, liver abscess	Commensal, nonpathogenic	Commensal, nonpathogenic	Commensal, nonpathogenic	Commensal, nonpathogenic	Diarrhea
Diagnostic methods[b]	EIA, IFA, HAI, serology, complement fixation, latex agglutination, immunodiffusion, immunoelectrophoresis, O&P exam, histologic exam using PAS stain	O&P exam	O&P exam	O&P exam	O&P exam	O&P exam

[a] Distinction between *E. histolytica* and *E. dispar* cannot be made based on morphology unless ingested red cells are seen in the cytoplasm of the trophozoite.

[b] Description of central body form.
EIA, enzyme immunoassay; IFA, immunofluorescent assay; HAI, hemagglutination inhibition; O&P exam, ova and parasites examination; PAS, periodic acid-Schiff.

and on whether the infection is confined to the intestinal tract or has spread to involve other organs. Intestinal amebiasis is the most common form of infection and is asymptomatic in most individuals (85% to 95% of cases). Asymptomatic individuals may continue to pass cysts in their stools for months and are an important epidemiologic link in the continued transmission of the organism. Although infections with *E. histolytica* may be asymptomatic (45, 51), based on zymodeme analysis, most asymptomatic individuals are infected with the noninvasive *E. dispar* (45, 48, 49 and 50). Between 5% and 15% of cases have definite signs and symptoms ranging from amebic colitis to severe amebic dysentery to extraintestinal complications (51). Amebic colitis has various symptoms related to localized tissue destruction in the large intestine. These intestinal lesions are characterized by small, pinpoint mucosal ulcerations that may expand into the submucosa to form typical flask-shaped ulcers. The more severe dysenteric form of the disease is characterized by severe abdominal cramping and bloody diarrhea. There is extensive invasion of

the mucosa with ulceration that may lead to intestinal perforation and peritonitis, hemorrhage, or secondary bacterial infection and abscess formation. Extraintestinal amebiasis most commonly involves the liver and is characterized by systemic signs of infection (fever, leukocytosis, weight loss) as well as localized tenderness and organomegaly. Rarely, amebic abscesses may appear in other organs, such as lung or brain, either by direct extension from a liver abscess (lung) or by hematogenous dissemination from an intestinal or hepatic focus.

Life Cycle and Epidemiology

The life cycle of *E. histolytica* is illustrated in Fig. 60.22. After ingestion of infective cysts in contaminated food or water, the cysts pass through the stomach, where exposure to gastric acid stimulates excystation and release of the pathogenic trophozoite in the duodenum. The number of trophozoites produced varies with the number of nuclei in the cyst stage. The trophozoites become established and proliferate by binary fission in the large intestine, where they produce extensive local necrosis. Under particular

conditions encystation occurs. The resulting cysts are passed in the feces, thus perpetuating the cycle.

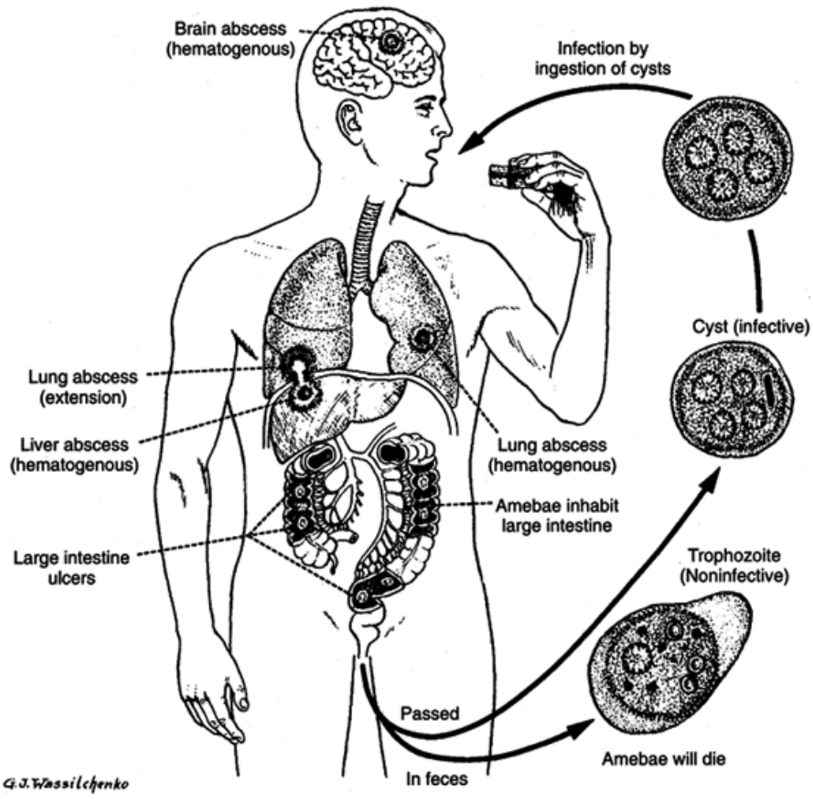


FIGURE 60.22. Life cycle of *Entamoeba histolytica*. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

E. histolytica has a worldwide distribution; however, its highest incidence is in underdeveloped tropical and subtropical regions with poor sanitation and contaminated water. Many of the infected individuals are asymptomatic carriers and represent a reservoir for the spread of *E. histolytica* to others. Both symptomatic and asymptomatic individuals infected with *E. histolytica* may pass infectious cysts, as well as noninfectious trophozoites, in their stools. The trophozoites are unable to survive for prolonged periods in the environment and are destroyed by gastric acid if ingested. Contamination of food and water with infectious cysts is a particular problem in hospitals for the mentally ill, military and refugee camps, prisons, and crowded day care facilities. Sexual transmission of cysts by oral-anal sexual practices is an important means of transmission of amebiasis in homosexual populations.

Diagnosis

Examination of stool specimens will usually allow the diagnosis of intestinal amebiasis. The identification of *E. histolytica* trophozoites and cysts in stools and trophozoites in tissue is diagnostic of amebic infection. Microscopic examination of stool specimens is insensitive, and diagnosis may require collection and examination of multiple stool specimens or a purged series. False-negative stool examinations are attributable to the fact that the parasites are not usually distributed homogeneously in the specimen and are concentrated in the intestinal ulcers and at the margins of abscesses, rather than in the stool or the necrotic center of the abscess. False-positive microscopic examinations may be owing to failure to distinguish *E. histolytica* from other nonpathogenic amebae (Table 60.6 and Fig. 60.19), as well as polymorphonuclear leukocytes. Extraintestinal amebiasis may be diagnosed using radiographic scanning procedures for the liver, lung, or brain. Specific serologic tests are frequently positive in individuals with extraintestinal or symptomatic intestinal disease and are negative in asymptomatic carriers and uninfected individuals. In addition to conventional microscopic and serologic tests, investigators have developed immunologic tests for the detection of antigen in feces as well as PCR and nucleic acid probe assays for the detection of pathogenic strains of *E. histolytica* (45, 49, 50 and 51). Although promising, these newer diagnostic tests are not yet widely available.

Treatment

Treatment of amebiasis varies with the clinical stage of the infection. Asymptomatic carriage may be eradicated with

iodoquinol, diloxanide furoate, or paromomycin. Acute amebic colitis is treated with metronidazole plus iodoquinol. Amoebic dysentery may be treated with emetine, dehydroemetine, or metronidazole followed by iodoquinol. Metronidazole is generally considered the agent of choice in the treatment of extraintestinal amebiasis. Drainage of large amebic abscesses may occasionally be necessary.

Other Intestinal Amoebae

Other amoebae can parasitize the oral and gastrointestinal tract of humans. These include *E. coli*, *E. gingivalis*, *E. hartmanni*, *E. polecki*, *Endolimax nana*, *I. butschlii*, and *B. hominis*. All these amebae, with the possible exception of *E. polecki* and *B. hominis* are considered nonpathogenic commensals. *E. polecki* most commonly infects pigs and monkeys but may cause mild diarrhea in humans. Infections caused by *E. polecki* are confirmed microscopically and are treated in a manner similar to amebiasis caused by *E. histolytica*.

The role of *B. hominis* in human disease remains controversial (45, 46 and 47). When *B. hominis* is present in large numbers in the absence of other pathogens, it may be the cause of gastrointestinal disease (45, 46 and 47). The most common symptoms include recurrent diarrhea, vomiting, and abdominal pain. The diagnosis is made by detection of the organism using routine microscopic examination of stool. Metronidazole appears to be the most appropriate choice of therapy.

The nonpathogenic amebae must be differentiated microscopically from *E. histolytica*. This is particularly true for *E. coli*, which is similar in size to *E. histolytica* and is frequently found in stool specimens from patients exposed to contaminated food or water. Careful microscopic examination of trophozoite and cyst forms present in stained and unstained specimens is essential for the accurate identification of intestinal amebae (Table 60.6 and Fig. 60.21).

Flagellates

There are three clinically significant species of flagellates found in humans, *Giardia lamblia*, *Trichomonas vaginalis*, and *Dientamoeba fragilis* (45). Nonpathogenic commensal flagellates include *Chilomastix mesnili*, *T. hominis*, *T. tenax*, *Enteromonas hominis*, and *Retortamonas intestinalis*.

The trichomonads and *D. fragilis* have only trophozoite stages in their life cycle, whereas the other flagellates have both trophozoite and cyst stages (Fig. 60.23). The flagellates other than *D. fragilis* are easily recognized by their characteristic rapid motility. Additional characteristics include shape, number of nuclei, and the presence of fibrils and other structures, such as undulating membrane, sucking disk, and cytostome (Table 60.7).

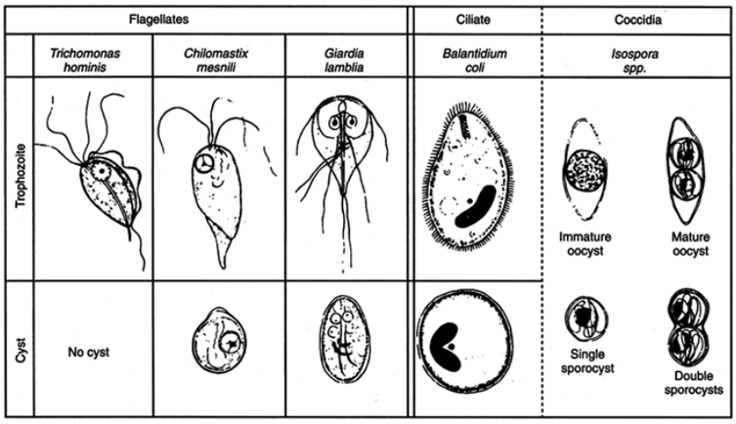


FIGURE 60.23. Protozoa found in human feces (flagellates, ciliate, coccidia). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

TABLE 60.7. MORPHOLOGY OF INTESTINAL AND UROGENITAL FLAGELLATE TROPHOZOITES AND CYSTS

	Giardia lamblia	*Dientamoeba fragilis*	*Trichomonas vaginalis*
Trophozoites			
Size (μm)	10–20	5–15	5–30
Shape	Pear-shaped	Oval to round	Pear-shaped
Motility	"Falling leaf" motility in very diluted preps, otherwise motility is back-and-forth beating of flagella	Ameboid	Jerky, nondirectional
Nuclei	2	2, no chromatin, small karyosome divided into 4–6 distinct granules	1
Characteristic features	Sucking disk on ventral surface	Cytoplasm is finely granular, vacuolated with bacterial inclusions	Short undulating membrane, no free posterior flagellum; axostyle protruding posteriorly
Cysts			
Size (μm)	8–19	No cyst form exists	No cyst form exists
Shape	Oval or elliptical		
Nuclei	4, usually located on one end		
Characteristic features	Flagella and fibrils are positioned longitudinally in cysts; others lie laterally in cyst		
Clinical manifestations	Acute: irritation of mucosal lining, increased mucus secretion, dehydration, and epigastric pain, diarrhea, weight loss, flatulence	Diarrhea, abdominal pain, nausea, anorexia, malaise, fatigue	Vaginal irritation and discharge, urethritis, prostatitis
	Chronic: epigastric pain, flatulence, diarrhea, weight loss, malabsorption		
Diagnostic methods	O&P exam, DFA stain, ELISA, Entero-Test (String test)	O&P exam with permanent stain	Direct, wet mount, culture, DFA stain

O&P exam, ova and parasite examination; DFA, direct fluorescent antibody; ELISA, enzyme-linked immunosorbent assay.

Diseases produced by flagellates are primarily the result of mechanical irritation and inflammation of intestinal and/or urogenital mucosa. The tissue invasion and destruction seen with *E. histolytica* is rare with flagellates.

Giardia lamblia

Classification

G. lamblia (also known as *G. duodenalis*) is characterized by presence of both actively motile trophozoites and sessile cyst forms in

the same specimen. *Giardia* is one of the most easily recognized intestinal parasites. The pear-shaped trophozoite is bilaterally symmetrical containing two nuclei and eight flagella (Fig. 60.23 and Fig. 60.24). The trophozoite ranges from 10 to 20 μm in length and contains a large ventral sucking disk that occupies the anterior half to three fourths of the body. Division is by longitudinal binary fission. The mature cyst form ranges from 8 to 19 μm in length and contains four nuclei. The nuclei are concentrated toward the anterior end and are not distinct in unstained preparations. *Giardia* cysts are recognized by their oval shape, the fibrils present in the cytoplasm, and the anteriorly located nuclei (Fig. 60.23 and Fig. 60.25).

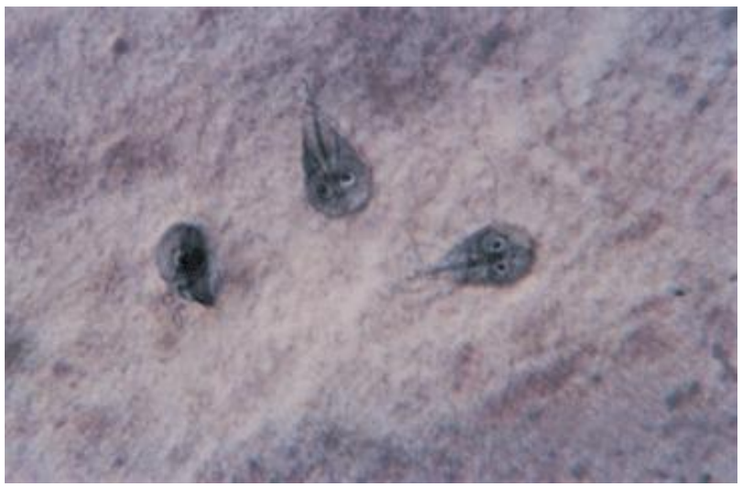


FIGURE 60.24. *Giardia lamblia* trophozoites in fecal smear (Kohn's chlorazol black E stain; original magnification, ×1,125). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

FIGURE 60.25. *Giardia lamblia* cysts in fecal smear (trichrome stain; original magnification, ×1,600). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

Clinical Manifestations

Giardia lamblia causes an infection of the small intestine (Fig. 60.26). Infection can be asymptomatic (in as many as 50% of individuals) or symptomatic. Symptomatic disease may range from mild diarrhea to a full malabsorption syndrome with diarrhea and steatorrhea. The disease may be accompanied by abdominal

bloating, cramps, and flatulence; however, blood and pus are rarely present in stools. The organisms appear to cause disease by both mechanical blockage of the absorptive surface of the small bowel and damage to the mucosal epithelium (45, 52). Frank tissue necrosis as seen in amebiasis does not occur, and spread beyond the gastrointestinal tract is rare. Although spontaneous recovery generally occurs after 10 to 14 days, a more chronic disease with multiple relapses may develop. Individuals with IgA deficiency are particularly prone to developing chronic giardiasis.

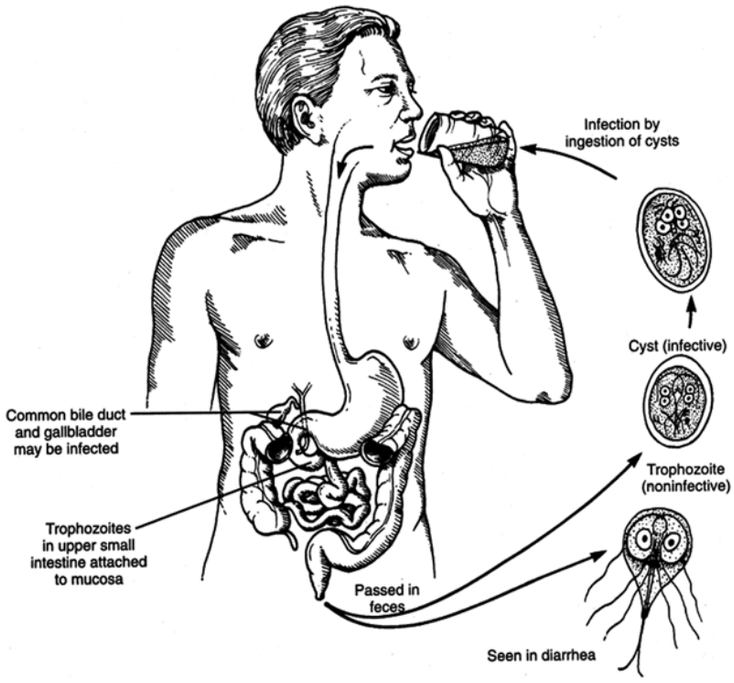


FIGURE 60.26. Life cycle of *Giardia lamblia*. (Modified from Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

Diagnosis

The predominant form seen on stool examination is the cyst form (Fig. 60.26). Trophozoites may be present in diarrheal stools or duodenal aspirates. Organisms may be visualized on a direct wet preparation with or without iodine staining. The use of stool concentration techniques and permanently stained (trichrome) fecal smears is necessary for complete stool workup (45, 52). In addition to conventional microscopy, several immunologic tests for detection of fecal antigen are available commercially. These tests include counterimmunoelectrophoresis, enzyme immunoassay, and indirect immunofluorescent staining. Reported sensitivities range from 88% to 98% and specificities from 87% to 100% (2, 7, 45, 52, 53).

Treatment

Treatment of both asymptomatic cyst passers as well as symptomatic individuals is necessary to interrupt the transmission of *G. lamblia*. The drug of choice is quinacrine hydrochloride (Atabrine). Metronidazole is also effective and may be less toxic. More than one course of therapy may be necessary in some individuals. Prevention and control of giardiasis involve avoidance of contaminated food and water.

Dientamoeba fragilis

Classification

D. fragilis, originally classified as an ameba, is now considered an amebalike flagellate more closely related to the genus *Trichomonas*. The organism lacks a cyst stage and is frequently binucleate. Electron microscopic studies have revealed internal structures typical of a flagellate.

Clinical Manifestations

Many individuals infected with *D. fragilis* are asymptomatic. In some individuals, the organism may cause mucosal irritation of the cecum and upper colon with accompanying symptoms of abdominal discomfort, flatulence, diarrhea, anorexia, and weight

loss (7, 45). The organism has been found in human bile ducts; however, its role in biliary tract disease is unknown.

Life Cycle and Epidemiology

D. fragilis resides in the mucosal crypts. The mode of transmission remains uncertain but most likely is fecal-oral in nature. It is postulated that the delicate trophozoite is transported from person to person inside the protective shell of helminth eggs such as the pinworm *Enterobius vermicularis*.

Diagnosis

Infection with *D. fragilis* is documented by detecting the labile trophozoites in appropriately preserved and stained stool specimens. The excretion of the parasite may fluctuate markedly from day to day, and thus collection of several stool samples may be necessary. Examination of a purged stool sample may also be useful.

Treatment

The antiamebic drugs iodoquinol and tetracycline are often stated to be the drugs of choice for *D. fragilis* infections; however, the cure rates are not high with these drugs. A more effective agent appears to be paromomycin. Metronidazole, furazolidone, and diloxanide furoate may also be effective against *D. fragilis* but require additional investigation.

Trichomonas vaginalis

Classification

T. vaginalis is a parasite of the urogenital system. *T. vaginalis* exists only as a trophozoite and is found in the urethras and vaginas of women and in the urethras and prostates of men. The organism is oval and 5 to 30 µg wide and possesses four flagella and a short undulating membrane that are responsible for motility (Fig. 60.27). There are a single large nucleus and an axostyle composed of microtubules that project from the posterior end. The organism is actively phagocytic and grows optimally under anaerobic conditions.

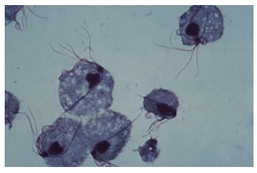


FIGURE 60.27. *Trichomonas vaginalis* trophozoites (Giemsa stain; original magnification, ×2,400). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

Clinical Manifestations

Infection with *T. vaginalis* may be asymptomatic in both women and men. Symptomatic infection in women may range from a mild watery discharge to severe vaginitis with extensive inflammation and erosion of the epithelial lining, associated with burning, itching, and dysuria. Infection in men may involve the urethra, prostate, and/or seminal vesicles. A urethral discharge accompanied by dysuria, nocturia, and prostatic tenderness may be observed.

Life Cycle and Epidemiology

T. vaginalis is strictly a parasite of the urogenital system, where it exists as a motile trophozoite. Transmission is primarily by sexual contact or, less commonly, by contaminated fomites (toilet articles, clothing). Perinatal acquisition of disease is reported secondary to passage through the mother's infected birth canal.

Diagnosis

Diagnosis of infection with *T. vaginalis* is usually made by examining wet mounts of vaginal discharge or prostatic secretions or by examination of urine from both men and women. Stained or unstained smears may be examined (Fig. 60.27). The diagnostic yield may be improved by culturing the organism (93% sensitivity) or using monoclonal fluorescent antibody staining (85% sensitivity) (45). Serologic tests may be useful in epidemiologic surveillance.

Treatment

The treatment of choice is metronidazole; however, resistance has been reported, and retreatment with higher doses may be required. Both male and female sexual partners should be treated to avoid reinfection.

Ciliates

Balantidium coli is the only ciliate known to cause infection in humans (45). The organisms inhabit the large intestine, cecum, and terminal ileum where they may cause extensive ulceration.

Balantidium coli

Classification

B. coli is a large ciliated protozoa ranging in size from 40 to more than 200 µm in greatest dimension. The trophozoite is covered uniformly with rows of hairlike cilia that aid in motility. The organism is structurally more complex than an ameba, with both macronuclei and micronuclei, a cytostome, and numerous cytoplasmic contractile and food vacuoles (Fig. 60.28). The organism may exist in both trophozoite and cyst stage. Reproduction is by binary fission, although conjugation may occur.

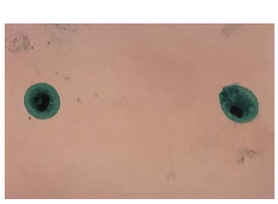


FIGURE 60.28. *Balantidium coli* cyst **(left)** and trophozoite **(right)** in fecal smear (trichrome stain; original magnification, ×180). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal protozoa*. Chicago: American Society of Clinical Pathologists, 1976.)

Clinical Manifestations

Infection with *B. coli* may be asymptomatic or present with a picture of severe intestinal involvement that can mimic the intestinal phase of amebiasis. Symptomatic disease is characterized by abdominal pain, cramping, and tenderness accompanied by nausea, tenesmus, and watery stools with blood and pus. Intestinal ulceration may be extensive, but metastatic (extraintestinal) spread is rare.

Life Cycle and Epidemiology

The life cycle of *B. coli* is simple, involving ingestion of cysts in contaminated food or water, excystation, and invasion of trophozoites into the mucosa of the large bowel. *B. coli* is distributed worldwide, but infection is rare. In the United States, epidemics have occurred in mental hospitals. Swine constitute an important reservoir in other parts of the world. Risk factors for human disease include contact with swine and substandard hygienic conditions.

Diagnosis

As in other intestinal infections caused by protozoa, the diagnosis of balantidiasis is confirmed by detection of the large cysts and trophozoites on microscopic examination of an appropriately collected stool specimen (Fig. 60.28). The organism is detected easily in fresh, wet microscopic preparations.

Treatment

The treatments of choice are tetracycline and iodoquinol. Metronidazole and paromomycin are alternative antimicrobials; however, mixed results have been reported with metronidazole. Preventive measures include appropriate personal hygiene and maintenance of sanitary conditions.

Coccidia

Coccidia are intracellular parasites that are part of a large group called Apicomplexa. This group includes intestinal parasites such as *Isospora*, *Cryptosporidium*, and *Cyclospora* as well as the blood and tissue parasites discussed earlier (*Plasmodium*, *Babesia*, *Toxoplasma*). All coccidia demonstrate both sexual (sporogonic) and asexual (schizogonic) reproductive cycles.

Isospora

Classification

Isosporiasis is caused by *I. belli*, a coccidian parasite of the intestinal epithelium. The organism is present in the stool in the form of oocysts, which average 30 µm in length by 12 µm in width. The mature oocyst contains two sporocysts, each with as many as four sausage-shaped sporozoites (Fig. 60.23).

Clinical Manifestations

Infections range from asymptomatic to significant gastrointestinal disease resembling giardiasis or cryptosporidiosis (54, 55). Symptoms include diarrhea, abdominal pain and cramping, flatulence, malaise, anorexia, weight loss, and low-grade fever. Symptomatic disease is usually self-limited; however, illness may persist for months or years, particularly in immunocompromised individuals such as those with AIDS.

Life Cycle and Epidemiology

Infection with *I. belli* is initiated with the ingestion of mature oocysts in contaminated food or water. After ingestion, excystation occurs in the small bowel releasing sporozoites, which in turn invade intestinal epithelial cells and become rounded trophozoites. The trophozoites undergo schizogony and produce merozoites, which ultimately rupture the host cell and are released to invade other epithelial cells. Some merozoites become sexual gametocytes. The product of sexual reproduction is the oocyst, which is released into the bowel lumen and eliminated in the feces. Fecal oocysts mature within 48 hours and are then infective.

Isospora is distributed worldwide but is rarely reported in humans. In recent years, *I. belli* has been recognized as an important opportunistic pathogen in individuals with AIDS (54, 55). Infection follows ingestion of contaminated food or water or oral-anal sexual contact. There is no evidence of an animal reservoir of *I. belli*.

Diagnosis

The presence of oocysts in stool may be difficult to detect without a concentration procedure. Zinc sulfate or sugar flotation is the most sensitive stool concentration technique. Stool concentrates may be stained with either iodine or a modified acid-fast procedure. Intestinal mucosal biopsy may be necessary to establish the diagnosis when stool specimens are negative.

Treatment

The treatment of choice is combined therapy with pyrimethamine and sulfadiazine or trimethoprim/sulfamethoxazole. Prevention and control are accomplished by adequate sanitation and avoidance of oral-anal sexual contact.

Cryptosporidium

Classification

Cryptosporidium species are coccidian parasites closely related to *I. belli* and *T. gondii*. *C. parvum* is the species associated with all well-documented cases of cryptosporidiosis in mammals (54, 55 and 56). *Cryptosporidium* has an apical complex but produces neither cilia nor flagella. The parasite is characterized by an oocyst without sporocysts and containing four sporozoites. Development of the oocyst takes place intracellularly just beneath the surface membrane of the intestinal epithelial cell and is thus extracytoplasmic.

Clinical Manifestations

Infection with *C. parvum* is usually asymptomatic or a mild, self-limited enterocolitis characterized by watery diarrhea without blood. Immunocompromised individuals may have severe diarrhea characterized by 50 or more stools per day with marked fluid loss, abdominal discomfort, nausea, anorexia, fever, and weight loss. Infection in these individuals may persist for months to years. In patients with AIDS, *Cryptosporidium* has been found in sputum, lung biopsy specimens, and the biliary tract (54, 55 and 56).

Life Cycle and Epidemiology

The life cycle of *Cryptosporidium* resembles that of other coccidian parasites (Fig. 60.29). Mature oocysts containing sporozoites are the infective stage. Infection is initiated with the ingestion of oocysts in contaminated food or water. After excysting from oocysts in the lumen of the intestine, sporozoites invade host cells within the brush border of the intestinal epithelium and develop into trophozoites. Both asexual and sexual reproduction occurs in the epithelial cells. Sporulated oocysts are released in the feces and serve to transmit the infection from one host to another.

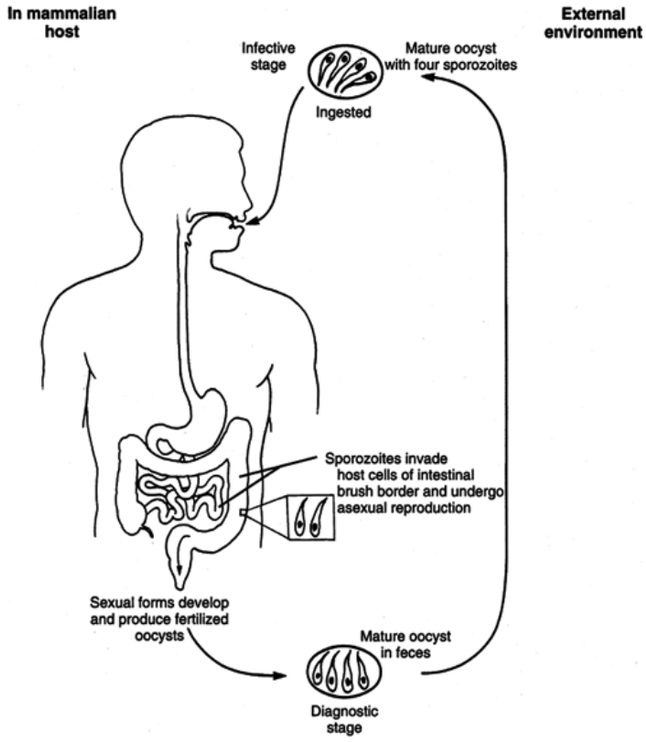


FIGURE 60.29. Life cycle of *Cryptosporidium* species.

Cryptosporidiosis is a zoonosis and can be spread from animal reservoirs to humans, as well as from person to person. The most common means of infection are the fecal-oral and oral-anal routes. Individuals at high risk include veterinarians, animal handlers, and homosexuals. Many outbreaks have now been described in day care centers where fecal-oral transmission is common.

Diagnosis

Cryptosporidium may be detected in large numbers in unconcentrated stool specimens obtained from immunocompromised individuals with diarrhea. Oocysts may be concentrated with the modified zinc sulfate centrifugal flotation technique or by Sheather's sugar flotation procedure (55, 56). Specimens may be stained using the modified acid-fast method or by an indirect immunofluorescence assay (2, 53, 55, 56). An enzyme immunoassay for detecting fecal antigen is also commercially available (2, 55, 56). The number of oocysts shed in stool may fluctuate; therefore, it is recommended that a minimum of three specimens be examined. Molecular (PCR) and serologic procedures for diagnosing

and monitoring infections are under investigation but are not widely available at present (55, 56 and 57).

Treatment

Treatment of symptomatic cryptosporidiosis in immunocompetent individuals is usually not undertaken because of the self-limited nature of the disease and the lack of a safe and effective therapy. Unfortunately, treatment of cryptosporidiosis in immunedeficient persons has been largely ineffective. No controlled studies have been published, and all therapeutic information is based on isolated reports and anecdotal information. Spiramycin, a macrolide antibiotic, may help to control the diarrhea in some patients with cryptosporidiosis during the early stages of AIDS but is ineffective in patients who have progressed to the later stages of AIDS (55, 56). Spiramycin was no more effective than placebo in treating cryptosporidial diarrhea in infants (55,56). Reports concerning the efficacy of azithromycin and paromomycin are promising but will require further confirmation. In the absence of an effective treatment, supportive measures to restore the fluid loss from watery diarrhea appears to be the only intervention available to most clinicians.

Cyclospora

Classification

Cyclospora is a coccidian parasite that is related taxonomically to *C. parvum*, *Isospora* species, and *T. gondii*. *C. cayetanensis* is the single species described thus far infecting humans (55, 58, 60). The oocysts of *Cyclospora* are spherical and, at 8 to 10 µm in diameter, are larger than those of *C. parvum* (4 to 6 µm). The oocysts of *Cyclospora* contain two sporocysts, each of which contain two sporozoites. The sporozoites contain a membrane bound nucleus and micronemes. As with *Cryptosporidium*, the development of the oocyst takes place intracellularly; however, *Cyclospora* develops within the cytoplasm of the host cell, whereas *Cryptosporidium* is extracytoplasmic.

Clinical Manifestations

Cyclosporiasis resembles cryptosporidiosis clinically and is marked by mild nausea, abdominal cramping, anorexia, and nonbloody, watery diarrhea. The diarrhea is generally self-limited in immunocompetent individuals but may be prolonged and last for weeks. Fatigue, malaise, flatulence, and bloating have also been reported to accompany the diarrhea. Clinical illness may be prolonged, severe, and marked by a high rate of recurrence in the immunocompromised host. In patients with AIDS, biliary tract infection has been reported (60).

Life Cycle and Epidemiology

The intermediary life cycle stages of *C. cayetanensis* occur in the cytoplasm of enterocytes located in the small intestine. The oocysts are excreted unsporulated and require 7 to 15 days outside of the host to become fully mature and infectious. Infection is initiated when the mature oocyst is ingested, usually via contaminated water or food (55, 59, 61). Within the lumen of the intestine, the sporozoites are released from the oocyst and invade the epithelial cells. The organism is found within vacuoles in the cytoplasm of jejunal epithelial cells, and its presence is associated with inflammatory changes, villous atrophy, and crypt hyperplasia.

Cyclospora is widely distributed in nature and infects a variety of birds, reptiles, and mammals. Direct animal-to-human or person-to-person spread has not been documented; however, it is now reasonably well established that *Cyclospora* infection is acquired via contaminated water (55, 58, 59, 61). The prevalence of infection ranges from 2% to 18% in endemic areas and from 0.1% to 0.5% in developed countries. In areas of endemicity, the peak incidence of cyclosporiasis coincides with the rainy season. Recent outbreaks in the United States occurred in the summer months and have been linked to contaminated fruit and water (55, 59, 61). *Cyclospora* oocysts are resistant to chlorination and remain viable and infectious when stored in potassium dichromate. Like *Cryptosporidium*, *Cyclospora* is not easily detected by methods used currently to ensure the safety of drinking water.

Diagnosis

The oocysts of *Cyclospora* may be detected in stool by light microscopy. In a wet mount preparation, the unstained oocysts appear as nonrefractile, spherical to oval, slightly wrinkled bodies measuring 8 to 10 µm in diameter with an internal cluster of membrane-bound globules. In fresh specimens, *C. cayetanensis* will fluoresce when examined with a fluorescence microscope fitted with a 365-nm excitation filter.

Cyclospora oocysts may be concentrated from stool with the modified zinc sulfate centrifugal flotation technique or by Sheather's sugar flotation technique. The organisms stain variably acid fast with the modified Ziehl-Nielsen stain or the Kinyoun acid-fast stain. The appearance on acid-fast staining ranges from unstained to mottled pink to deep red (Fig. 60.30).

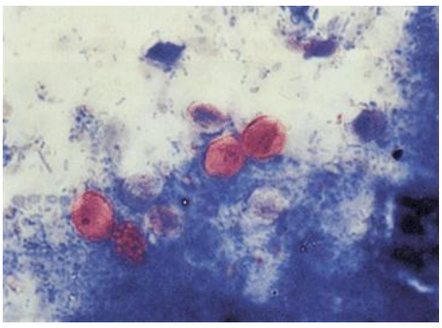


FIGURE 60.30. *Cyclospora* oocysts, acid-fast stain. (From Ortega YR. *Cryptosporidium*, *Cyclospora* and *Isospora*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1406–1412.)

The relative sensitivity and specificity of the available methods for diagnosing *Cyclospora* infection are unknown. Currently there are no commercially available immunodiagnostic methods to aid in the diagnosis and monitoring of these infections. A PCR assay has been described but has not been evaluated for use in a clinical laboratory (55, 60). The lack of understanding of the disease process coupled with rather poor diagnostic techniques all contribute to an underrecognition of *Cyclospora* infection.

Treatment

In HIV-infected individuals with cyclosporiasis, it appears that the high rate of recurrence may be attenuated with long-term suppressive therapy with trimethoprim/sulfamethoxazole (55, 59). The efficacy of trimethoprim/sulfamethoxazole therapy has been demonstrated in anecdotal reports, a large open-label study of HIV-positive individuals, and a placebo-controlled trial (55, 59, 60, 62). The effectiveness of numerous additional agents, including metronidazole, norfloxacin, quinacrine, nalidixic acid, tinidazole, and diloxanide furoate, has not been proven.

Prevention of *Cyclospora* infection is difficult and is compounded by the organism's resistance to chlorination. Regardless, treatment of water by chlorination and filtration remains a reasonable practice. Furthermore, the same methods of improved personal hygiene and sanitation used for other intestinal protozoa should be employed as preventive measures for this disease.

Microsporidia

Classification

Microsporidia are obligate intracellular pathogens belonging to the phylum Microspora. The parasites are characterized by the structure of their spores, which have a complex tubular extrusion mechanism used for injecting the infective material (sporoplasma) into cells. Microsporidia are considered to be primitive eukaryotic organisms because they lack mitochondria, peroxisomes, Golgi membranes, and other typically eukaryotic organelles.

Microsporidia have been detected in human tissues and have been implicated as participants in human disease (63, 64 and 65). To date, six genera (*Enterocytozoon*, *Encephalitozoon*, *Nosema*, *Pleistophora*, *Vittaforma*, and *Trachipleistophora*) and unclassified *Microsporidium* species have been reported in humans. *Enterocytozoon bieneusi* and *Encephalitozoon* (*Septata*) *intestinalis* have gained increasing attention as important causes of chronic diarrhea in patients with AIDS (66, 67); both *Encephalitozoon*-like and *Enterocytozoon*-like organisms have been reported in the tissues of AIDS patients with hepatitis and peritonitis (64, 65 and 66); *Pleistophora*, *Trachipleistophora*, and *Nosema* are known to cause myositis in immunocompromised patients (64, 66); *Nosema* has caused localized keratitis as well as disseminated infection in a child with severe combined immunodeficiency (64, 65 and 66); and *Microsporidium* species and *Encephalitozoon hellem* have caused infection of the human cornea (64, 66).

Clinical Manifestations

Clinical signs and symptoms of microsporidiosis are quite variable in the few human cases reported. Intestinal infection owing to *E. bieneusi* in AIDS patients resembles that which occurs in patients with cryptosporidiosis and cyclosporiasis. Patients experience chronic watery diarrhea without blood or mucus. The diarrhea is persistent and debilitating, leading to dehydration and extreme wasting. The clinical presentation of infection with other species of *Microspora* is dependent on the organ system involved and ranges from localized ocular pain and loss of vision (*Microsporidium* and *Nosema* species) to neurologic disturbances and hepatitis (*Encephalitozoon cuniculi*) to a more generalized picture of dissemination with fever, vomiting, diarrhea, and malabsorption (*Nosema*). In the latter, disseminated infection with *N. connori*, the organism was seen involving the muscles of the stomach, bowel, arteries, diaphragm, and heart and the parenchymal cells of the liver, lungs, and adrenals (64, 65).

Life Cycle

Microsporidia are obligate intracellular parasites with a wide host range among invertebrate and vertebrate animals. Transmission is by ingestion of spores that have been shed in the urine and feces of infected animals. After ingestion by a new host, pressure builds inside the spore, causing the polar filament to evert and injecting the sporoplasm with its nuclear material into an adjacent cell in the intestine, where it multiplies and is disseminated (64, 65). Although some species are highly selective in the cell type that they invade, collectively, the Microspora are capable of infecting every organ of the body.

Once inside a suitable host cell, the microsporidia multiply extensively, either within a parasitophorous vacuole or free within the cytoplasm of the host cell. The intracellular multiplication includes a phase of repeated divisions by binary fission (merogony) or multiple fission (schizogony) and a phase culminating in spore formation (sporogony). Parasites spread from cell to cell to continue merogony, which may overlap sporogony. During sporogony, sporonts divide into sporoblasts and sporoblasts mature into spores containing the infective sporoplasm. The mature spores are then excreted into the environment and the cycle continues.

Diagnosis

There are several techniques that have been useful for recognizing microsporidia in human material (63, 64 and 65, 68, 69). Diagnosis is frequently made by identification of the organisms in biopsy material and by light microscopic examination of cerebrospinal fluid and urine. Spores measuring between 1.0 and 2.0 µm may be visualized by Gram (gram-positive), acid-fast, periodic acid–Schiff, Calcofluor, and Giemsa staining techniques (Fig. 60.31). Giemsa-stained organisms are easier to see in histologic sections than those stained with hematoxylin and eosin. Immunohistochemical techniques may also be useful. Electron microscopy is considered the gold standard for diagnostic confirmation of microsporidiosis; however, its sensitivity is unknown. Species identification requires transmission electron microscopy.

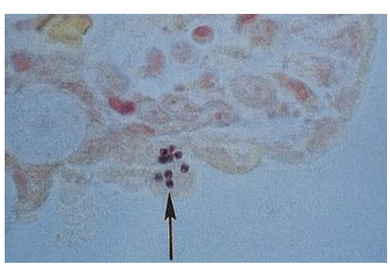


FIGURE 60.31. Gram-positive spores of microsporidia (*arrow*) in jejunal biopsy (Brown-Brenn stain; original magnification, ×1,750). (From Weber R, Bryan RT, Owen RL, et al. Improved light-microscopical detection of microsporidia spores in stool and duodenal aspirates. *N Engl J Med* 1992;326:161–166.)

Chromotrope-based staining techniques for light microscopic detection of microsporidial spores in stool or duodenal aspirates have proven useful (64, 65, 69, 70). The chromotrope-based techniques detect microsporidia spores on light microscopy of unconcentrated formalin-fixed fecal material. The spores show a characteristic pinkish-red staining pattern and can be distinguished

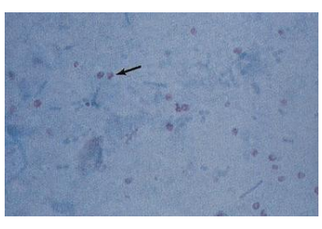


FIGURE 60.32. Smear of formalin-fixed stool specimen, showing pinkish red–stained microsporidia spores, some with a distinct pinkish beltlike stripe (*arrow*). Bacteria are stained faint green (chromotrope-based stain; original magnification, ×1.800). (From Weber R, Bryan RT, Owen RL, et al. Improved light-microscopical detection of microsporidia spores in stool and duodenal aspirates. *N Engl J Med* 1992;326:161–166.)

easily from bacteria, yeasts, and fecal debris (Fig. 60.32). This technique may serve as a practical noninvasive means of diagnosing microsporidiosis.

Additional diagnostic techniques, including culture and serologic testing, are currently under investigation (63, 64 and 65, 68). Nucleic acid-based methods, such as PCR, have been developed to detect or identify microsporidia but are largely relegated to research laboratories at the present time (63). The role of these techniques in the diagnosis of microsporidiosis has yet to be validated in clinical and epidemiologic studies (63, 64 and 65).

Treatment

There is no known effective treatment for microsporidial infections (63, 64 and 65, 71, 72, 73 and 74). Treatment with metronidazole has resulted in temporary improvement in patients with intestinal microsporidiosis. Likewise, some patients have been treated with sulfa drugs or albendazole and survived (63). *In vitro* studies have demonstrated activity of trimethoprim/sulfamethoxazole and the antibiotic fumagillin against Encephalitozoon (71, 72). Oral fumagillin has been shown to eradicate *E. bieneusi* from patients with HIV-associated diarrhea but was associated with unacceptable toxicity (74).

HELMINTHS

Part of "60 - Parasitology"

Helminths may assume both free-living and parasitic life forms and are classified into nematodes (roundworms) and platyhelminths (flatworms). The helminths are generally macroscopic, and the adult worms may vary tremendously in size from barely visible to 10 m in length.

Appropriate and timely diagnosis of a helminthic infection is dependent on a clear understanding of the life cycles of the parasites, the tissues likely to be involved, and the geographic distribution of the organisms (75, 76). The life cycles of helminths may be quite complex and include both direct and indirect cycles. The direct cycle requires only the definitive host, whereas indirect cycles require an intermediate host in which larval stages develop as well as definitive hosts, which harbor the adults.

The clinical signs and symptoms of helminthic infections depend on the location of the organisms and may be caused by adults, larvae, or eggs. The host response to the presence of organisms may be prominent and often includes eosinophilia, especially in the early stages of infection when the parasites are in tissue.

The final diagnosis is usually dependent on detection and identification of mature or developmental (larvae, embryo, egg) stages of the parasite. Occasionally, the diagnosis must be made clinically or serologically.

The majority of helminths produce characteristic eggs that are passed in the feces and serve as the chief means of diagnosing the infections. The identification of eggs should be approached in a systematic manner taking into account its size and shape, the thickness of the shell, and the presence or absence of specialized structures such as spines, knobs, or opercula (Fig. 60.33). The presence and characteristics of larvae present within the eggs may also be useful.

Nematodes

Nematodes are the helminths most frequently parasitizing humans and include intestinal nematodes as well as blood and tissue nematodes. The most common nematodes of medical importance are those inhabiting the intestinal tract (Table 60.8). Most of these organisms have a direct life cycle, and their presence may be confirmed by detecting the characteristic eggs in feces.

TABLE 60.8. MORPHOLOGY OF INTESTINAL NEMATODES

	Ascaris lumbricoides	*Enterobius vermicularis*	*Trichuris trichiura*	Hookworm (*Necator americanus* and Ancylostoma duodenale)	*Strongyloides stercoralis*
Method of infection	Ingestion of egg	Ingestion of egg	Ingestion of egg	Skin penetration by *Necator* and skin penetration and ingestion by *Ancylostoma*	Skin penetration
Location in host	Small intestine, migration in bile duct and lungs	Development in intestines, adults in cecum, appendix, colon, rectum	Cecum, appendix, colon, development in intestines	Migration into lungs, adults attached to mucosa of small intestine	Migration into lungs, females in mucosal epithelium of small intestine, autoinfection
Egg morphology	Fertilized egg, 45–75 × 35–50 µm; infertile eggs 85–95 × 43–47 µm	Smooth, thick-shelled eggs 50–60 × 20–32 µm, flattened on one side	Thick-shelled with bipolar plugs 50–54 × 20–23 µm	Thin-shelled 56–75 × 36–40 µm, unable to distinguish between species	Thin-shelled, rarely seen
Larva morphology				*Ancylostoma*: First-stage rhabditiform 250–350 × 17 µm, long buccal canal, genital primordium is small; infective third-stage filariform 625–675 µm long, pointed tail, sheath not striated. *Necator*: first-stage rhabditiform 250–300 × 17 µm, long buccal canal, genital primordium is small. Infective third-stage filariform 580–620 µm long, pointed tail, sheath striated	First-stage rhabditiform 180 × 14–20 µm wide, short buccal capsule, prominent genital primordium. Infective third-stage filariform 500 µm long, tail is notched, no sheath about infective larvae
Adult worm morphology	Females, 20–35 cm long × 3–6 mm wide with straight tail; males, 15–31 cm long × 2–4 mm wide with curved tail, 3 well-developed lips	Females, 8–13 mm long × 0.3–0.5 mm wide with pointed tail; males, 2–5 mm long × 0.1–0.2mm wide with curved tail	Females, 35–50 mm long with straight posterior ends; males, 30–45 mm long with 360˚coil at the caudal end	*Ancylostoma*: adult males 7–11 mm long × 0.4–0.5 mm wide; Females, 10–13 mm × 0.5–0.7 mm wide. Adults have a long buccal capsule containing 2 pairs of teeth. *Necator* adult males and females, 7–11 × 0.3 mm. Adults have a buccal capsule containing cutting plates rather than teeth	There is no parasitic male; females, 2.1–2.7 mm long × 30–40 µm wide
Clinical manifestations	Nutritional deficiencies, pneumonitis, eosinophilia, tissue reaction	Pruritus, perianal itching, nervousness, insomnia, nightmares	Abdominal cramps, severe rectal tenesmus, rectal prolapse	Iron-deficiency anemia, edema of the face and feet, listlessness; may cause hemoglobin levels of ≤ 5 g/dL pruritus, and erythema	Cutaneous: pruritus, erythema, lesions. Pulmonary: pneumonia, cough, shortness of breath, wheezing, fever, transient pulmonary infiltrates. Intestinal: peptic ulcer, abdominal pain; may mimic Crohn's disease, leukocytosis with peripheral eosinophilia
Diagnostic methods	O&P exam, ELISA, CF, HAI	Cellophane-tape prep	O&P exam	O&P exam	O&P exam, HAI
Specimen	Stool	Cellophane-tape slide	Stool	Stool	Stool, sputum
Common name	Large intestinal roundworm	Pinworm	Whipworm	*Necator*: New World hookworm *Ancylostoma*: Old World hookworm	Threadworm

O&P exam, ova and parasite examination; ELISA, enzyme-linked immunosorbent assay; HAI, hemagglutination inhibition; CF, complement fixation.

The filariae are among the most important of the blood and tissue nematodes. They are long, slender roundworms that parasitize the blood, lymph, and subcutaneous and connective tissues of humans. All filariae are transmitted by insect vectors, and most produce larvae called microfilariae that may be demonstrated in the blood, lymph, or connective tissue of the human host.

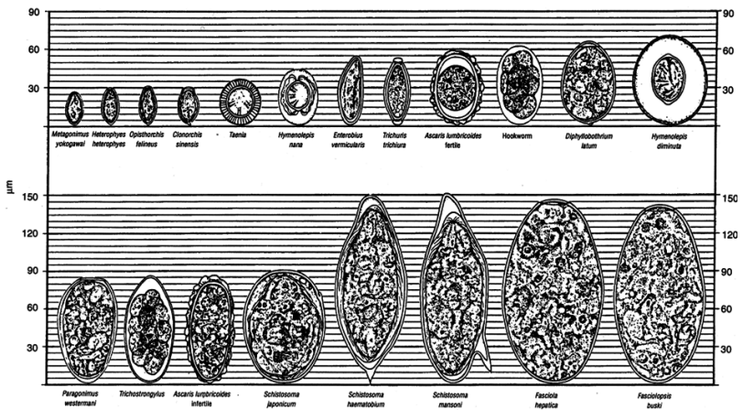


FIGURE 60.33. Helminths found in human specimens: relative sizes. (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Enterobius vermicularis

Infection with *E. vermicularis*, the common pinworm, is the most common helminthic infection in North America. Both female and male worms inhabit the cecum and large bowel. The female worm measures as much as 13 mm in length and is cream colored with bilateral ridges or alae and a sharply pointed tail (Table 60.8 and Fig. 60.34). The male is smaller and rarely seen.

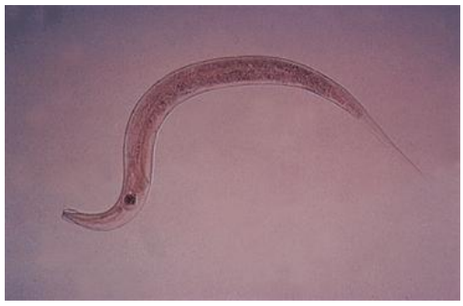


FIGURE 60.34. *Enterobius vermicularis*, adult female (whole mount). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

FIGURE 60.35. *Enterobius vermicularis*, embryonated eggs (low power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

The gravid female migrates from the cecum to the perianal area, where as many as 20,000 eggs are deposited on the perianal skin. The eggs may be transmitted from hand to mouth or inhaled as fomites. The ingested eggs hatch in the small intestine and migrate to the cecum, where they mature into adults in 2 to 6 weeks, thus completing the cycle.

Clinical signs and symptoms of pinworm infection are generally mild. Intestinal pathology is rare; however, inflammation and granuloma formation around eggs has been reported. The role of the worm in appendicitis is questionable despite its occasional presence in the lumen of the resected appendix.

The most common clinical presentation is anal pruritus caused by the secretions of migrating worms. Scratching may lead to further irritation and excoriation, occasionally complicated by secondary bacterial infection. Migration of the worms into the vagina or urinary bladder may produce genitourinary problems and granulomas.

The diagnosis of enterobiasis is usually suggested by the clinical manifestations and confirmed by detection of the characteristic eggs (Fig. 60.35) on the anal mucosa. Detection of the eggs is accomplished by applying the sticky surface of a piece of cellophane tape to the mucocutaneous junction of the anus, applying the tape to the surface of a microscope slide and examining the slide under low power. Multiple samplings may be required to detect the eggs. Occasionally the female worm may be observed on the surface of the stool or on the perianal skin. Sampling should be obtained at night or in the morning before bathing. Systemic signs of infection such as eosinophilia are rare.

The treatment of choice is pyrantel pamoate with mebendazole as an alternative. Because of the potential for cross-infection in the family environment, treatment of all family members is recommended. Although cure rates are high, reinfection is common.

Trichuris trichiura

T. trichiura is the causative agent of whipworm infection. The worm is characterized by a long, thin, threadlike anterior portion and a bulbous posterior portion. The adult whipworm is 30 to 50 mm in length, and the female worm may produce 3,000 to 10,000 eggs per day. The adult worms reside in the cecum and large intestine for as long as 8 years. The whipworm is a cosmopolitan parasite and may infect as many as half a billion people worldwide.

Trichuris is a soil-transmitted helminth with a simple life cycle. The distinctive eggs (Fig. 60.36) are passed in the stool and deposited on soil, where they must mature for at least 10 days before they contain infective larvae. When ingested by the appropriate host, larvae are released in the small intestine and migrate to the cecum, where they penetrate the mucosa and mature to

adults. The adult worms remain attached to the mucosa by their slender anterior portions for the remainder of their lifetime.

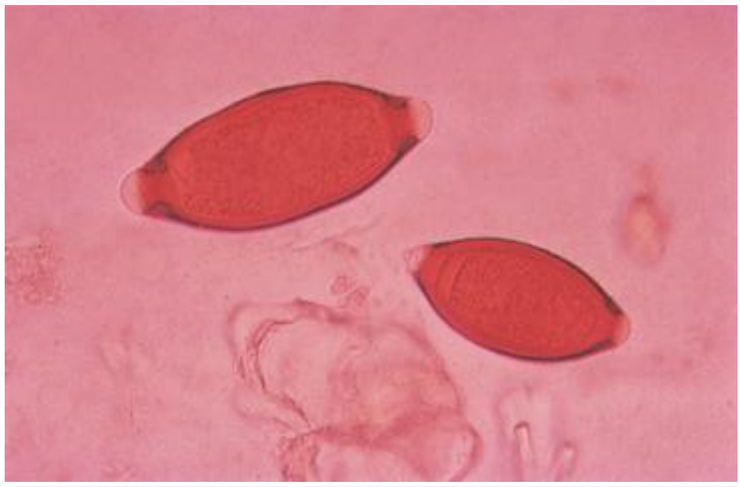


FIGURE 60.36. *Trichuris trichiura*, unembryonated fertile eggs (high power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

The clinical manifestations of trichuriasis are generally related to the intensity of the worm burden. Most infections are with small numbers of *Trichuris* and are usually asymptomatic. Infections with larger numbers of organisms may result in abdominal pain, bloody diarrhea, weakness, and weight loss. Rectal prolapse may occur as a result of straining with defecation. Attachment of the adult worms to the intestinal mucosa may produce localized inflammation, ulceration, and hemorrhage with associated anemia and eosinophilia.

The diagnosis of trichuriasis is confirmed by detection of the characteristic bile-stained eggs with polar plugs (Fig. 60.36) on examination of stool. In light infections, a concentration step

may be required; however, this is not usually necessary in symptomatic infections because they usually produce more than 10,000 eggs per gram of feces.

The drug of choice for symptomatic infections is mebendazole. Asymptomatic infections generally do not require therapy. Prevention depends on adequate sanitation and avoidance of human feces as fertilizer.

Ascaris lumbricoides

A. lumbricoides is the most common and widely distributed of the soil-transmitted helminths. It is characterized by its large size, fusiform shape, and firm pink cuticle (Table 60.8). The adult Ascaris is 20 to 36 cm in length, and the female worm may produce as many as 200,000 eggs per day. The adult worms reside in the small intestine and have a relatively short life span of 6 to 18 months. It is estimated that as many as a billion people, including four million Americans, are infected with *Ascaris* worldwide.

Ascaris has a slightly more complex life cycle than *Trichuris* but is otherwise typical for a soil-transmitted roundworm. The large (35 to 55 µm) mammillated eggs (Fig. 60.33 and Fig. 60.37) are passed in the stool and deposited in soil, where they must mature for 2 to 3 weeks. The eggs are highly resistant to environmental stresses and may remain viable for as long as 6 years. When ingested by the host, the eggs hatch in the small intestine, releasing larvae, which penetrate the mucosa and enter the bloodstream. The larvae are carried to the lungs, where they develop for several days in the alveolar capillary bed. When they develop to the appropriate stage, the larvae rupture into the alveolar space, are coughed up and swallowed, and eventually reach the small intestine, where they grow to adulthood. The entire process from ingested egg to adult takes approximately 2 months.

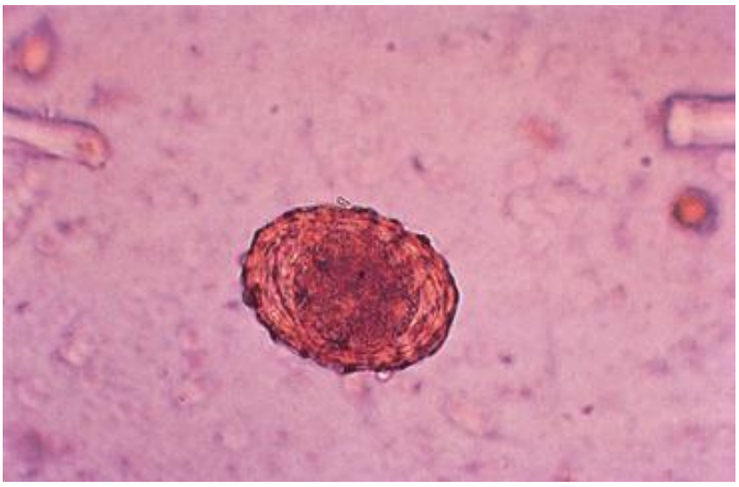


FIGURE 60.37. *Ascaris lumbricoides*, unembryonated fertile egg (low power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Ascariasis is a disease of warm climates and poor sanitation. It is prevalent in areas where human feces are used as fertilizer and may be maintained by small children who defecate outdoors, contaminating the soil, and pick up the infectious eggs on their hands during play. Both food and water may be contaminated with *Ascaris* eggs.

The clinical manifestations of ascariasis may result from either the migration of the larvae through the tissues or the presence of the adult worms in the intestinal lumen. Light infections are generally asymptomatic and may become apparent only when the adult worm is vomited up or passed in the stool. Pulmonary involvement owing to the migration of the larvae to the lungs is related to the degree of hypersensitivity induced by previous infections and the intensity of the current exposure. Pulmonary infiltrates and asthma like signs and symptoms are common and are accompanied by eosinophilia and oxygen desaturation (75). Death from respiratory failure may occur in massive infections.

In the intestinal phase of the infection, a large worm burden may produce nonspecific abdominal discomfort and diarrhea or may cause intestinal obstruction owing to a tangled bolus of worms. Migration of the adult worm to the appendix, bile duct, or pancreatic duct may cause obstruction, inflammation, and tissue damage of the organ. Intrahepatic abscesses may form if the organism reaches the liver. Direct perforation of the intestine resulting in peritonitis has also been reported (75).

Diagnosis of the infection is confirmed by detecting the characteristic eggs in feces or occasionally by examining an adult worm that has been passed in the feces or vomited. The eggs are oval with a thick shell and an irregular, bile-stained, mammillated external layer (Fig. 60.37). They may be fertilized or unfertilized and are usually quite numerous. Large worm burdens may be detected on roentgenographic examination of the small bowel. The pulmonary phase of the disease may be diagnosed by the finding of larvae and eosinophils in sputum.

Treatment of symptomatic infection is highly effective an may be accomplished with either pyrantel pamoate or mebendazole. Mebendazole is preferred if co-infection with *T. trichuris* is also present. Administration of mass therapy at 6-month intervals has been used as a means of community-wide control. Prevention depends on adequate sanitation and avoidance of human feces as fertilizer

Hookworms

The two principal species of hookworms causing disease in humans are *Ancylostoma duodenale*, the Old World hookworm, and *Necator americanus*, the New World hookworm. The adults of both species are pinkish-white, measure approximately 10 mm in length, and have mouthparts designed for sucking blood from injured intestinal mucosa. *A. duodenale* has a long buccal cavity containing two pairs of chitinous teeth, whereas the buccal cavity of *N. americanus* contains shearing chitinous plates rather than teeth (Table 60.8). Both male and female parasites attach to the mucosa of the small intestine. Males may be distinguished from females by the presence of a fan-shaped copulatory bursa at the posterior end. The female worm releases 10,000 to 20,000 thin-shelled eggs daily, which are passed in the feces.

Both *N. americanus* and *A. duodenale* are soil-transmitted helminths, and their life cycles are indistinguishable (Fig. 60.38). The eggs are passed in the feces and embryonate in the soil, where they hatch within 48 hours, releasing rhabditiform larvae. The first-stage larvae develop in the soil over the period of a week to become infective filariform larvae, which then may penetrate intact human skin and migrate via the lymphohematogenous system to the lungs. Similar to *Ascaris*, the larvae rupture into the alveoli, are coughed up, swallowed, and pass to the small intestine, where they mature in 1 to 2 months.

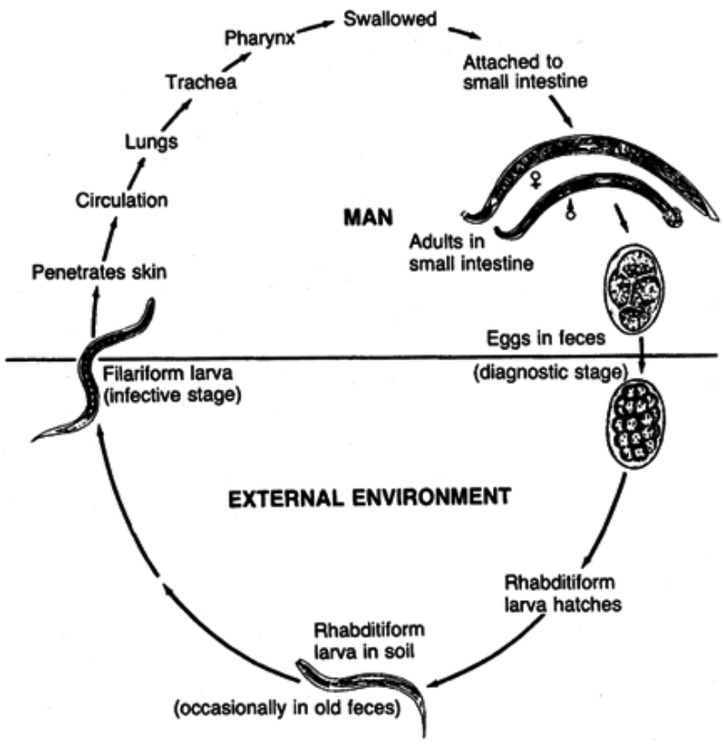


FIGURE 60.38. Life cycle of hookworm. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

Transmission of hookworm infection requires deposition of egg-containing feces on shady, well-drained soil and is favored by tropical or subtropical conditions. It is estimated that approximately 900 million individuals worldwide are infected with hookworms, including 700,000 in the United States (75). Similar to infection with other intestinal helminths, most hookworm infections are asymptomatic because of the very low worm burden. Initial symptomatology may be related to the penetration of the skin by the filariform larvae (pruritic rash) or the migration of the larva through the lung (pneumonitis). The intestinal phase of infection may produce epigastric pain, nausea, vomiting, and diarrhea; however, the major consequences of hookworm infection are anemia and hypoalbuminemia secondary to chronic blood loss. The severity of the anemia is directly related to the worm burden and intake of dietary iron. Daily blood loss is estimated at 0.15 to 0.25 mL for each adult *A. duodenale* and 0.03 mL for each adult *N. americanus*. Severe anemia may develop over a period of months or years and may contribute to mental, sexual, and physical development retardation.

The diagnosis of hookworm infection is confirmed by detection of the distinctive eggs (Fig. 60.39) in direct or concentrated stool preparations. The oval eggs measure 56 to 75 µm long by 36 to 40 µm wide and have a thin shell. The eggs of *A. duodenale* and *N. americanus* are indistinguishable. Rhabditiform larvae may be present if the stool is allowed to stand too long at ambient temperature and must be differentiated from those of *Strongyloides stercoralis* (Table 60.8). Quantitative egg counts in direct wet mounts may be useful in estimating the worm burden:

fewer than five eggs per slide denotes light infections, whereas more than 25 denotes heavy infections.

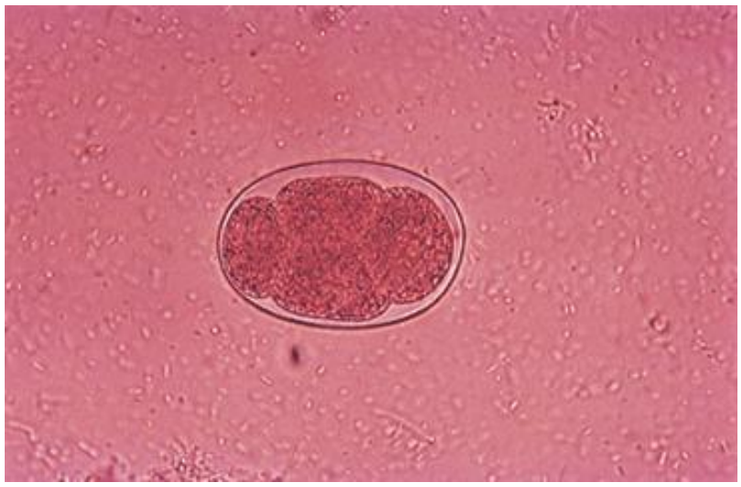


FIGURE 60.39. Hookworm, segmented egg (low power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Treatment of hookworm infection involves antihelminthic chemotherapy and correction of anemia when present. The drug of choice is mebendazole; pyrantel pamoate is an effective alternative. Anemia may be corrected by iron replacement or, if severe, by blood transfusion. Prevention may be accomplished by the wearing of shoes and by improved sanitation.

Strongyloides stercoralis

S. stercoralis, at 2 mm in length, is the smallest intestinal nematode infecting humans. Adult females burrow into the mucosa of the duodenum and reproduce parthenogenetically. The eggs are deposited within the mucosa, where they hatch, releasing rhabditiform larvae that are passed in the stool. Thus, eggs are almost never found in the stools. The rhabditiform larvae, which measure 180 × 16 µm in length, are distinguished from the larvae of hookworms by their short buccal capsule and large genital primordium (Table 60.8).

The life cycle of *S. stercoralis* may be either direct or indirect (Fig. 60.40). The direct cycle is similar to that observed with the hookworms, in which infective filariform larvae develop in soil for at least 1 week, and infection is acquired by the larvae penetrating the skin and migrating through the lungs to the intestine, where the female matures in approximately 1 month. Each female produces approximately a dozen eggs per day, which hatch within the mucosa and release rhabditiform larvae into the lumen of the bowel. The rhabditiform larvae are passed in the stool and may either continue the direct cycle by developing into infective filariform larvae or develop into free-living adult worms and initiate the indirect cycle. In the indirect cycle, free-living adult worms develop in the soil and produce eggs, which ultimately may give rise to infective third-stage filariform larvae. In humans, especially in debilitated individuals or in patients who are immunosuppressed, there may be a third cycle, that of internal autoinfection (75, 77, 78). This autoinfection occurs when the rhabditiform larva's passage through the intestine is delayed, allowing it to transform into an infective filariform larva while still in the body of the host. The larva may then invade the intestinal mucosa or the perianal skin without undergoing the external (soil) phase of the life cycle. This process may give rise to massive superinfection with larvae in many tissues and may lead to death (hyperinfection syndrome) (77, 78).

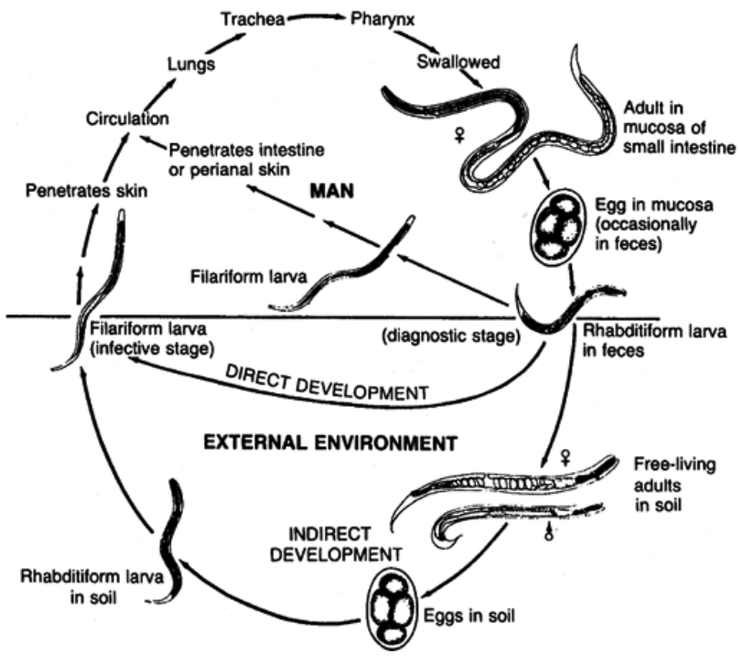


FIGURE 60.40. Life cycle of *Strongyloides stercoralis*. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

S. stercoralis is similar to hookworm in its requirements for warm temperatures and moisture for the external development of larvae in soil. It infects approximately 35 million individuals throughout the tropical and subtropical areas of the world (75, 77).

Although individuals with strongyloidiasis do not manifest cutaneous symptoms (“ground itch”) secondary to the penetration of the infective larvae, they do present with pulmonary symptoms similar to those seen in both ascariasis and hookworm infection. The intestinal infection is usually asymptomatic; however, heavy worm loads may involve the biliary and pancreatic ducts, the entire small bowel, and the colon, causing inflammation and ulceration and leading to epigastric pain and tenderness, vomiting, diarrhea (occasionally bloody), and malabsorption. Symptoms mimicking peptic ulcer disease coupled with peripheral eosinophilia should strongly suggest the diagnosis of strongyloidiasis.

Diagnosis of strongyloidiasis may be difficult owing to intermittent passage of low numbers of first-stage larvae in the stool. In fresh stool specimens, the detection of rhabditiform larvae is indicative of *Strongyloides* infection. Differentiation of *Strongyloides* larvae from those of hookworm may be necessary when stool examination is delayed. First-stage rhabditiform *Strongyloides* larvae have a short buccal cavity and a prominent genital primordium compared to the long buccal cavity and inconspicuous genital primordium present in hookworm larvae. In contrast to hookworm infections, eggs are generally not present in the feces of patients with strongyloidiasis. When absent from the stool, larvae may be detected in duodenal aspirates or in sputum in cases of massive infection. Serologic tests are not generally available.

All infected patients should be treated to prevent autoinfection and potential dissemination (hyperinfection) of the parasite (77, 78). The drug of choice for both mild and severe infection is thiabendazole. Mebendazole is an alternative therapeutic agent. Stools should be checked after therapy to ensure adequate treatment. Patients in endemic areas should be examined thoroughly for the presence of this parasite both before and during steroid treatment or immunosuppressive therapy. Strict infection control measures should be enforced when caring for individuals with hyperinfection syndrome, as stool, saliva, vomitus, and body fluids may contain infectious filariform larvae. Similar to hookworm, *Strongyloides* control requires education, proper sanitation, and prompt treatment of existing infections.

Trichinella spiralis

T. spiralis causes the disease trichinosis. The adult form of *T. spiralis* lives in the duodenal and jejunal mucosa of flesh-eating mammals worldwide. The larval form is present in undercooked or contaminated meat. Among domestic animals, swine are most frequently involved. Humans become infected by ingestion of inadequately cooked meat that contains the encysted larval form of the organism. On ingestion, the encysted larvae are freed by gastric digestion, penetrate the mucosa of the intestine, and mature just above the lamina propria. The adult male and female worms mate in the intestinal mucosa, and the female produces new larvae, which enter the vascular system and are eventually deposited in skeletal muscle. A single female worm may produce larvae over a period of 4 to 16 weeks, generating as many as 1,500 larvae, each measuring 6 × 100 μm (76). Larvae penetrating striated muscle continue to grow and molt and are eventually enclosed in a fibrous capsule, which may calcify over a period of 6 to 18 months. The muscles invaded most frequently include the extraocular muscles of the eye; the tongue; the deltoid, pectoral, and intercostal muscles; the diaphragm; and the gastrocnemius. The encysted larvae remain viable for many years and would be infectious if ingested by a new animal host.

Trichinosis is one of the few tissue parasitic diseases still seen in the United States (76). As with other parasitic infections, most patients have minimal or no symptoms. The clinical presentation depends largely on the tissue burden of organisms and the location of the migrating larvae. Patients in whom 10 or fewer larvae per gram of tissue are deposited are usually asymptomatic, those with 100 or more generally have significant disease, and those with 1,000 to 5,000 have a very serious course that occasionally ends in death (76). The most common clinical manifestations of trichinosis are fever, myalgias, weakness, malaise, periorbital edema, and headache. Less frequent manifestations include skin rash, generalized edema, diarrhea, nausea, splinter hemorrhages, and subconjunctival bleeding. Symptoms generally appear approximately 2 weeks after initial infection and persist for several weeks. Diarrhea, if it occurs at all, tends to be early in infection when adult worms are still in the intestinal tract. Weakness occurs late as an end-stage manifestation of the presence of the calcifying larvae in the muscle tissue.

Trichinosis occurs throughout most of the world except Australia and the Pacific Islands. In the United States, infections are becoming less frequent owing to decreased consumption of pork and pork products, federal guidelines for the commercial preparation of pork products, and federal legislation regarding the preparation and composition of livestock feed to avoid inclusion of infected meat. Nevertheless, it is estimated that more than 1.5 million Americans carry live *Trichinella* cysts in their musculature and that 150,000 to 300,000 acquire new infection annually (76).

The diagnosis of trichinosis is confirmed by performing a muscle biopsy and visualizing the encysted larvae. Key to making the diagnosis is clinical suspicion of the disease and pursuit of the biopsy. The deltoid or gastrocnemius muscle will usually provide the best yield. Often several individuals are infected because of a common exposure to the same poorly prepared meat product. Along with the wide variety of clinical symptoms that are seen, patients usually have an eosinophilia ranging from 15% to 50% of the peripheral white cell count. Together, these epidemiologic, clinical, and laboratory features should lead to suspicion of the disease.

A serologic test using bentonite flocculation is a possible approach to confirming infection. Significant antibody titers are usually absent before the third week of illness but then may persist for years. ELISA tests are also under development, and one such test has been used as a screening method to detect infection in animals as part of the inspection process before the sale of meat products (76, 79).

Treatment of trichinosis is primarily symptomatic because there are no good antiparasitic agents for the tissue larvae. Patients with severe infections may be treated with corticosteroids plus either mebendazole or thiabendazole. Treatment of the adult worms in the intestine with mebendazole may halt the production of new larvae. Prevention is based on thorough cooking of pork and other meat.

Toxocara canis

T. canis, a large intestinal ascarid of canines, is the etiologic agent of toxocariasis or visceral larval migrans. The complete life cycle of this parasite occurs in dogs and is similar to that of the human *Ascaris*. Canines ingest *Toxocara* eggs, which hatch into larvae in the small intestine. These larvae migrate to liver, lung, and tracheae, are swallowed again, and mature into adult worms. Each female worm discharges approximately 200,000 eggs daily into the feces. Passage of larvae from pregnant females to fetuses can also occur and leads to ongoing infection when puppies are born. After reaching the soil, the eggs must embryonate for 2 to 3 weeks and thereafter are infectious for both canines and humans. The eggs may remain viable in soil for months to years.

Human infection occurs on ingestion of infectious eggs. The eggs hatch in the small intestine and the liberated larvae pass through the pulmonary capillaries and reach the systemic circulation. The larvae continue to grow, penetrate the vascular walls, and migrate through various tissues. Rarely, if ever, do the larvae reach the intestine and complete the maturation to adulthood in the human host.

The clinical manifestations of toxocariasis in humans are related to the migration of the larvae through tissues (76). The larvae may invade any tissue of the body, where they can induce necrosis, bleeding, and the formation of eosinophilic granulomas. Patients may be asymptomatic or have only eosinophilia but can also have serious disease that is directly related to the number and location of the lesions caused by the migrating larvae and the degree to which the host is sensitized to the larval antigens. The organs most frequently involved include the lungs, heart, kidney, liver, skeletal muscles, eye, and central nervous system. Signs and symptoms include cough, wheezing, fever, rash, anorexia, seizures, fatigue, and abdominal discomfort. On examination, patients may have hepatosplenomegaly as well as nodular, pruritic skin lesions. Death may result from respiratory failure, cardiac arrhythmia, or brain damage. Ocular disease can also occur with the movement of larvae through the eye and may be mistaken for malignant retinoblastoma. Prompt diagnosis is required to avoid an unnecessary enucleation.

T. canis is a cosmopolitan parasite. Wherever infected canines are present, the eggs are a threat to humans. The incidence of infection appears to be higher in the southeastern sections of the United States. Seroprevalence studies indicate that approximately 4% to 20% of the population has ingested *T. canis* eggs at some time. Children are particularly susceptible to infection, presumably owing to more frequent exposure to contaminated soil and the tendency to put objects in their mouth.

Diagnosis of visceral larval migrans is based on clinical findings, the presence of eosinophilia, known exposure to dogs, and serologic confirmation. ELISA assays are readily available and appear to offer the best serologic marker for disease. Examination of feces from infected patients is not useful because egg-laying adults are not present. Tissue examination for larvae may provide a definitive diagnosis but may be negative owing to sampling error.

Treatment is primarily symptomatic because antiparasitic agents are not of proven benefit. Agents including diethylcarbamazine, thiabendazole, and mebendazole have been tried. Corticosteroid therapy may be lifesaving if the patient has serious pulmonary, myocardial, or central nervous system involvement because a major component of the infection is an inflammatory response to the organism.

Ancylostoma braziliense

Cutaneous larva migrans is an infection of the skin caused by the dog and cat hookworm *A. braziliense*. This species of hookworm naturally infects dogs and cats and accidentally infects humans. Eggs discharged in the feces are deposited on warm, moist, sandy soil, where they develop into filariform larvae capable of penetrating skin on contact. The larvae can penetrate intact human skin but can develop no further in the human host. The larvae remain trapped in the skin and subcutaneous tissue for weeks to months, during which time they continue to migrate, creating serpentine tunnels and provoking a severe erythematous and vesicular reaction. Clinically, the patient notes pruritic, raised, red, linear lesions approximately 10 to 20 cm long (80). Scratching of the irritated skin may lead to secondary bacterial infection. Approximately half of the patients will develop transient pulmonary infiltrates with peripheral eosinophilia (Löffler's syndrome), presumably owing to pulmonary migration of the larvae. Larvae are rarely found in sputum or skin biopsies, and the diagnosis must be made on clinical grounds. The infection is treated adequately with thiabendazole. Antihistamines may be helpful in controlling the pruritus. Preventive measures include wearing shoes, deworming pets, and improvement of sanitation.

Filariae

The filariae are nematodes that are widely distributed in nature. The adults are long and slender, measuring many centimeters in length, and may inhabit virtually any tissue, including blood and lymphatic vessels, pleural and peritoneal cavities, subcutaneous tissue, heart, and brain. Some species migrate in tissues, and others remain localized and may become encased in a fibrous tissue reaction. The adults mate in the tissues, producing many progeny called microfilariae. The microfilariae are small (200 to 300 μm), slender, motile forms, which may be found in the circulating blood or migrating in the subcutaneous tissues, depending on the species.

Microfilariae are classified by morphologic characteristics, geographic location, and type of clinical infection seen (76). The major division of microfilariae is by the presence or absence of a sheath surrounding the parasite (Fig. 60.41). Unsheathed microfilariae include *Onchocerca volvulus*, *Mansonella ozzardi*, *M. perstans*, and *M. streptocerca*. The sheathed microfilariae include *Wuchereria bancrofti*, *Brugia malayi*, and *Loa loa*. Definitive identification to species is based on the presence and number of nuclei seen in the tail of the microfilaria. Alternatively, these parasites can be divided as causes of cutaneous, lymphatic, or body cavity infection. Species identification of blood microfilariae is particularly important because some may cause serious disease although others rarely do.

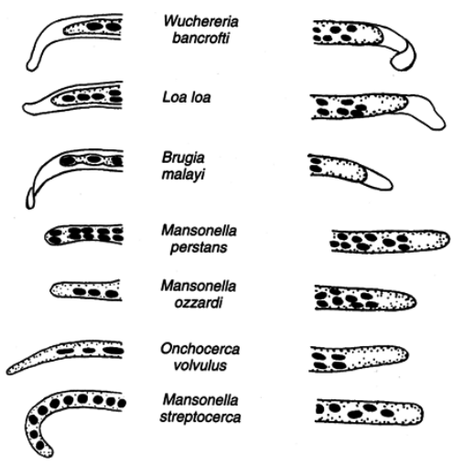


FIGURE 60.41. Differentiation of microfilariae, based on posterior ends **(left)** and anterior ends **(right)** of the microfilariae. Note the distribution of nuclei, their presence or absence in the extreme caudal portion, and the presence or absence of a sheath. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

Onchocerca volvulus

Onchocerciasis, a major cause of blindness worldwide, is caused by the skin filaria *O. volvulus*. Humans become infected by the

bite of black flies of the genus *Simulium*. The filariform larvae of the parasite are injected by the bite of the fly into the skin and migrate to subcutaneous connective tissue, where they develop into adult worms. The 20- to 30-mm adults become encased in fibrous subcutaneous nodules within which they may remain viable for as long as 15 years. The female worm gives birth to as many as 2,000 microfilariae each day. These larval forms exit the capsule to migrate for as long as 2 years in the subcutaneous tissues, skin, and eye. The life cycle is completed by ingestion of the microfilariae by new black flies in which the microfilariae transform into infectious filariform larvae.

Clinical onchocerciasis is characterized by infection involving the skin, subcutaneous tissue, lymph nodes, and eyes. Patients present with pruritus, subcutaneous nodules, or visual impairment. The clinical manifestations of the infection are owing to the acute and chronic inflammatory reaction to antigens released by the microfilariae as they migrate through the tissues. Within the skin, this inflammatory process results in loss of elasticity and areas of depigmentation, thickening, and atrophy. Subcutaneous inflammation may produce fibrosing, obstructive lymphadenitis, which may result in elephantiasis. Firm, nontender, mobile nodules may develop, especially over bony prominences.

Onchocerciasis affects more than 50 million people in tropical Africa and Central and South America and causes blindness in approximately 5% of infected individuals. The mechanism for development of eye disease is thought to be a combination of both direct invasion by the microfilariae and antigen-antibody complex deposition within the ocular tissues (81). Patients progress from conjunctivitis with photophobia to punctate and sclerosing keratitis. Internal eye disease with anterior uveitis, chorioretinitis, and optic neuritis may also occur. The disease is one of the leading causes of blindness worldwide. In some communities in West Africa, as many as half of the adult male population is blind owing to this infection.

Diagnosis of onchocerciasis is made by demonstration of microfilariae in skin snip preparations from the infrascapular or gluteal region. A sample is obtained by raising the skin with a needle and shaving the epidermal layer with a razor. The specimen is incubated in saline for several hours, then inspected with a dissecting microscope for the presence of nonsheathed microfilariae (Fig. 60.42). In patients with ocular disease, the organism may also be seen in the anterior chamber with the aid of a slit lamp. Serologic and culture methods are not helpful, although efforts to develop serologic detection methods are ongoing.

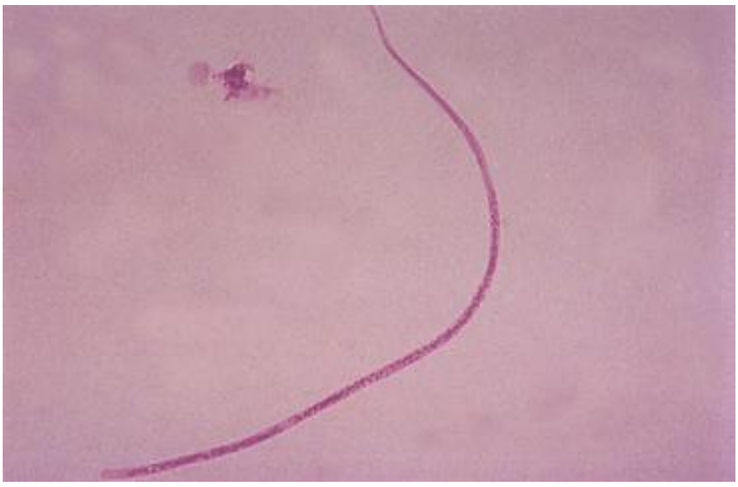


FIGURE 60.42. *Onchocerca volvulus*: Microfilariae teased from skin snip (hematoxylin and eosin stain). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

Treatment includes the use of ivermectin to eradicate the microfilariae and either surgical excision of nodules or administration of suramin for the adult worms. A single dose of ivermectin (150 µg/kg) greatly reduces the number of microfilariae in the skin and eyes, thus diminishing the likelihood of developing disabling onchocerciasis. Suppression of dermal microfilariae with ivermectin reduces the transmission of this vector-borne disease and may prove to be a successful preventive strategy (82).

Mansonella

Filarial infections caused by *Mansonella* species (*M. perstans*, *M. ozzardi*, and *M. streptocerca*) are generally asymptomatic but may cause dermatitis, lymphadenitis, hydrocele, and rarely lymphatic obstruction resulting in elephantiasis (76). All the *Mansonella* species produce nonsheathed microfilariae (Fig. 60.41) in blood and subcutaneous tissues, and all are transmitted by biting midges (*Culicoides*) or black flies (*Simulium*).

These filariae are found in South America and Central Africa. All are treatable with diethylcarbamazine. Species identification, if desired, can be accomplished with blood smears noting the structure of the microfilariae. Serologic tests are also available.

Wuchereria and Brugia

Lymphatic filariasis constitutes a group of diseases characterized by lymphatic obstruction that results in a condition known as

elephantiasis (76). The two species of filariae most commonly involved are *W. bancrofti* and *B. malayi*. The adult forms of these filariae parasitize the lymphatic vessels, primarily in the arms, legs, or groin, where they may exist for as long as 10 years. The female worms produce sheathed microfilariae (Fig. 60.41) that enter the bloodstream of the host and circulate with a nocturnal periodicity. The microfilariae may be found in greatest numbers in the peripheral circulation between 9:00 p.m. and 2:00 a.m., a time period that corresponds to the peak feeding time of the mosquito vectors (*Culex*, *Aedes*, *Anopheles*, or *Mansonia*). The microfilariae that are ingested by the mosquito vector transform first into rhabditiform and then into filariform larvae within the mosquito. The infective filariform larvae actively penetrate the feeding site when the mosquito takes its next meal and migrate to the lymphatic vessels, where they reach adulthood in 6 to 12 months.

Clinical manifestations of infection with *W. bancrofti* and *B. malayi* vary from asymptomatic infection with positive serology but no detectable microfilariae to lymphatic inflammation and chronic lymphatic obstruction with development of lymphoceles, lymphangitis, and elephantiasis. Acute symptoms may include fever, headache, and back pain in association with lymphangitis. The acute presentation is thought to be caused by the inflammatory response to the presence of molting adolescent worms and dead or dying adults within the lymphatic vessels. The acute response is followed by the formation of granulomas, fibrosis, and permanent lymphatic obstruction. With repeated exposure over many years, individuals develop heavy infections and massive lymphatic blockade. The resultant lymphedema and fibrosis may be complicated by recurrent bacterial infections, which contribute to the tissue damage. Continuation of this process over a period of decades results in elephantiasis, usually most severe in the lower extremities and genitalia (76). Occasionally ascites and pleural effusions secondary to rupture of enlarged lymphatics into the peritoneal or pleural cavities may be observed.

It is estimated that lymphatic filariasis currently infects between 80 and 100 million individuals in Africa, Latin America, the Pacific Islands, and Asia. *W. bancrofti* is distributed throughout these areas, whereas *B. malayi* is limited to Southeast Asia and Indonesia.

Eosinophilia is usually present during acute inflammatory episodes; however, demonstration of microfilariae in blood is required for definitive diagnosis. Microfilariae can be demonstrated in Giemsa-stained blood films as for malaria (Fig. 60.43 and Fig. 60.44). If the parasitemia is scant, the blood specimen may be concentrated before it is examined (76). Given the nocturnal periodicity of the microfilaremia, it is recommended that blood specimens be obtained between 9:00 p.m. and 2:00 a.m.

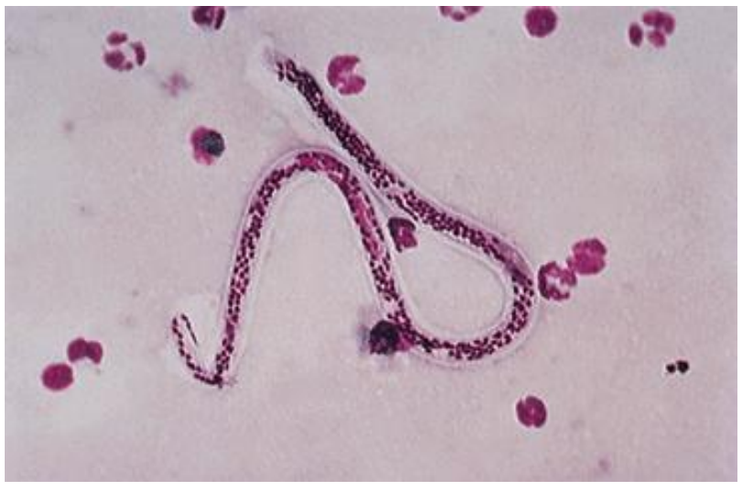


FIGURE 60.43. *Wuchereria bancrofti*: Microfilariae in thick film (Giemsa stain). (From Smith JW, et al. *Diagnostic medical parasitology: blood and tissue parasites*. Chicago: American Society of Clinical Pathologists, 1976.)

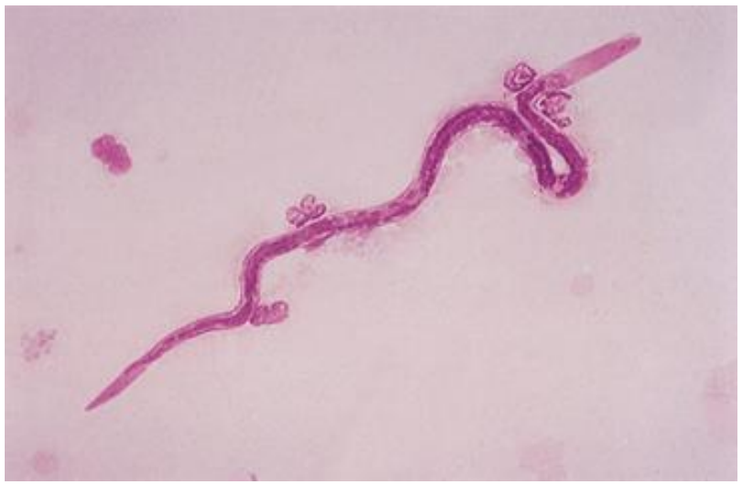


FIGURE 60.44. *Brugia malayi*: Microfilariae in thin film (Giemsa stain). (From Smith JW, et al. Diagnostic medical parasitology: blood and tissue parasites. Chicago: American Society of Clinical Pathologists, 1976.)

Both *W. bancrofti* and *B. malayi* produce sheathed microfilariae. Further identification is based on the study of head and tail structures (Fig. 60.41). Clinically, an exact species identification is not critical, as the treatment for all filarial infections is identical (diethylcarbamazine), with the possible exception of *O. volvulus* (ivermectin plus suramin).

Serologic testing is available but may lack sensitivity and specificity. Detection of circulating filarial antigens is promising but not widely available as a diagnostic test.

Treatment is of little benefit in most cases of lymphatic filariasis. Diethylcarbamazine has been used but has little apparent effect on adult worms. Ivermectin appears promising, although controlled studies have not yet been performed. Supportive and surgical therapy for lymphatic obstruction may be of some cosmetic help. Control programs combine mosquito control with mass treatment of the entire population.

Loa loa

Loiasis is a filarial disease found primarily in the rain forests of Africa, where *L. loa* is transmitted by biting flies of the genus *Chrysops*. The adult worms can persist in the human host for 17 years or more and migrate continuously through the subcutaneous tissues. During migration, the parasites elicit an intense eosinophilia and produce localized areas of allergic subcutaneous inflammation known as Calabar swellings. These swellings reach 10 to 20 cm in diameter, last 2 to 3 days, and may be accompanied by fever, pruritus, and pain. Occasionally, the adult worm may migrate across the eye subconjunctivally, producing tearing, pain, edema, and impaired vision.

The adult female produces sheathed microfilariae, which may be detected in the bloodstream during day time hours. The

sheath of the microfilariae does not stain with Giemsa, and the tail has nuclei extending to the rounded tip (Fig. 60.41).

The diagnosis of loiasis may be made by recovering the adult worm from the eye or by detecting microfilariae in the blood or aspirates of Calabar swellings. Eosinophilia is constant and may constitute 50% to 70% of the peripheral white count. Diethylcarbamazine is effective against both adults and microfilariae; however, destruction of the parasites may induce severe allergic reactions, which require treatment with corticosteroids. The role of ivermectin remains undefined for this infection.

Dirofilaria immitis

The canine heartworm *D. immitis* is endemic in the southeastern United States. It is transmitted by mosquitoes and has a typical filarial life cycle, with the dog serving as the definitive host. Accidental infections in humans may be acquired by the bite of the mosquito vector, resulting in one or more worms becoming lodged in the right heart or pulmonary artery. Because humans are unsuitable hosts for *Dirofilaria*, the parasite dies before maturing and is embolized to smaller branches of the pulmonary vasculature, where it may cause a small area of infarction that heals and may appear as a coin lesion on a chest radiograph. Throughout this process, most patients remain asymptomatic. The pulmonary coin lesion may pose a problem because it resembles a malignancy requiring surgical removal. A definitive diagnosis is made when a thoracotomy specimen is examined microscopically, revealing the typical cross-sections of the parasite (76, 83).

Unfortunately, no laboratory test currently available can provide an accurate diagnosis of dirofilariasis (76). Peripheral eosinophilia is rare, and the radiographic features are insufficient to allow one to distinguish pulmonary dirofilariasis from bronchogenic carcinoma. Serologic tests are not sufficiently sensitive or specific to be clinically useful (76).

Dracunculus medinensis

Dracunculiasis is a chronic subcutaneous infection caused by the tissue-invading nematode *D. medinensis*. *D. medinensis* has a simple life cycle depending on fresh water and a copepod of the genus *Cyclops*, which serves as the intermediate host. The infection in humans is initiated when a copepod infected with the larval stage of *D. medinensis* is ingested in drinking water. After ingestion, the larvae are released from the copepods in the stomach and small intestine, penetrate the wall of the digestive tract, and migrate to the retroperitoneal space, where they mature. Male and female worms mate in this location. The gravid female then migrates to the subcutaneous tissues, usually in the extremities, where a vesicle is formed in the host tissue. The vesicle ulcerates and the worm protrudes a loop of uterus through the ulcer. On contact with water, the larval worms are released. The larvae are then ingested by the *Cyclops* in fresh water, where they may in turn infect another mammalian host. When the female has discharged all the larvae, it may retreat into deeper tissue, where it is gradually absorbed, or it may simply be expelled from the site.

Dracunculiasis is generally asymptomatic until the gravid female creates the vesicle and the ulcer in the skin for the liberation of larvae. The time between initial exposure and ulceration is usually approximately 1 year. The site of the ulceration is painful and erythematous, and the adult worm may be visible in the center of the lesion. At the time of ulceration, the patient may develop local or generalized pruritus, nausea, diarrhea, and shortness of breath. Fluid present at the site of ulceration may contain many visible larvae. Patients with multiple adult worms may have several ulcerations occur simultaneously. Occasionally, the site may become secondarily infected with bacteria, leading to abscess formation and further tissue destruction.

D. medinensis occurs in many parts of Asia and equatorial Africa, infecting an estimated 10 million individuals. Reservoir hosts include dogs and other mammals. Humans perpetuate the cycle when infected individuals bathe or stand in the wells or ponds from which drinking water is obtained. Larvae are discharged from the lesions on the arms, legs, feet, and ankles to infect the copepods in the water, which are in turn collected with the drinking water.

Diagnosis of this infection is usually straightforward clinically, with the adult worm present in the ulcerated skin lesion. Larvae can also be detected under the microscope in samples of fluid from the ulcer site (76).

Treatment may be accomplished by administration of niridazole, metronidazole, or thiabendazole. Alternatively, the worm may be removed surgically. The ancient method of slowly wrapping the worm on a twig, thus extracting it from the ulcer, is still used in many endemic areas. This process may be complicated by infection and/or toxic reactions if the worm is broken during removal.

Prevention requires education concerning the life cycle of the parasite and avoidance of water contaminated with *Cyclops*. Ideally, adequate treatment of water supplies would result in elimination of this infection (84).

Trematodes

The trematodes or flukes are members of the Platyhelminthes and are generally flat, fleshy, leaf-shaped worms. They vary in length from a few millimeters to several centimeters and possess two muscular suckers: one oral, surrounding the opening to the primitive digestive tract, and one ventral for attachment. The digestive system consists of a muscular pharynx and esophagus and bilateral ceca that end blindly near the posterior aspect of the worm.

The flukes may be divided into two major categories based on their reproductive systems (75, 76). Most flukes are hermaphrodites. The adult hermaphrodite contains both male and female gonads and produces operculate eggs (Fig. 60.45). The schistosomes constitute the second major category and include organisms with separate sexes. The female schistosome deposits only nonoperculated eggs. Schistosomes also have more rounded bodies than do the hermaphroditic flukes. Both schistosomes and hermaphroditic flukes have similar life cycles that include one or more intermediate hosts. Mollusks (snails or clams) are essential, first intermediate hosts for all trematodes. The mollusks are infected by ciliated larvae or miracidia, which are released from eggs that have been deposited into fresh water. Within the first intermediate host, the miracidia undergo asexual reproduction to produce thousands of motile cercariae, which are released into the water where they actively swim about in search of their next host. In the case of schistosomes, the cercariae directly penetrate the skin of humans and produce infection.

The cercariae of the hermaphroditic flukes encyst in or on an aquatic plant or animal (crustacean or fish), where they develop into infective metacercariae. Their cycle is completed when the second intermediate host is ingested by a human.

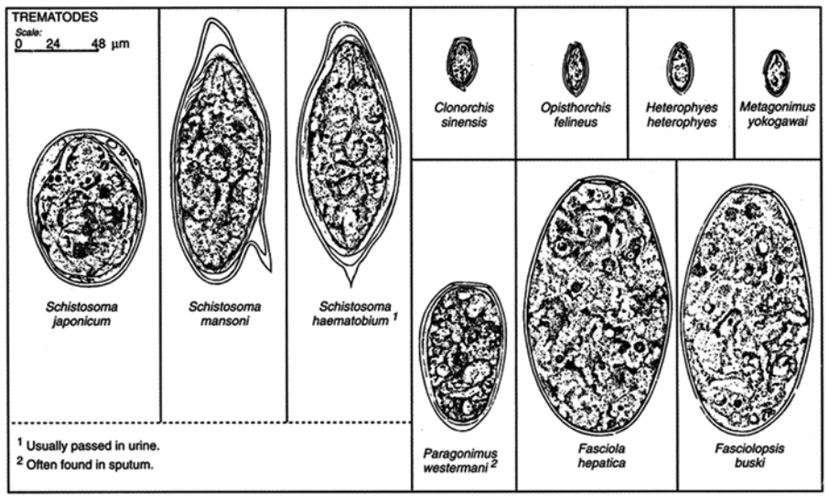


FIGURE 60.45. Trematode eggs found in human feces. (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Another means of classifying the trematodes that cause human infection is by the anatomic location of the parasite in the human host (75, 76). Thus, we have the intestinal trematodes (Table 60.9), the liver and lung trematodes (Table 60.10), and the blood

TABLE 60.9. MORPHOLOGY OF INTESTINAL TREMATODES

	Fasciolopsis buski	*Heterophyes heterophyes*	*Metagonimus yokogawai*
Method of infection	Ingestion of water chestnut, bamboo shoots	Ingestion of freshwater fish	Ingestion of freshwater fish
Location in host	Attached to wall of the small intestine	Crypts and lumen of the small intestine	Crypts and lumen of the small intestine
Reservoir host	Dogs, pigs, rabbits	Fish-eating mammals	Fish-eating mammals
Intermediate host	*Segmentia* or *Hippeutis* snails	*Pironella* and *Cerithidea* snails	*Semisulcospira* snails
Adult worm	Large, up to 75 mm long × 20 mm wide	Small, pyriform-shaped, 1–2.5 mm long × 0.3–0.7 mm wide; single ovary is situated anterior to the posterior testes	Small, pyriform-shaped, 1–2.5 mm long × 0.3–0.7 mm wide; single ovary is situated anterior to the posterior testes
Egg morphology	Ellipsoidal, unembryonated, operculated, 130–150 × 63–90 µm	Ovoid, operculated, 20–30 × 15–17 µm, contain a miracidium when discharged	Ovoid, operculated, 20–30 × 15–17 µm, contain a miracidium when discharged
Clinical manifestations	Attachment of worm may cause hypersecretion of mucus, hemorrhage, ulcerations, abscess, bowel obstruction, acute ileus, eosinophilia, leukocytosis, malabsorption	Abdominal pain, mucous diarrhea, ulceration of intestinal wall, may invoke pathologic lesions in the brain and heart	Abdominal pain, mucous diarrhea, ulceration of intestinal wall; may invoke pathologic lesions in the brain and heart
Diagnostic methods	O&P exam	O&P exam	O&P exam
Specimen	Feces	Feces	Feces
Common name	Large intestinal fluke	Heterophid fluke	Heterophid fluke

O&P exam, ova and parasites examination.

trematodes (Table 60.11). Of the many trematodes infecting humans, only seven representative agents will be discussed: the blood flukes, all of which are members of the genus *Schistosoma* (*S. mansoni*, *S. haematobium*, and *S. japonicum*), the liver flukes (*Fasciola hepatica* and *Opisthorchis sinensis*), the lung fluke (*Paragonimus westermani*), and the intestinal fluke (*Fasciolopsis buski*). The other flukes causing infection in humans are similar to these representative agents in terms of epidemiology, clinical syndromes, and therapy (75, 76). Basic details of additional intestinal and tissue flukes are provided in Table 60.9 and Table 60.10.

TABLE 60.10. MORPHOLOGY OF LIVER AND LUNG TREMATODES

	Opisthorchis sinensis	*Opisthorchis viverrini*	*Fasciola hepatica*	*Paragonimus westermani*
Method of infection	Ingestion of raw fish	Ingestion of raw fish	Ingestion of raw water plants	Ingestion of crabs, crawfish
Location in host	Bile ducts of the liver	Bile ducts of the liver	Bile ducts of the liver	Encapsulated in parenchyma of the lung
Reservoir host	Dogs, cats, fish-eating mammals	Dogs, cats, fish-eating mammals	Herbivores	Dogs, cats, tigers, lions
Intermediate host	1st: *Parafossarulus, Bulinus, Semisulcospira, Alocinma, Melanoides* snails	1st: *Bithynia* snail; 2nd: freshwater fish	1st: *Parafossarulus, Bulinus, Semisulcospira, Alocinma, Melanoides* snails	1st: *Semisulcospira* and *Brotia* snails; 2nd: Crabs, crawfish
Adult worm	Flattened, spatulate, 10–25 mm long × 3–5 mm wide, hermaphroditic with single round ovary anterior to testes	10–25 mm long, 2 testes lying one behind the other in posterior end, ovary anterior to testes	Large, fleshy, up to 30 mm long × 14 mm wide; anterior end is cone shaped, all internal organs are branched	Ovoid, 7.5–12 × 4–6 mm, hermaphroditic, lobed ovary located anterior to testes, testes lie side by side in posterior part of body
Egg morphology	Ovoid, thick-shelled, seated operculum; 27–35 × 12–19 µm; knob at abopercular end; prominent shoulders, contains miracidium	19–29 × 12–17 µm, operculated; prominent shoulders; contains miracidium	Large, ellipsoidal, unembryonated; operculated 130–150 × 63–90 µm; operculum is small	Ovoid, large, thick-shelled; unembryonated when passed in sputum or feces; 80–120 × 45–70 µm; thickened at abopercular end
Clinical manifestations	Biliary tract obstruction, bile retention, acute pancreatitis, cholecystitis, cholelithiasis, fever, chills, diarrhea, epigastric pain, enlarged tender liver, jaundice	Infection confined to biliary tract system, weakness, pain in right upper quadrant, elevated serum IgE levels (3 or 4 times)	Fever, right upper quadrant pain, eosinophilia associated with migratory phase, biliary obstruction, cholangitis, acute epigastric pain, fever, jaundice, enlarged liver, eosinophilia	Cough with increased production of blood tinged sputum, chest pain, dyspnea with chronic bronchitis, fever, headache, nausea, vomiting, visual disturbances, motor weakness, localized or general paralysis
Diagnostic methods	CF, HAI, O&P exam	O&P exam	CF, HAI IIF, immunodiffusion, immunoelectrophoresis, countercurrent electrophoresis, O&P exam	CF, O&P exam
Specimen	Feces	Feces	Feces	Feces
Common name	Chinese liver fluke		Sheep liver fluke	Oriental lung fluke

IgE, Immunoglobulin E; CF, complement fixation; HAI, hemagglutination inhibition; IIF, indirect immunofluorescence; O&P exam, ova and parasite examination.

Intestinal Trematodes

Several intestinal trematodes are recognized, including *F. buski*, *Heterophyes heterophyes*, and *Metagonimus yokogawai* (Table 60.9). *F. buski*, the largest, most prevalent, and important intestinal fluke, is a large (75 mm long by 20 mm wide) fleshy worm (Fig. 60.46) that is a common parasite of pigs and humans in China and Southeast Asia. An estimated 10 million individuals are infected with this intestinal parasite.

The life cycle of *F. buski* is typical of hermaphroditic trematodes. Metacercariae are encysted on aquatic plants, which serve as the second intermediate host. Ingestion of the metacercariae along with the freshwater vegetation results in establishment of the infection in the small intestine. The flukes live attached to the mucosa of the duodenum and jejunum by means of the ventral sucker. Self-fertilization occurs and egg production begins 3 months after the initial infection with the metacercariae. The eggs, which are broadly ellipsoidal (130 to 150 µm × 60 to 90 µm) and unembryonated with a small, indistinct operculum

TABLE 60.11. MORPHOLOGY BLOOD TREMATODES

	Schistosoma mansoni	*Schistosoma japonicum*	*Schistosoma haematobium*
Method of infection	Cercaria penetrates skin of humans, loses tail, becomes a schistosomulum, and migrates to blood vessel where it develops to an adult	Cercaria penetrates skin of humans, loses tail, becomes a schistosomulum, and migrates to blood vessel where it develops to an adult	Cercaria penetrates skin of humans, loses tail, becomes a schistosomulum, and migrates to blood vessel where it develops to an adult
Location in host	Venous plexus of colon and lower ileum; portal system of liver	Venous plexus of small intestine	Plexus of bladder, rectum
Reservoir host	Humans, nonhuman primates	Dogs, cats, cattle, water buffalo, pigs	Humans
Intermediate host	*Biomphalaria* snail	*Oncomelania* snail	*Bulinus* snail
Adult worm	Sexes are separate; males, 6.4–12 mm long, 6–9 testes in anterior half of body; females, 7.2–17 mm long	Sexes are separate; sizes are comparable; contains more eggs *in utero* than *S. mansoni* or *S. haematobium*	Sexes are separate, males, 10–15 mm long; females, 20 mm long; tegument has minute tuberculations; female may contain 20–100 eggs *in uterus*
Egg morphology	114–175 × 45–70 µm, nonoperculated, transparent shell, prominent lateral spine, contain miracidium when passed in feces	Round, large, nonoperculated, 70–100 × 55–65 µm, transparent shell, small inconspicuous spine	Large, nonoperculated, transparent shell, prominent terminal spine, 112–170 × 40–70 µm
Clinical manifestations	Cercarial dermatitis, high fever, hepatosplenomegaly, lymphadenopathy, eosinophilia, dysentery, abdominal pain, liver tenderness, urticaria, general malaise, blood and mucus in stools	Cercarial dermatitis, high fever, hepatosplenomegaly, lymphadenopathy, eosinophilia, dysentery, abdominal pain, liver tenderness, urticaria, general malaise, blood and mucus in stools	Cercarial dermatitis, high fever, hepatosplenomegaly, lymphadenopathy, eosinophilia, dysentery, abdominal pain, liver tenderness, urticaria, general malaise, hematuria, dysuria
Diagnostic methods[a]	CF, HAI, IIF, immunoelectrophoresis, ELISA, O&P exam	CF, HAI, IIF, immunoelectrophoresis, ELISA, O&P exam	CF, HAI, IIF, immunoelectrophoresis, ELISA, O&P exam
Specimen	Stool, rectal biopsy, serology	Stool, rectal biopsy, serology	Urine, serology
Common name	Manson's blood fluke	Blood fluke	Bladder fluke

CF, complement fixation; HAI, hemagglutination inhibition; IIF, indirect immunofluorescence; ELISA, enzyme-linked immunosorbent assay; O&P exam, ova and parasite examination.

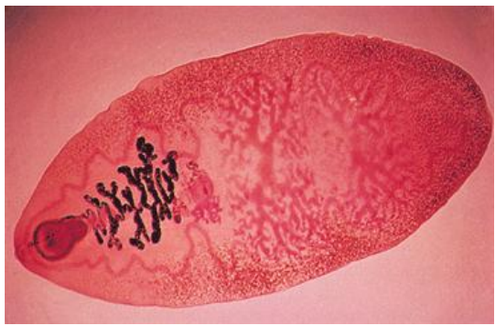


FIGURE 60.46. *Fasciolopsis buski*, adult worm (carmine stain, whole mount). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

(Fig. 60.45 and Fig. 60.47), are passed in the feces and gain access to fresh water, where they require a prolonged maturation period. After maturation, the eggs release a ciliated, free-swimming larval stage (miracidium) that seeks a particular species of snail for further development.

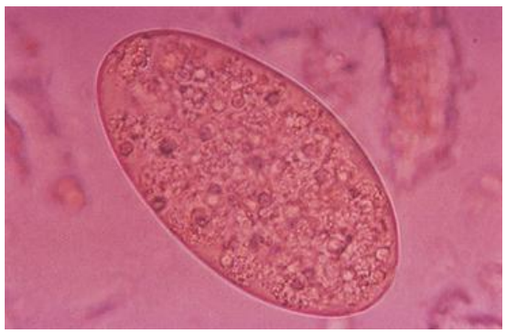


FIGURE 60.47. *Fasciolopsis buski*, egg (high power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Symptoms of *F. buski* infection are directly related to the worm burden in the small intestine. Ulceration of the superficial mucosa may occur, accompanied by inflammation and hemorrhage. Diarrhea and epigastric pain similar to that of a duodenal ulcer are the main presenting symptoms. Anorexia, nausea, and vomiting can occur, accompanied by eosinophilia. A malabsorption syndrome similar to giardiasis is common, and intestinal obstruction may occur if there are enough worms.

The diagnosis is made by finding the large, oval, bile-stained, operculate eggs in the stool (Fig. 60.47). The measurements and appearance of *F. buski* eggs are similar to those of the liver fluke *F. hepatica*, and differentiation of the eggs of these species alone is not usually possible (75). The large adult flukes can rarely be found in feces or specimens collected at surgery.

The treatment of choice is praziquantel. Alternatively,

niclosamide may be used. Prompt treatment of infected humans may minimize the spread of infection. Prevention may also include proper sanitation and elimination of the reservoir hosts through the use of molluscacides.

Liver Trematodes

There are several liver flukes in various geographic regions that parasitize the biliary passages of humans (76). These include *F. hepatica*, *Opisthorchis sinensis*, *O. viverrini*, *O. felineus*, and *Dicrocoelium dendriticum* (Table 60.10). The most important of these, *F. hepatica* and *O. sinensis*, are discussed in this chapter.

Fasciola hepatica

The largest liver fluke is *F. hepatica*, a common parasite of sheep and cattle with a worldwide distribution.

The large (30 × 13 mm), fleshy adult flukes live in the biliary tree, where they lay eggs that are passed in the feces. The eggs develop in water and follow a life cycle similar to that of *F. buski*, with human infection resulting from ingestion of fresh water vegetation that harbors the encysted metacercariae. When ingested, the infective larvae are released from their cysts and migrate through the duodenal wall, across the peritoneal cavity, and into the bile ducts (via the liver), where they become adult worms. As with *F. buski*, self-fertilization occurs and egg production begins approximately 3 to 4 months after the initial infection. The large (130 to 150 µm × 60 to 90 µm), operculated eggs (Fig. 60.45) are excreted in the stool.

Early clinical signs and symptoms are owing to the migration of the larvae through the liver. Right upper quadrant pain and tenderness, hepatomegaly, chills and fever, with marked eosinophilia may be observed. The presence of the adult worms in the biliary tree elicits an inflammatory response with resultant hyperplasia of the epithelium and eventual fibrosis. Biliary obstruction caused by bile duct fibrosis and stone formation may occur. Penetration of adult worms into the liver parenchyma occurs rarely but may result in necrotic foci and abscess formation known as “liver rot.”

The diagnosis is established by detecting the characteristic eggs in stools. The unembryonated, operculate eggs cannot be distinguished from those of *F. buski* (Fig. 60.45 and Fig. 60.47). Distinction between the two flukes is necessary for therapeutic purposes because *F. buski* responds favorably to praziquantel and *F. hepatica* does not. The two species may be distinguished by sampling the patient's bile to detect the eggs of *F. hepatica*. Care must be taken to differentiate true infection from the spurious presence of eggs resulting from the ingestion of infected liver. Several stool samples should be examined to rule out such spurious observations and ensure that there is a true infection.

F. hepatica responds poorly to praziquantel. Treatment with bithionol or triclabendazole has been effective. Control measures are similar to those for *F. buski*.

Opisthorchis sinensis

The Chinese liver fluke *O. sinensis* is the most common liver fluke that infects humans. *O. sinensis* infects approximately 19 million individuals in China, Japan, Korea, and Vietnam (76, 85). It is also seen among residents of the United States and can be traced to the consumption of raw, pickled, smoked, or dried freshwater fish harboring viable metacercariae. Cats and dogs serve as important reservoir hosts.

The adult fluke is small and delicate, measuring 10 to 25 mm long × 3 to 5 mm wide (Fig. 60.48). *O. sinensis* may survive in the biliary tract for as long as 50 years producing approximately 2,000 eggs per day. The tiny (15 × 30 µm), urn-shaped eggs enter the duodenum with the bile and are discharged in the feces. On reaching fresh water, the eggs are ingested by the molluscan first intermediate host, transformed into cercariae, and released to penetrate the tissues of freshwater fish, in which they encyst to become infective metacercariae. Infection is acquired by ingestion of uncooked fish. The ingested larvae excyst in the duodenum and migrate via the ampulla of Vater and the common bile duct into the bile ducts of the liver, where they mature and live.

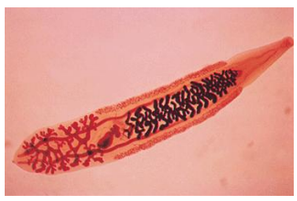


FIGURE 60.48. *Opisthorchis sinensis*, adult worm (carmine stain, whole mount). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Infection in humans is usually mild and asymptomatic. Severe infections involving large numbers of flukes may produce fever, diarrhea, epigastric pain, hepatomegaly, anorexia, and occasionally jaundice. Biliary obstruction may occur, and individuals with severe longstanding infection may develop bile duct carcinoma. Invasion of the gallbladder or pancreatic ducts may occur, causing cholecystitis and pancreatitis.

The diagnosis is made by detecting the distinctive eggs in stools. The eggs measure 27 to 35 µm × 12 to 19 µm and are characterized by a distinct operculum with prominent shoulders and a tiny knob at the posterior (abopercular) pole (Fig. 60.49). In mild infections, repeated examinations of stool or duodenal aspirates may be necessary. In acute symptomatic infection, there is usually an eosinophilia and elevation of serum alkaline phosphates levels. Radiographic imaging procedures may detect abnormalities of the biliary tract.

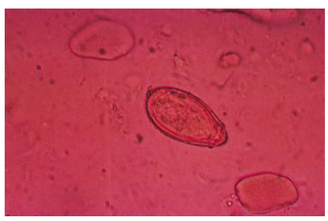


FIGURE 60.49. *Opisthorchis sinensis*, egg (oil immersion). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Treatment is with praziquantel. Prevention requires avoidance of improperly cooked fish and the implementation of proper sanitation policies.

Lung Trematodes

The principal lung fluke of humans is *Paragonimus westermani*, which is found in the Orient; however, several additional species of *Paragonimus* parasitizing humans and other animals in Africa and South, Central, and North America have been described. *P. westermani* is a small (10 mm × 5 mm), reddish-brown worm that is a common parasite of shore-feeding animals (pigs, wild

boars, monkeys) as well as humans in China, Japan, and Southeast Asia. Its prevalence is directly related to the consumption of uncooked freshwater crabs and crayfish. It is estimated that approximately 3 million individuals are infected with this lung fluke. As many as 1% of all Indochinese immigrants to the United States are infected with *P. westermani* (76, 86).

The adult worms are usually found encysted in the pulmonary parenchyma of their definitive host. Infection is acquired by the ingestion of raw or partially cooked crabs or crayfish containing encysted metacercariae in their tissues. When ingested, the metacercariae excyst in the duodenum and migrate through the intestinal wall into the peritoneal cavity, through the diaphragm, and finally reach maturity in the lungs 5 to 6 weeks later. Occasionally organisms undergo aberrant migration and are retained in the intestinal wall and mesentery or migrate to other ectopic foci such as the liver, pancreas, skeletal muscle, subcutaneous tissue, or the central nervous system.

On reaching the pulmonary parenchyma, the flukes become encapsulated by the host's fibrous reaction. The adult worms deposit operculate, golden-brown eggs within the fibrous capsule. The capsule ultimately erodes into a bronchiole and discharges the eggs into the respiratory tree. The eggs may then be coughed up and spat out or swallowed and passed in the stool.

The clinical manifestations of paragonimiasis may be attributable to larvae migrating through tissues or to adults established in the lungs or other ectopic sites. The onset of the disease coincides with larval migration and is associated with fever, chills, and high eosinophilia. When established in the lungs, the adult worms stimulate an inflammatory reaction that results in fever and a chronic cough with increased sputum production and occasional hemoptysis. Evacuation of the capsule or cyst containing the adult worms and eggs produces a cavity, which may become secondarily infected with bacteria. The expectorated sputum is blood-tinged and contains numerous golden-brown eggs. Severe infections produce a chronic pneumonia with abscess formation and pulmonary fibrosis. Dyspnea, severe chest pain, and pleural effusion may be seen. Location of larvae, adults, and eggs in ectopic sites may produce severe clinical symptoms depending on the site involved. Cerebral paragonimiasis occurs in approximately 1% of Oriental cases and may produce a severe neurologic disease including seizures, paralysis, and visual disturbances.

The laboratory diagnosis of paragonimiasis depends on the detection of the golden-brown, operculated eggs in sputum and feces (Fig. 60.50). Pleural effusions, when present, should also be examined for eggs. Marked eosinophilia is common, and chest radiographs often show nodular shadows, calcifications, or patchy infiltrates. Serologic procedures are available through reference laboratories and may be useful, particularly in cases with extrapulmonary (e.g., central nervous system) involvement.

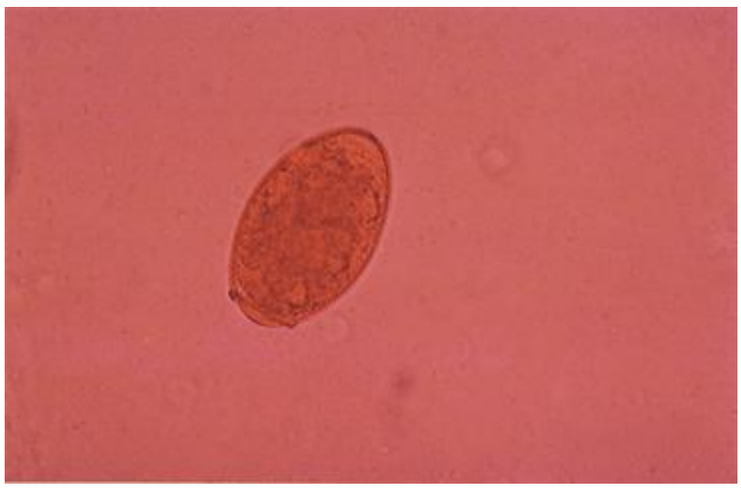


FIGURE 60.50. *Paragonimus westermani*, egg (high power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

The disease responds adequately to treatment with either praziquantel or bithionol. Control requires adequate cooking of freshwater crustaceans before ingestion.

Blood Trematodes (Schistosomiasis)

Schistosomiasis is a major parasitic infection of tropical areas with some 200 million infections worldwide (76). The three schistosomes most frequently associated with human disease are *S. mansoni*, *S. japonicum*, and *S. haematobium*. As discussed previously, schistosomes differ from the hermaphroditic flukes in body structure and in having separate sexes, nonoperculated eggs, and a life cycle that includes only one intermediate host. They are also obligate intravascular parasites inhabiting the venules of the intestine (*S. mansoni* and *S. japonicum*) or the bladder (*S. haematobium*) rather than the cavities, ducts, and other tissues parasitized by the hermaphroditic flukes.

Infection with schistosomes is acquired by exposure to fresh water containing ciliated, free-swimming cercariae. The cercariae penetrate intact skin, causing intense pruritus, enter the circulation, and develop in the intrahepatic portal circulation. Adult male and female worms pair up early in development. The flat 1- to 2-cm male has a longitudinal (gynecophoral) canal resulting from folding the lateral aspects of the body toward the center. The long, slender, cylindrical female resides in the canal, and the two worms stay together in this fashion for the rest of their lives. After mating in the portal vein, the paired adult worms migrate to their final location in the small venules of the mesentery (*S. mansoni* and *S. japonicum*) or venous plexuses of the bladder (*S. haematobium*). In general, *S. japonicum* resides in the venous radicals of the small intestine, *S. mansoni* resides in the vicinity of the descending colon and rectum, and *S. haematobium* resides in the veins of the bladder and other pelvic organs. Once in the submucosal venules, the worms initiate oviposition, which may continue at the rate of 300

to 3,000 eggs daily for 4 to 35 years. Although the host inflammatory response to the adult worms, which are coated with host proteins, is minimal, the eggs elicit an intense inflammatory reaction with both mononuclear and polymorphonuclear cellular infiltrates and the formation of microabscesses. In addition, the larvae inside the eggs produce enzymes that aid in tissue destruction and allow the eggs to pass through the mucosa and into the lumen of the bladder and bowel, where they are passed to the external environment in the urine and feces, respectively.

The eggs are fully embryonated when they are passed to the environment and hatch quickly on reaching fresh water to release motile miracidia. The miracidia then invade the appropriate snail host, where they develop into thousands of infectious cercariae. The free-swimming cercariae are released into the water, where they are immediately infectious for humans and other mammals.

The infection is similar in all three species of human schistosomes in that disease results primarily from the host immune response to the eggs rather than from the adult worms. As noted above, the very earliest signs and symptoms of infection are owing to the penetration of the cercariae through the skin. Immediate and delayed hypersensitivity to parasite antigens results in an intensely pruritic papular skin rash.

The onset of oviposition, 1 to 2 months after primary exposure, results in the Katayama syndrome, characterized by fever, chills, cough, urticaria, arthralgias, lymphadenopathy, splenomegaly, and abdominal pain. Laboratory abnormalities include leukocytosis, eosinophilia, and a polyclonal gammopathy with elevated levels of IgG, IgM, and IgE. This symptom complex may persist for 3 months or more and is thought to be caused by the massive release of parasite antigens with subsequent immune complex formation.

The more chronic and significant phase of schistosomiasis is owing to the presence of the eggs in the various tissues and the resulting formation of granulomas and fibrosis (87). The retained eggs induce extensive inflammation and scarring, the clinical significance of which is directly related to the location and number of eggs. In *S. haematobium* infection, egg deposition in the walls of the bladder results in hematuria and dysuria and eventually may cause scarring and loss of bladder capacity, obstructive uropathy, and bladder carcinoma. Infections with *S. mansoni* and *S. japonicum* may produce hepatic and intestinal abnormalities. Deposition of eggs in the bowel mucosa results in inflammation and thickening of the bowel wall with associated abdominal pain, diarrhea, and blood in the stool. Importantly, eggs of *S. mansoni* and *S. japonicum* may be carried by the portal vein to the liver, where the resulting inflammation can lead to periportal fibrosis and eventually to portal hypertension with its associated manifestations. Although *S. mansoni* and *S. japonicum* eggs are primarily deposited in the intestine, eggs may appear in the brain, spinal cord, lungs, and other sites. Deposition of eggs in the central nervous system (brain and spinal cord) occurs more commonly with *S. japonicum* and may produce severe neurologic problems (87).

The geographic distribution of the various species of *Schistosoma* is dependent on the availability of a suitable snail host. *S. mansoni* is the most widespread of the blood flukes, occurring in Africa, the Middle East, South America, Puerto Rico, and several Caribbean islands. *S. haematobium* is largely confined to Africa and the Middle East where its distribution overlaps that of *S. mansoni*. *S. japonicum*, the Oriental blood fluke, is found only in Japan, China, the Philippines, and the Celebes. Schistosomiasis may be considered a disease of economic progress, as the development of massive land irrigation projects in desert and tropical areas has resulted in the dispersion of infected humans and snails to previously uninvolved areas.

The diagnosis of schistosomiasis is usually established by the demonstration of characteristic eggs in urine or feces. Concentration techniques may be necessary in light infections. Biopsy of bladder or rectal mucosa may be positive when repeated examination of urine and stool specimens are negative. In general, the three species of *Schistosoma* may be readily differentiated by their characteristic egg morphology (Table 60.11 and Fig. 60.45 and Fig. 60.51). The eggs of *S. mansoni* are oval, possess a sharp lateral spine, and measure 114 to 175 µm × 45 to 70 µm. Those of *S. haematobium* differ primarily in the terminal location of the spine. The eggs of *S. japonicum* are more nearly circular, possess a minute lateral spine, and measure 70 to 100 µm × 55 to 65 µm.

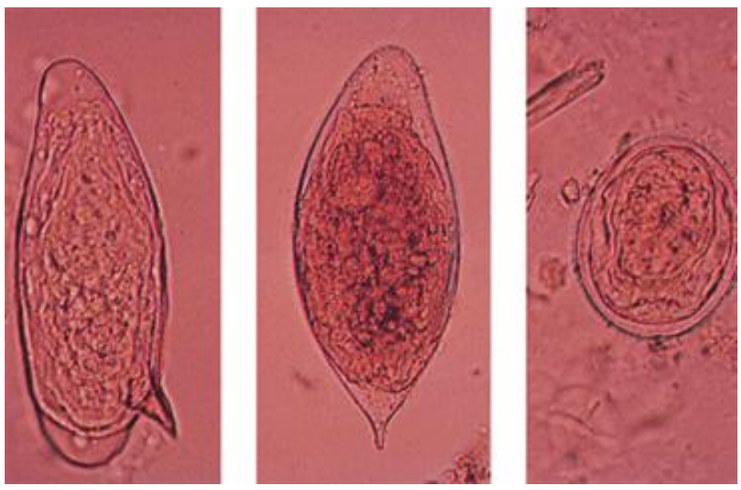


FIGURE 60.51. Eggs of *Schistosoma mansoni* **(left)**, *S. haematobium* **(center)**, and *S. japonicum* **(right)** (high power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

Examination of the eggs for viability is necessary to confirm the presence of active infection (76). This is accomplished most easily by the detection of movement of flame cell cilia within the egg on microscopic examination. Alternatively, the eggs may be hatched in water. Quantitation of egg output in stool or urine is useful in estimating the severity of infection and the response to therapy.

Serologic testing is available but is largely of epidemiologic interest only. The development of newer tests using stage-specific antigens may allow the distinction of active from inactive disease and thus have greater clinical utility.

Specific therapy with praziquantel or oxamniquine may be indicated for individuals with moderate or severe active infections. Antihelmintic therapy may terminate oviposition but will not affect lesions caused by eggs already deposited in tissues. Schistosomal dermatitis and the Katayama syndrome may be treated with the administration of antihistamines and corticosteroids.

Prevention of schistosomiasis has been difficult (76). Proper sanitation and the use of molluscacides are necessary steps in interrupting the transmission of the parasite. Vaccine development is an area of intense research interest, but a candidate vaccine is not available for human use at this time.

Cestodes

The cestodes or tapeworms, members of the Platyhelminthes, are generally long, flat, segmented, ribbonlike intestinal parasites (75).

The adult worms vary in length; some species are barely visible to the naked eye, whereas others are 20 to 25 feet long. The adult tapeworms are divided into three distinct body parts: the anterior portion or scolex, a generative neck, and a long, segmented body, the strobila. The head or scolex is equipped with various structures for attachment to the intestinal mucosa. These attachment structures vary with species but usually include muscular suckers (*Taenia* species), long, lateral grooves or bothria (*Diphyllobothrium latum*), and a retractable rostellum armed with a crown of chitinous hooks (*T. solium*) (Fig. 60.52). Posterior to the scolex is the neck, which gives rise to the strobila, consisting of individual body segments or proglottids.

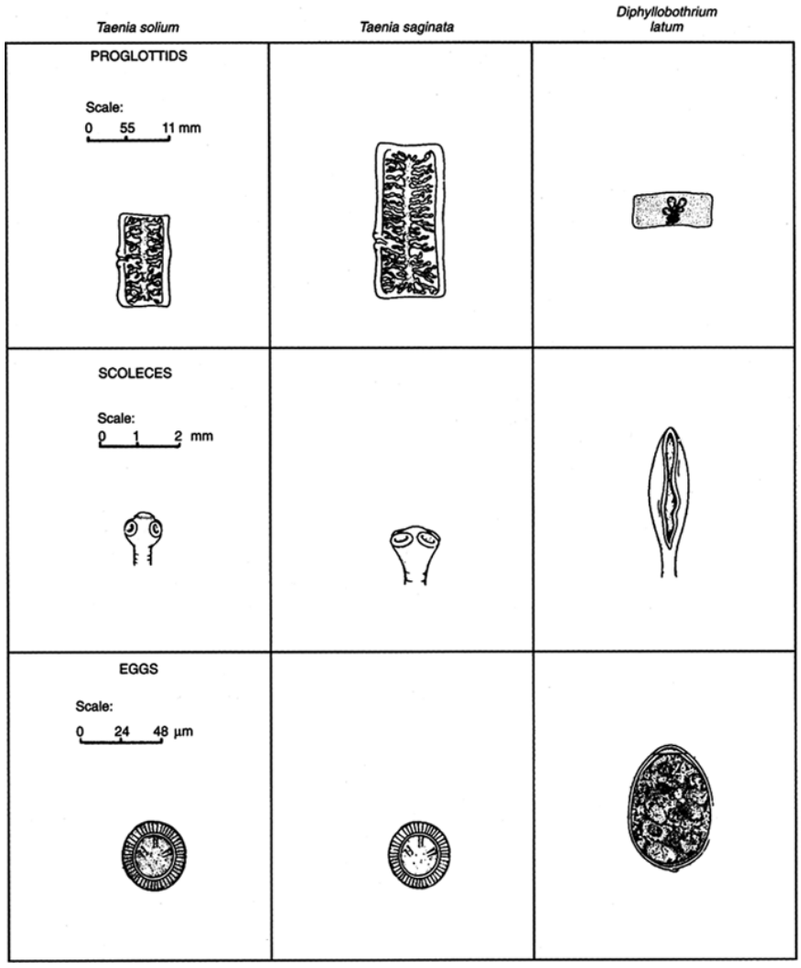


FIGURE 60.52. Diagnostic features of *Taenia solium*, *T. saginata*, and *Diphyllobothrium latum*. (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

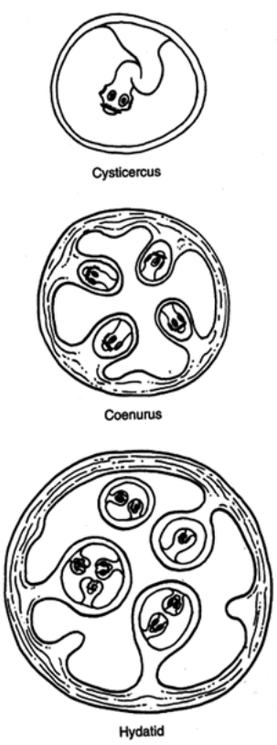


FIGURE 60.53. Tapeworm cysts. (Adapted from Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

All tapeworms are hermaphroditic, and each proglottid contains both male and female reproductive organs. Proglottids develop from immature to mature to egg producing, with the most developed segments located farthest from the neck. Each self-contained proglottid is joined to the rest of the worm by a common cuticle, nerve trunks, and excretory canals. Tapeworms have no digestive system, and nutrients are absorbed from the host intestine through the complex cuticle of the worm. Gravid proglottids release their eggs by rupturing, disintegrating, or passing the eggs through a uterine pore. In some species (*T. solium* and*T. saginata*), eggs are not released from intact proglottids; instead, gravid proglottids separate from the remainder of the worm and are passed intact in the feces. The eggs of most cestodes are nonoperculated and contain a fully developed, six-hooked (hexacanth) embryo. In contrast, the eggs of the fish tapeworm *D. latum* are immature at the time of deposition and possess an operculum, similar to fluke eggs, through which the embryo exits when fully developed (Fig. 60.52).

The life cycles of all cestodes are complex and, with the exception of *Hymenolepis nana*, require passage of the larvae through one or more intermediate hosts. For most cestodes, the eggs are ingested by the specific intermediate host and hatch within its gut. The released larvae then penetrate the intestinal mucosa and migrate to the tissues, where they encyst and develop to the infective stage. The life cycle is completed when infective larvae are ingested by the appropriate definitive host. *D. latum*, whose eggs are immature on release, requires two intermediate hosts to complete its larval development: aquatic invertebrates (copepods) and fish. The larval stages present in the tissues of intermediate hosts vary with the different species of cestode and include cysts with a single immature scolex or proscolex (cysticercus; *Taenia* species), cysts with multiple proscolices (coenurus; *Multiceps* species), and cysts containing multiple daughter cysts, each with multiple proscolices (hydatid cyst; *Echinococcus* species) (Fig. 60.53). In some instances, humans serve as an intermediate host that harbors larval stages. These extraintestinal larval infections are named according to the larval stage found in tissue (cysticercosis, coenurosis, hydatid disease) and are usually more serious than the presence of adult worms in the intestine.

Numerous cestodes infect humans, including *T. saginata*, *T. solium*, *D. latum*, *Echinococcus granulosus*, *E. multilocularis*, *Hymenolepis diminuta*, *H. nana*, *Dipylidium caninum*, and *Multiceps multiceps* (75). Together they infect approximately 65 million individuals, producing embarrassment, discomfort, anemia, and occasionally death. Of the many cestodes causing intestinal and tissue infections in humans, only four representative agents will be discussed: the pork tapeworm and cause of cysticercosis (*T. solium*), the beef tapeworm (*T. saginata*), the fish tapeworm (*D. latum*), and *E. granulosus*, the agent of hydatid disease. Basic details of the common intestinal cestodes of medical importance are provided in Table 60.12.

TABLE 60.12. MORPHOLOGY OF CESTODES

	Diphyllobothrium latum	*Taenia saginata*	*Taenia solium*	*Hymenolepis nana*	*Hymenolepis diminuta*
Method of infection	Ingestion of raw fish	Ingestion of undercooked beef	Ingestion of undercooked pork and ova	Ingestion of cysticercoid in infected arthropod, ingestion of egg, autoinfection	Ingestion of cysticercoid in infected arthropod
Location in host	Adult in small intestine	Adult in small intestine	Adult in small intestine;	Adult in small intestine; larvae in villi of small intestine	Adult in small intestine
Adult worm	4–10 m in length, gravid proglottids 3 × 11 mm, rosette-shaped central uterus, genital pore on mid-ventral surface	4–8 m in length, no rostellum or hooklets, mature proglottids, the ovary has 2 lobes and a vaginal sphincter muscle, gravid proglottids 18–20 mm × 5–7 mm, 15–30 lateral branches on each side of central uterine stem	3–5 m in length, mature proglottids, the ovary has 2 lobes and accessory lobe; vaginal sphincter muscle is absent; gravid proglottids 11 × 5 mm, 7–13 lateral branches on each side of central uterine stem	Small, 2.5–4 cm long, proglottids are wider than they are long	20–60 cm long, proglottids are wider than they are long
Scolex	Small, 3 × 1 mm, spatulate with 2 shallow grooves	Small, 1–2 mm in diameter with 4 suckers	Small, 1 mm in diameter, 4 suckers, rostellum with 2 rows of hooklets	Small, knoblike with 4 suckers, rostellum, and hooklets	Knoblike, 4 suckers, rostellum, no hooklets
Egg morphology	Ovoid, operculated, thick-shelled, 58–75 × 40–50 µm	Spherical, 31–43 µm, thick, radially striated; contains 6-hooked embryo (oncosphere)	Spherical, 31–43 µm thick, radially striated; contains 6-hooked embryo (oncosphere)	Spherical, thin, hyaline shell 30–47 µm in diameter, 6 hooked oncosphere, polar filaments	Spherical, thick-shelled, 70–85 × 60–80 µm, 6-hooked oncosphere, no polar filaments
Clinical manifestations	Intestinal obstruction, diarrhea, abdominal pain, or anemia, possible vitamin B_{12} deficiency	Obstruction, diarrhea, hunger pains, weight loss, appendicitis	Adult worm: hunger pains, indigestion, diarrhea, constipation, eosinophilia; cysticercus: epileptiform seizures, abnormal behavior, transient paresis, intermittent obstructive hydrocephalus, disequilibrium meningoencephalitis	Headache, dizziness, anorexia, abdominal pain, diarrhea	Infection is usually tolerated; very few symptoms
Diagnostic methods	O&P exam	O&P exam, HAI, ELISA	HAI, O&P exam, ELISA	O&P exam	O&P exam
Specimen	Feces	Feces, proglottids, scolex	Feces, proglottids, scolex	Feces	Feces
Common name	Broad fish tapeworm	Beef tapeworm	Pork tapeworm	Dwarf tapeworm	Rat tapeworm

O&P exam, ova and parasite examination; HAI, hemagglutination inhibition; ELISA, enzyme-linked immunosorbent assay.

Taenia solium

T. solium, the pork tapeworm, inhabits the human jejunum, where it may survive for decades and grow to a maximum length of 3 to 5 m. Its small scolex is armed with four sucking disks and a rostellum with a double row of hooklets (Fig. 60.52). Differentiation of *T. solium* from *T. saginata* may be accomplished by examination of the scolex and proglottids (Fig. 60.52). Gravid proglottids of *T. solium* are smaller than those of *T. saginata* and contain only seven to 13 lateral uterine branches versus 15 to 30 for the beef tapeworm. *T. solium* is widely distributed throughout much of the world and is particularly common in Eastern Europe, Asia, Africa, and Latin America. It is seen infrequently in the United States.

As indicated previously, the gravid proglottids of *Taenia* species become detached from the strobila and either rupture, releasing the characteristic eggs within the intestine, or are passed intact in the feces (Fig. 60.54). The eggs are spherical, 30 to 40 µm in diameter, and possess a thick, radially striated shell containing a six-hooked hexacanth embryo (Fig. 60.55). Both pigs and humans may become intermediate hosts when they ingest food contaminated with viable eggs. Autoinfection may occur when eggs are transferred from the perianal area to the mouth on contaminated fingers. When ingested, the eggs hatch in the

stomach of the intermediate host, releasing the hexacanth embryo or oncosphere. The oncosphere penetrates the intestinal wall and migrates in the circulation to the tissues, where it develops into a cysticercus over 3 to 4 months. The cysticerci may develop in muscle, connective tissue, brain, lung, and eyes and remain viable for as long as 5 years. Humans may become the definitive host for this parasite when they ingest inadequately cooked pork containing cysticerci. Digestion of the infected meat releases the cysticercus, which then everts, exposing the scolex, and attaches to the intestinal mucosa, where it develops into an adult worm, thereby completing the cycle.

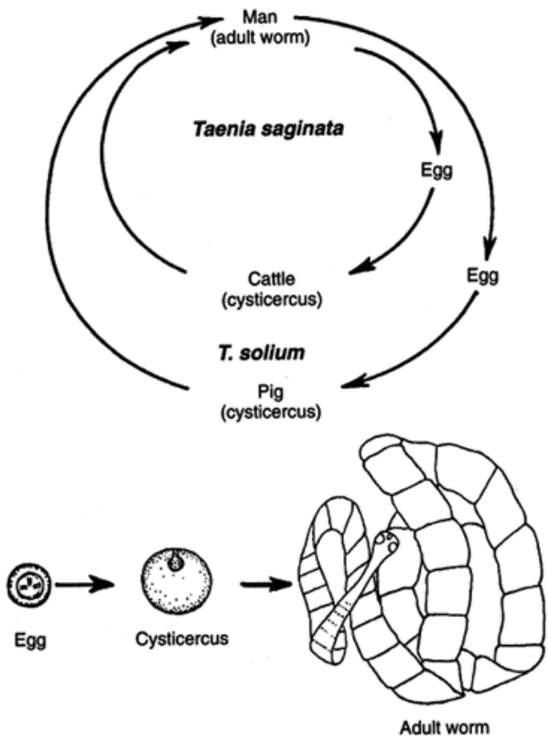


FIGURE 60.54. Life cycles of *Taenia saginata* and *T. solium*. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.)

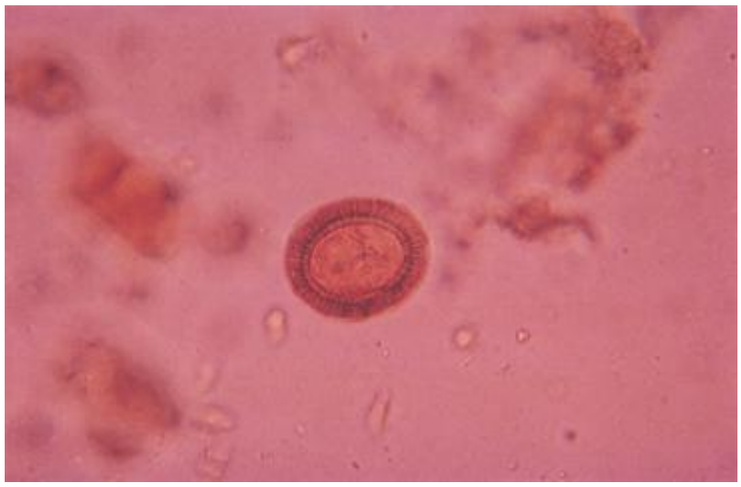


FIGURE 60.55. *Taenia* species, egg (high power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

The symptoms produced by the presence of the adult *T. solium* in the intestine are minimal. Most patients are asymptomatic and become aware of the infection only when they see proglottids in their feces. Occasionally, individuals report abdominal discomfort, nausea, diarrhea, and weight loss.

An entirely different clinical picture may present when humans are the intermediate host. Cysticercosis may be of little consequence if the larvae lodge in the muscles or subcutaneous tissues; however, serious disease may follow as cysticerci develop in vital areas such as the brain or eyes (88, 89). In the brain, they may produce seizures, meningoencephalitis, hydrocephalus, and cranial nerve damage. In the eye, loss of visual acuity and visual field defects may occur. Tissue reaction to viable larvae may be only moderate, thus minimizing symptoms; however, the death of the larvae stimulates a marked inflammatory reaction with exacerbation of symptoms and resulting fever, muscle pains, and eosinophilia.

The laboratory diagnosis of intestinal infection is made by finding eggs or proglottids in the stool (75). As the eggs of *T. solium* and *T. saginata* are identical (Fig. 60.55), it is necessary to examine a proglottid to identify the species correctly (Fig. 60.52). The diagnosis of cysticercosis is usually established by detection and surgical removal of soft-tissue nodules, the appearance of calcified cysticerci in soft-tissue roentgenograms, and detection of cysts in the eye (87). Central nervous system lesions may be detected by computed tomography, radioisotope scanning, or ultrasonography (88, 89). Serologic studies may be useful, but false positives may occur in individuals with other helminthic infections.

The treatment of choice for intestinal infection is niclosamide, which acts directly on the worm. Praziquantel, paromomycin, and quinacrine are effective alternatives. The drug of choice for cysticercosis is praziquantel (89). Concomitant steroid administration may be useful in minimizing the inflammatory response to the dying larvae. Surgical removal of cerebral and ocular cysts may be necessary. Preventive measures include proper sanitation, adequate cooking of all pork products, and prompt treatment of all human cases of *T. solium* infection to minimize egg transmission.

Taenia saginata

T. saginata, the beef tapeworm, closely resembles *T. solium* and may parasitize the jejunum and small intestine of humans for as long as 25 years, attaining a length of 10 m. *T. saginata* bears a scolex armed with four muscular suckers but no hooklets, and its proglottids contain 15 to 30 uterine branches. *T. saginata* is rare in the United States but is highly prevalent in Africa, the Middle East, eastern Europe, and South America.

The life cycle of *T. saginata* is virtually identical to that described for *T. solium* (Fig. 60.54). A major difference is that humans rarely, if ever, act as intermediate hosts for *T. saginata* (76). As with *T. solium*, humans are infected when they ingest inadequately cooked meat containing cysticerci.

The clinical syndrome associated with *T. saginata* is similar to intestinal infection with *T. solium*. Patients are generally asymptomatic or may complain of vague abdominal pains, chronic epigastric discomfort, and weakness.

As with *T. solium*, the diagnosis is made by detection of eggs and proglottids in feces. Motile proglottids may be observed in freshly passed stool. Differentiation of *T. saginata* from *T. solium* is made by examination of mature proglottids or recovery of the entire worm with a scolex lacking hooklets.

Treatment is identical to that described for the intestinal phase of *T. solium*. A single dose of niclosamide is highly effective in elimination of the adult worm. Prevention depends on proper sanitation and adequate cooking of beef.

Diphyllobothrium latum

D. latum, the fish tapeworm, inhabits the human jejunum and ileum, where it may survive for many years and grow to a maximum

length of 4 to 20 m. Its elongated, fusiform scolex is armed with two lateral sucking grooves or bothria (Fig. 60.52). Unlike the *Taenia* species, the proglottids are uniformly wider than they are long and contain a centrally positioned, rosette-shaped uterus. The large (58 to 75 µm × 40 to 50 µm), ovoid, operculated eggs (Fig. 60.52 and Fig. 60.56) are discharged through the uterine pore of the proglottid. More than 1 million eggs are released per day into the fecal stream. *D. latum* is found worldwide, most prevalently wherever raw, pickled, or undercooked freshwater fish (including salmon) is eaten by humans (75). Human infections are found in central and northern Europe, Scandinavia, Siberia, China, and Japan. In North America, endemic foci have been described in Alaska, Canada, and the Great Lakes region of the United States.

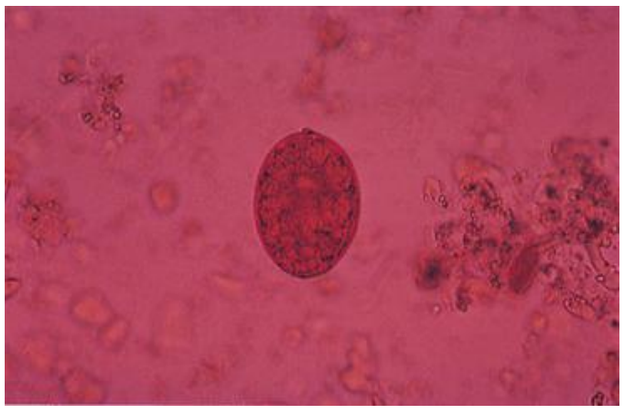


FIGURE 60.56. *Diphyllobothrium latum*, egg (low power). (From Smith JW, et al. *Diagnostic medical parasitology: intestinal helminths*. Chicago: American Society of Clinical Pathologists, 1976.)

The life cycle of *D. latum* differs from that of the *Taenia* species in that it requires two intermediate hosts (Fig. 60.57). The unembryonated, operculate eggs are shed in the feces and, on reaching fresh water, require a period of 2 to 4 weeks to develop a ciliated, free-swimming larval form called a coracidium. The fully developed coracidium leaves the egg via the operculum and is ingested by small freshwater crustaceans (e.g., *Cyclops* and *Diaptomus*), where it develops into a procercoid larval form. The crustacean is then eaten by a fish, and the infectious plerocercoid larvae develop in the musculature of the fish. Humans are infected when they eat raw or undercooked fish containing the plerocercoid larval forms.

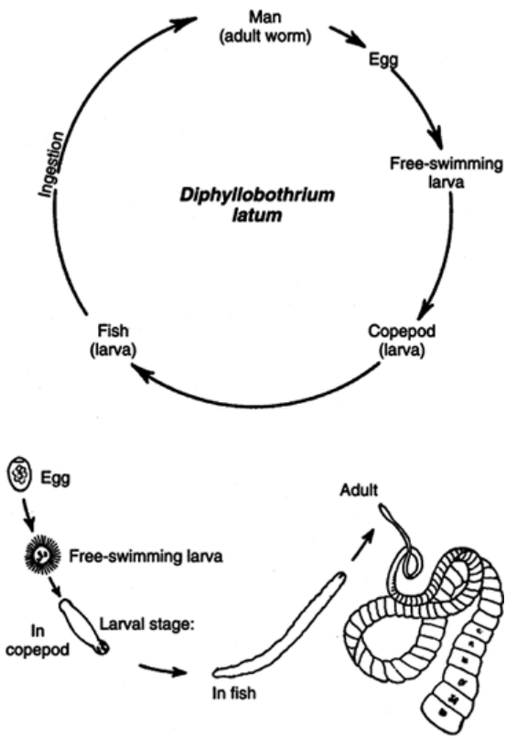


FIGURE 60.57. Life cycle of *Diphyllobothrium latum*. (From Markell EK, Voge M, John DT. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.).

Clinically, as is the case with most tapeworm infections, most *D. latum* infections are asymptomatic. Occasionally individuals complain of epigastric pain, abdominal cramping, nausea, vomiting, and weight loss. As many as 40% of persons infected with *D. latum* may have low serum vitamin B_{12} levels, presumably as a result of the competition between the host and the worm for dietary vitamin B_{12}, but only 0.1% to 2% develop clinical signs of vitamin B_{12} deficiency, including megaloblastic anemia and neurologic manifestations such as numbness, paresthesia, and loss of vibration sense.

The diagnosis of *D. latum* infection is made by detection of the bile-stained operculated egg (Fig. 60.56) or typical proglottids in stool. Concentration techniques are usually not necessary, as the worms produce large numbers of ova.

Treatment is similar to that described for *T. saginata* infections: single-dose therapy with niclosamide is highly effective. Vitamin B_{12} supplementation may be necessary in individuals with evidence of clinical vitamin B_{12} deficiency (e.g., anemia and neurologic symptoms). Thorough cooking of all salmon and freshwater fish and attention to proper sanitation are necessary preventive measures.

Echinococcus granulosus

Echinococcosis or hydatid disease is a tissue infection of humans caused by the larvae of the cestode *E. granulosus*. The adult *E. granulosus* is a small (5 mm) intestinal parasite of canines (e.g., dog, fox, wolf, coyote). The worm consists of a *Taenia*-like scolex with four sucking disks and a double row of hooklets and a strobila containing three proglottids: one immature, one mature, and one gravid. The eggs, which are identical in appearance to those of *Taenia* species (Fig. 60.55), are released from the terminal gravid proglottid and are shed in the feces, where they may be ingested by one of a number of mammalian intermediate hosts, including sheep, goats, deer, moose, and humans (Fig. 60.58). The ingested eggs hatch and the embryos penetrate the intestinal mucosa and are carried by the circulation to lodge in the liver, lungs, brain, heart, bone, and other tissues. Within the tissues, the larvae form hydatid cysts characterized by a laminated

germinative membrane from which daughter cysts or brood capsules arise, each containing multiple inverted proscolices (Fig. 60.53). The cysts expand slowly, accumulating fluid and hydatid sand composed of scoleces, hooks, and calcareous corpuscles from disintegrating brood capsules. Cysts generally reach a diameter of 1 cm over 6 months and ultimately may attain a size of 5 to 20 cm over a period of several years. Spillage of cyst contents may lead to the development of cysts in other sites, as the proscolices have the germinative potential to form new cysts. When hydatid-containing tissues of the intermediate hosts are ingested by a canine, thousands of scolices are released in the intestine to develop into adult worms. Humans and herbivores do not become infected with the adult worm.

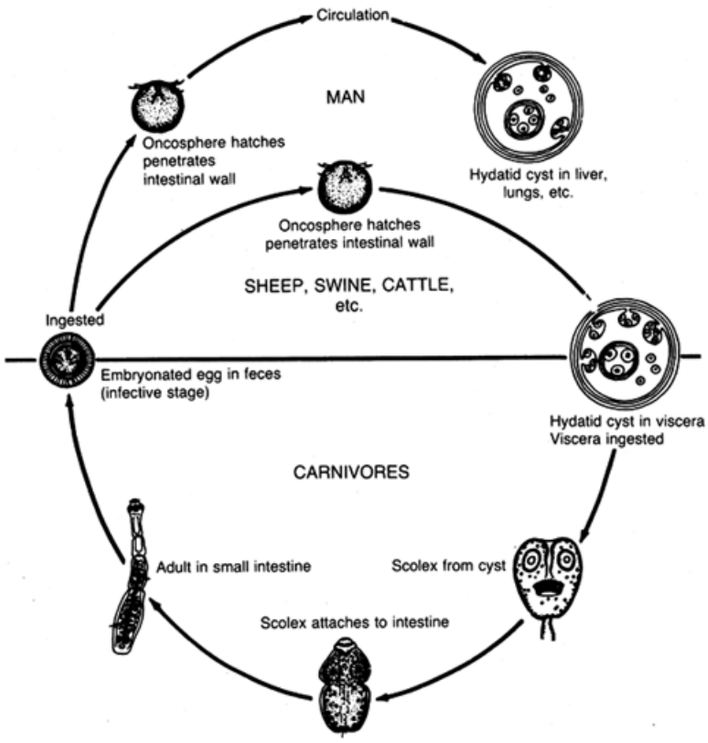


FIGURE 60.58. Life cycle of *Echinococcus granulosus*. (From Strickland GT. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.)

Human infection with *E. granulosus* most commonly occurs in countries where domestic herbivores such as sheep, cattle, and goats are raised in close contact with dogs (76). The highest incidence of hydatid disease is seen in Australia, New Zealand, South America, South and East Africa, Central Europe, and the Middle East. It also occurs in Alaska, Canada, and the sheep farming regions of the United States (Utah, New Mexico, Arizona, California). Humans become infected after ingestion of contaminated food and water, as well as from hand-to-mouth transmission of eggs as a result of handling dogs infected with the adult worm.

Human infection typically is asymptomatic for as long as 5 to 20 years after acquisition of the infection. The clinical manifestations of hydatid disease are owing to the space-occupying nature of the slowly expanding cyst and thus are directly related to the number, anatomic location, and rate of growth of the cysts. In the majority of cases, the cysts are in the lung or the liver, and symptoms include cough, hemoptysis and chest pain, or abdominal pain and tenderness. Rupture of the cysts occurs in 20% of cases, producing fever, urticaria, and occasionally anaphylactic shock and death owing to the release of antigenic cyst contents. Cyst rupture may also lead to dissemination of infection caused by the release of thousands of proscolices. Cysts in bone may cause pathologic fractures, whereas those in the brain present as tumorlike, space-occupying lesions with seizures, visual disturbances, and hydrocephalus. Over time, the cyst may die and become calcified (87).

Diagnosis of hydatid disease is difficult and depends primarily on clinical, radiographic, and serologic findings. Cysts may be observed on routine roentgenograms or by computed tomographic or ultrasonic scanning procedures (90). Although aspiration of cyst contents and demonstration of hydatid sand (e.g., scolices, hooks) may be diagnostic, it is contraindicated because of the risk of anaphylaxis and dissemination of infection. Serologic testing may be useful but is negative in 10% to 40% of infections.

Surgical excision of hydatid cysts is the therapy of choice for symptomatic patients. Aspiration of cyst contents and instillation of formalin, hypertonic saline, silver nitrate, or cetrimide to kill the proscolices and detoxify the remaining fluid may be indicated. If surgery is contraindicated, medical therapy with high-dose albendazole, mebendazole, or praziquantel may be considered. Preventive measures include attention to personal hygiene

and hand washing and deworming of infected dogs. Dogs should never be allowed to feed on the viscera of slain animals.

REFERENCES

1. Britto C, Cardoso MA, Monteiro Vanni CM, et al. Polymerase chain reaction detection of *Trypanosoma cruzi* in human blood samples as a tool for diagnosis and treatment evaluation. *Parasitology* 1995;110:241–247.

2. Garcia LS, Shimizu RY. Evaluation of nine immunoassay kits (enzyme immunoassay and direct fluorescence) for detection of *Giardia lamblia* and *Cryptosporidium parvum* in human fecal specimens. *J Clin Microbiol* 1997;35:1526–1529.

3. Maddison SE. Serodiagnosis of parasitic diseases. *Clin Microbiol Rev* 1991;4:457-469.

4. Oelemann WMR, Teixeira MDGM, DaCosta GCV, et al. Evaluation of three commercial enzyme-linked immunosorbent assays for diagnosis of Chagas' disease. *J Clin Microbiol* 1998;36:2423–2427.

5. Weiss JB. DNA probes and PCR for diagnosis of parasitic infections. *Clin Microbiol Rev* 1995;8:113–130.

6. Wilson SM. DNA based methods in the detection of *Leishmania* parasites: field applications and practicalities. *Ann Trop Med Parasitol* 1995;89(Suppl 1):95–100.

7. Garcia LS, Bruckner DA. *Diagnostic medical parasitology*, 3rd ed. Washington, DC: American Society for Microbiology, 1997.

8. Markell EK, Voge M, John DT, eds. *Medical parasitology*, 7th ed. Philadelphia: Saunders, 1992.

9. Sherris JC, ed. *Medical microbiology: an introduction to infectious diseases*, 2nd ed. New York: Elsevier, 1990.

10. Strickland GT, ed. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991.

11. Boland PB, Neafie RC, Marty AM. Malaria: a re-emerging disease. In: Nelson AM, Horsburgh CR Jr, eds. *Pathology of emerging infections* 2. Washington, DC: American Society for Microbiology, 1998:283–316.

12. Strickland GT. Malaria. In: Strickland GT, ed. *Hunter's tropical medicine*, 7th ed. Philadelphia: Saunders, 1991:486–602.

13. Barker RH, Banchongaksorn T, Courval JM, et al. A simple method to detect *Plasmodium falciparum* directly from blood using the polymerase chain reaction. *Am J Trop Med Hyg* 1992;46:416–426.

14. Craig MH, Sharp BL. Comparative evaluation of four techniques for the diagnosis of *Plasmodium falciparum* infections. *Trans R Soc Trop Med Hyg* 1997;91:279–282.

15. Rickman LS, Long GW, Oberst R, et al. Rapid diagnosis of malaria by acridine organe staining of centrifuged parasites. *Lancet* 1989;1:68–71.

16. Shiff CJ, Premji Z, Minjas JN. The rapid manual ParaSight®-F test. *Trans R Soc Trop Med Hyg* 1993;87:646–648.

17. Barid JK, Bassi H, Purnomo MJ, et al. Resistance to chloroquine by *Plasmodium vivax* in Irian Jaya, Indonesia. *Am J Trop Med Hyg* 1991;44:547–552.

18. Garavelli PL, Corti E. Chloroquine resistance in *Plasmodium vivax*: the first case in Brazil. *Trans R Soc Trop Med Hyg* 1992;86:128.

19. Myat-Phone-Kyaw, Myint-Oo, Myint-Lwin, et al. Emergence of chloroquine-resistant *Plasmodium vivax* in Myanmar (Burma). *Trans R Soc Trop Med Hyg* 1993;87:687–689.

20. Phillips EJ, Keystone JS, Kain KC. Failure of combined chloroquine and high-dose primaquine therapy for *Plasmodium vivax* malaria acquired in Guyana, South America. *Clin Infect Dis* 1996;23:1171–1173.

21. Rieckmann KH, Campbell GH, Sax LJ, et al. *Plasmodium vivax* resistance to chloroquine? *Lancet* 1989;2:1183–1184.

22. Schuurkamp GJ, Spicer PE, Kereu RK, et al. Chloroquine-resistant *Plasmodium vivax* in Papua New Guinea. *Trans R Soc Trop Med Hyg* 1992;86:121–122.

23. Garcia LS, Sulzer AJ, Healy GR, et al. Blood and tissue protozoa. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 6th ed. Washington, DC: American Society for Microbiology, 1995:1171–1195.

24. Bruckner DA, Labarca JA. *Leishmania* and *Trypanosoma*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1365–1373.

25. Anonymous. Problems with leishmaniasis. *BMJ* 1978;2:1179–1180.

26. Nuzum E, White F III, Thakur C, et al. Diagnosis of symptomatic visceral leishmaniasis by use of the polymerase chain reaction on patient blood. *J Infect Dis* 1995;171:751–754.

27. Berman JD. Chemotherapy for leishmaniasis: biochemical mechanisms, clinical efficacy, and future strategies. *Rev Infect Dis* 1988;10:560–586.

28. Berman JD. Human leishmaniasis: clinical diagnostic, and chemotherapeutic developments in the last 10 years. *Clin Infect Dis* 1997;24:684–703.

29. Van Meirvenne N, LeRay D. Diagnosis of African and American trypanosomiasis. *Br Med Bull* 1985;41:156–161.

30. McCabe RE, et al. Ketoconazole protects against infection with *Trypanosoma cruzi* in a murine model. *Am J Trop Med Hyg* 1983; 32:960-962.

31. Wilson M, McAuley JB. *Toxoplasma*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1374–1382.

32. Luft BJ, Remington JS. Toxoplasmic encephalitis. *J Infect Dis* 1988;157:1–6.

33. Oksenhendler E, Cadranel J, Sarfati C, et al. *Toxoplasma gondii* pneumonia in patients with the acquired immunodeficiency syndrome. *Am J Med* 1990;88(Suppl 5):18N–21N.

34. Crider SR, Horstman WG, Massey GS. *Toxoplasma orchitis*: report of a case and a review of the literature. *Am J Med* 1988;85:421–424.

35. Hohlfeld P, Daffos F, Costa JM, et al. Prenatal diagnosis of congenital toxoplasmosis with a polymerase-chain reaction test on amniotic fluid. *N Engl J Med* 1994;331:695–699.

36. Wilson M, Remington JS, Clavet C, et al. Evaluation of six commercial kits for detection of human IgM antibodies to *Toxoplasma gondii*. *J Clin Microbiol* 1997;35:3112–3115.

37. Pedrol E, Gonzalez-Clemente JM, Gatell JM, et al. Central nervous system toxoplasmosis in AIDS patients: efficacy of an intermittent maintenance therapy. *AIDS* 1990;4:511–517.

38. McCabe R, Chirurgi V. Issues in toxoplasmosis. *Infect Dis Clin North Am* 1993;7:587–604.

39. Visvesvara GV. Pathogenic and opportunistic free-living amebae. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. In: *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1383–1390.

40. Visvesvara GS, Schuster FL, Martinez AJ. *Balamuthia mandrillaris*, new genus, new species, agent of amebic meningoencephalitis in humans and animals. *J Eukaryot Microbiol* 1993;40:504–514.

41. May LP, Sidhu GS, Buchness MR. Diagnosis of *Acanthamoeba* infection by cutaneous manifestations in a man seropositive to HIV. *J Am Acad Dermatol* 1992;26:352–355.

42. Murakawa GJ, McCalmont T, Altman J, et al. Disseminated acanthamoebiasis in patients with AIDS. A report of five cases and a review of the literature. *Arch Dermatol* 1995;131:1291–1296.

43. Sisan JP, Kemper CA, Loveless M, et al. Disseminated *Acanthamoeba* infection in patients with AIDS: case reports and review. *Clin Infect Dis* 1995;20:1207–1216.

44. Valenstein P, Pfaller M, Yungbluth M. The use and abuse of routine stool microbiology: a College of American Pathologists Q-Probes study of 601 institutions. *Arch Pathol Lab Med* 1996;120:206–211.

45. Leber Al, Novak SM. Intestinal and urogenital amebae, flagellates and ciliates. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. Manual of clinical microbiology, 7th ed. Washington, DC: American Society for Microbiology, 1999:1391–1405.

46. Garcia LS, Bruckner DA, Clancy MN. Clinical relevance of *Blastocystes hominis*. *Lancet* 1984;11:1233–1234.

47. Zierdt CH. *Blastocystis hominis*: past and future. *Clin Microbiol Rev* 1991;4:61–79.

48. Diamond LS, Clark CG. A redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) separating it from *Entamoeba dispar* Brumpt, 1925. *J Eukaryot Microbiol* 1994;40:340–344.

49. Ong SJ, Cheng MY, Liu KH, et al. Use of the ProSpect® microplate enzyme immunoassay for the detection of pathogenic and nonpathogenic *Entamoeba histolytica* in faecal specimens. *Trans R Soc Trop Med Hyg* 1996;90:248–249.

50. Sanuki J, Asai T, Okuzawa E, et al. Identification of *Entamoeba histolytica* and *E. dispar* cysts in stool by polymerase chain reaction. *Parasitol Res* 1997;83:96–98.

51. Bruckner DA. Amebiasis. *Clin Microbiol Rev* 1992;5:356–369.

52. Wolfe MS. Giardiasis. *Clin Microbiol Rev* 1992;5:93–100.

53. Garcia LS, Shum AC, Bruckner DA. Evaluation of a new monoclonal antibody combination reagent for direct fluorescence detection of *Giardia* cysts and *Cryptosporidium* oocysts in human fecal specimens. *J Clin Microbiol* 1992;30:3255–3257.

54. Flynn PM. Emerging diarrheal pathogens: *Cryptosporidium parvum*, *Isospora belli*, *Cyclospora* species, and microsporidia. *Pediatr Ann* 1996;25:480–487.

55. Ortega YR. *Cryptosproidium*, *Cyclospora*, and*Isospora*. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1406–1412.

56. Current WL, Garcia LS. Cryptosporidiosis. *Clin Microbiol Rev* 1991;4:325–358.

57. Balatbat AB, Jordan GW, Tang YJ, et al. Detection of *Cryptosporidium parvum* DNA in human feces by nested PCR. *J Clin Microbiol* 1996;34:1769–1772.

58. Ortega YR, Sterling CR, Gilman RH, et al. *Cyclospora* species: a new protozoan pathogen of humans. *N Engl J Med* 1993;328:1308–1312.

59. Soave R. *Cyclospora*: an overview. *Clin Infect Dis* 1996;23:429–437.

60. Pape JW, Verdier RI, Boney M, et al. *Cyclospora* infection in adults infected with HIV. Clinical manifestations, treatment, and prophylaxis. *Ann Intern Med* 1994;121:654–657.

61. Herwaldt BL, Ackers M-L, the *Cyclospora* Working Group. An outbreak in 1996 of cyclosporiasis associated with imported raspberries. *N Engl J Med* 1997;336:1548–1558.

62. Hoge CW, Shlim DR, Ghimire M, et al. Placebo-controlled trial of co-trimoxazole for *Cyclospora* infections among travellers and foreign residents in Nepal. *Lancet* 1995;345:691–693.

63. DeGroote MA, Visvesvara GS, Wilson ML, et al. Polymerase chain reaction and culture confirmation of disseminated *Encephalitozoon cuniculi* infection in a patient with AIDS: successful therapy with albendazole. *J Infect Dis* 1995;171:1375–1378.

64. Weber R, Bryan RT, Schwartz DA, Owen Rl. Human microsporidial infections. *Clin Microbiol Rev* 1994;7:426–461.

65. Weber R, Canning EU. Microsporidia. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1413–1420.

66. Canning EU, Hollister WS. *Enterocytozoon bieneusi* (*Microspora*): prevalence and pathogenicity in AIDS patients. *Trans R Soc Trop Med Hyg* 1990;84:181–186.

67. Orenstein JM, Chiang J, Steinberg W, et al. Intestinal microsporidiosis as a cause of diarrhea in human immunodeficiency virus-infected patients: a report of 20 cases. *Hum Pathol* 1990;21:475–481.

68. Enriquez FJ, Ditrich O, Palting JD, et al. Simple diagnosis of *Encephaloitozoon* spp. infections by using a panspecific antiexospore monoclonal antibody. *J Clin Microbiol* 1997;35:724–729.

69. Ignatius R, Henschel S, Liesenfeld O, et al. Comparative evaluation of modified trichrome and Uritex 2B stains for detection of low numbers of microsporidial spores in stool specimens. *J Clin Microbiol* 1997;35:2266–2269.

70. Weber R, Bryan RT, Owen RL, et al. Improved light-microscopical detection of microsporidia spores in stool and duodenal aspirates. *N Engl J Med* 1992;326:161–166.

71. Colbourn NI, Hollister WS, Curry A, et al. Activity of albendazole against *Encephalitozoon cuniculi in vitro*. *Eur J Protistol* 1994;30:211–220.

72. Didier ES. Effects of albendazole, fumagillin, and TNP 470 on microsporidial replication *in vitro*. *Antimicrob Agents Chemother* 1997;41:1541–1546.

73. Goguel J, Katalama C, Sarafati C, et al. Remission of AIDS-associated intestinal microsporidiosis with highly active antiretroviral therapy. *AIDS* 1997;11:1658–1659.

74. Molina JM, Goguel J, Sarafati C, et al., for the Microsporidiosis Study Group, ANRS, France. Potential efficacy of fumagillin in intestinal microsporidiosis due to *Enterocytozoon bieneusi* in patients with HIV infection: results of a drug screening study. *AIDS* 1997;11:1603–1610.

75. Ash LR, Orihel TC. Intestinal helminths. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1421–1435.

76. Orihel TC, Ash LR. Tissue helminths. In: Murray PR, Baron EJ, Pfaller MA, et al., eds. *Manual of clinical microbiology*, 7th ed. Washington, DC: American Society for Microbiology, 1999:1436–1448.

77. Mahmoud AAF. Strongyloidiasis. *Clin Infect Dis* 1996;23:949–953.

78. Pelletier LL. Chronic strongyloidiasis in World War II Far East ex-prisoners of war. *Am J Trop Med Hyg* 1984;33:55–61.

79. Waldrop MM. Testing for trichinosis. *Science* 1985;227:621–624.

80. Davies HD, Sakuls P, Keystone JS. Creeping eruption. A review of clinical presentation and management of 60 cases presenting to a tropical disease unit. *Arch Dermatol* 1993;129:588–591.

81. MacKenzie CD, Williams JF, Sisley BM, et al. Variations in host responses and the pathogenesis of human onchocerciasis. *Rev Infect Dis* 1985;7:802–808.

82. Pacque M, Munoz B, Greene BM, et al. Community-based treatment of onchocerciasis with ivermectin: safety, efficacy, and acceptability of yearly treatment. *J Infect Dis* 1991;163:381–385.

83. Ro JY, Tsakalakis PF, White VA, et al. Pulmonary dirofilariasis: the great imitator of primary or metastatic lung tumor. A clinicopathologic analysis of seven cases and a review. *Hum Pathol* 1989;20:69–76.

84. Peries H, Carincross S. Global eradication of Guinea worm. *Parasitol Today* 1997;13:431–437.

85. Rim HJ. The current pathology and chemotherapy of clonorchiasis. *Korean J Parasitol* 1986;24(Suppl 3):1–141.

86. Johnson RJ, Jong EC, Dunning SB, et al. Paragonimiasis, diagnosis and the use of praziquantel in treatment. *Rev Infect Dis* 1985;7:200–206.

87. Orihel TC, Ash LR. *Parasites in human tissues*. Chicago: American Society of Clinical Pathologists, 1995.

88. Tsang VCW, Wilson M. *Taenia solium* cysticercosis: an under-recognized but serious public health problem. *Parasitol Today* 1995;11:124–126.

89. White AC Jr. Neurocysticercosis: a major cause of neurological disease world wide. *Clin Infect Dis* 1997;24:101–115.

90. von Sinner WN. Imaging of cystic echinococcosis. *Acta Trop* 1997;67:67–89.

Section 10
Immunopathology

Immunopathology - Introduction

David F. Keren

Section Chief

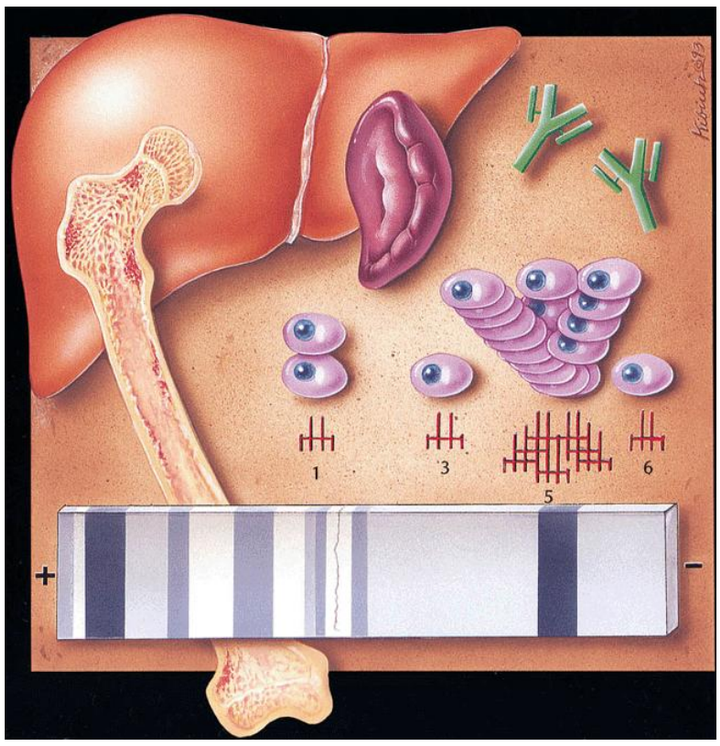


Figure.

The field of diagnostic immunology has become considerably broader in the past decade. The advent of flow cytometry, monoclonal reagents, and chemiluminescent probes, and the development of a clearer understanding of the immunopathology of autoimmune diseases have resulted in a tremendous increase in the number of types of diagnostic tests that are available. This section deals with the exploitation of immunologic techniques for clinical diagnosis.

Each chapter reviews features of diseases relevant to diagnostic immunology and discusses the techniques involved in the pathologic diagnosis. There is necessarily a modest degree of redundancy between some chapters within this section and those within other sections. For instance, some of the serology discussed here may also be relevant in the infectious disease sections. Some of the ligand assays discussed by Dr. England and Dr. Smart are complementary to information on radioimmunology in the chemistry section. Nonetheless, each chapter stands on its own with regard to the utility of the procedures discussed for making clinical diagnoses.

An attempt has been made to coordinate the chapters in this section to cover the field of diagnostic immunology as it relates to the diagnostic world today. Only minimal attention has been paid to the more historic aspects of the field. This is for the sake of readability and relevance to today's diagnostic procedures and is not meant to suggest that earlier techniques were not useful.

In each chapter, authors express their own opinions about particular procedures being discussed. This gives the reader important viewpoints from experts in the field.

61
Basic Principles of Immunodiagnosis

Roger S. Riley

F. Philip Anderson

The immune system is part of the body's defense against environmental agents (e.g., viruses, bacteria, fungi) and against our own cells that turn cancerous or learn to react against us. During the process of evolution, the immune system has acquired an impressive array of capabilities that are essential for its normal effective function. More than one trillion cells of various types constitute the immune system. Many of these cells are capable of differentiating “self” from “nonself,” rapidly destroying or disabling an invading organism, and “remembering” the encounter, so that an infectious organism is prevented from ever causing harm again. During the defense process, immune cells are constant and effective communicators through direct cell-to-cell contact or through the production of a wide variety of chemical intermediaries.

Scientists have applied an increasingly sophisticated array of tools and techniques to the investigation of the immune system. As a result, cells and cell products (antibodies, biological modifiers) with specific reactivity can be produced through laboratory manipulation of the immune response. These developments have greatly increased our understanding of the immune system and offer the promise of a natural, effective therapy for many immune and nonimmune diseases. In addition, immunologic products have been applied to medical diagnosis in a variety of imaginative ways and have become fundamental reagents in the clinical laboratory (1).

The term *immunoassay* refers to a biological test system that utilizes one or more immunological products or reagents. The basis for most immunoassays is the binding of an immunoglobulin molecule (antibody) to an antigen or hapten. However, other immunologic reagents, such as complement, are used in some immunoassays, and cell-to-cell interactions are the basis for others. In an immunoassay, the antigen may be another immunoglobulin molecule, a component of the cell membrane, or any other substance against which a specific antibody can be generated. The great specificity of the antigen-antibody reaction and the development of techniques that allow the large-scale production of antibodies have been the major factors in the clinical popularity of the immunoassay. To be useful in diagnosis, the product of an antigen-antibody interaction (immune complex) must be visualized or quantitatively measured by some means. Direct visualization with the naked eye is possible if the antigen is attached to cells or particles that clump together or agglutinate on interaction with the antibody. Although this simple interaction was the basis for early immunoassays and is still being used today, precise measurement of the antigen-antibody interaction requires that one or the other be labeled or tagged by some means.

BASIC IMMUNOLOGY

Part of "61 - Basic Principles of Immunodiagnosis"

The Immune Response

Animals respond to foreign substances (e.g., viruses or bacteria) by producing protein molecules in the blood called antibodies. Antibodies bind to the foreign substance as a part of the body's efforts to counteract it. Antibodies have the property of being highly specific, i.e., they will bind only to the foreign substance that caused them to be produced. Thus, the general term antigen designates any substance recognized by an animal as foreign and causing that animal to make antibodies.

The biological phenomenon of antibody production has been used in laboratory tests to identify substances found in human tissue. Because antibodies are highly specific for the antigen that caused them to be formed, they can be used to chemically locate that antigen in tissue. This is generally accomplished by the following method:

- A sample of the human substance (e.g., protein, enzyme, hormone) is injected into an animal (e.g., rabbit). Because the substance is of human origin, the animal recognizes it as foreign and makes antibodies against it. The human substance is the antigen causing the animal to produce antibodies. This process is known as immunization.
- A blood sample is then withdrawn from the animal and the anti-human antibodies are isolated and retrieved from the serum. These antibodies are highly specific for the human substance (antigen) that caused them to be formed and will not chemically bind with any other antigen in the tissue.
- In a laboratory, the antibodies can be applied to patient cell or tissue preparations. The antibody molecules will react with the type of human antigen that caused them to be made if it is present in the tissue. If the antigen is not present, no specific reaction will occur.

A *de novo* immune response begins with phagocytosis of the foreign invader, digestion and processing of the material, and presentation of some of the processed antigen on the cell surface (2,3,4 and 5). In conjunction with the self-HLA markers on the phagocyte and chemical mediators released from the phagocyte (monokines), the processed antigen is recognized by one kind of T lymphocyte (T-helper cell). These activated T-helper lymphocytes secrete a number of soluble substances (lymphokines, immune system hormones) that stimulate other immune cells, including suppressor T lymphocytes, which have a negative regulatory effect on the immune system. Another type of lymphocyte (B lymphocyte) produces soluble proteins termed antibodies (immunoglobulins). Each B cell is programmed to produce an antibody molecule that is highly specific for a single antigenic substance. When activated by lymphokines or by encounter with its antigen, a B cell rapidly reproduces. Each of these identical B cells differentiates into an antibody-producing factory (plasma cell) that can release thousands of antibody molecules per second (Fig. 61.1).

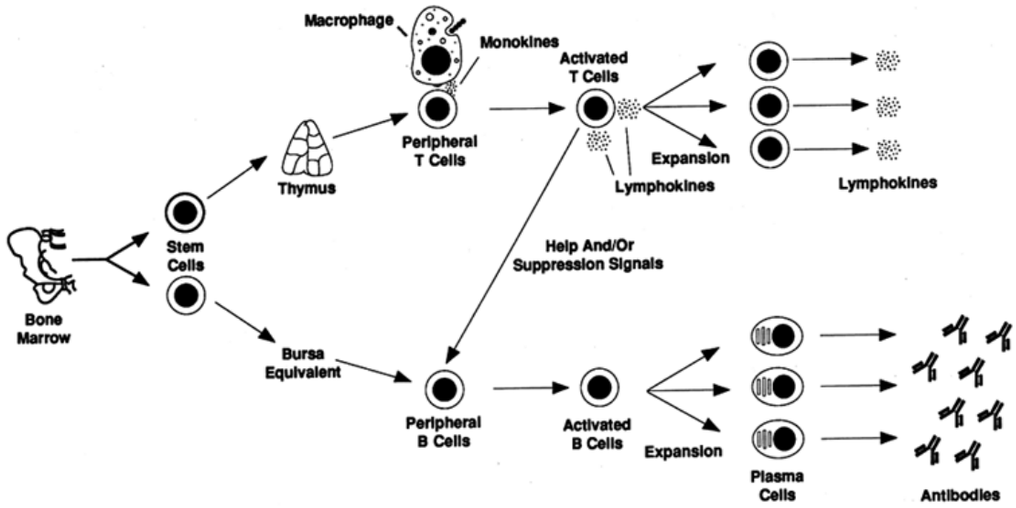


FIGURE 61.1. A schematic representation of the immune response.

Most T and B lymphocytes involved in an immune response die within several days. However, after each encounter, clones of memory cells persist in the body for life. If the offending antigen is ever encountered again, these cells are primed to quickly go into action without the delay faced during the first encounter.

A group of peptides termed cytokines is extremely important for immunoregulation and intracellular communication within the immune and inflammatory systems. These substances include the interleukins, interferons, tumor necrosis factors, and hemopoietic growth factors [colony-stimulating factors (CSF)] (6,7 and 8). Most of these substances were initially identified based on a single biological effect, but many have been subsequently found to have diverse local and even systemic effects. At the cellular level, cytokines interact through high-affinity surface receptors. The intracellular mechanism of action of many cytokines is currently unknown, although several cytokine receptors initiate the activation of adenylate cyclase, protein kinase C, tyrosine kinase, or other enzymes. The availability of immunoassays for the accurate measurement of cytokines in normal and diseased individuals and the discovery of recombinant DNA techniques to produce these substances in virtually unlimited quantity have led to a great interest in their potential diagnostic and therapeutic applications.

Polyclonal Antibodies

The natural immune response to an antigen is the generation of antibodies. Multiple clones of plasma cells respond to an antigenic challenge, and each produces antibodies that differ in antibody class, light-chain isotypes, idiotypic specificity, and combining ability (avidity). The number of antibody specificities generated depends on the complexity of the antigen. Although synthetic materials with repetitive epitopes may produce antibodies of a single specificity, hundreds to thousands of different antibodies are usually released in response to a large protein or cellular structure with multiple epitopes. Commercially, these polyclonal antibodies are usually obtained from animal sources. However, some antibodies, such as those with specificity for the HLA antigens, are commonly obtained from the serum of multiparous females or patients who have received multiple blood transfusions.

The commercial production of polyclonal antibodies from animals comprises a large industry that is closely protected by trade secrets. As a result, few details have been published in the literature (9,10). One exception is a review by Ritchie (11), who reported his experience gained from more than 10,000 animal bleedings.

The major considerations in the production of an antigen (immunogen) include:

- The size and chemical composition of the immunogen,
- The animal species used as the host,
- The age and general health of the animal host,
- The route of immunization,
- The dose of the immunogen,
- The immunization schedule,
- The site and method of protein conjugation (for haptens),
- The nature of the conjugate protein (for haptens).

Many animal species have been utilized for the preparation of polyclonal antibodies. The choice of animal depends on many factors, including the size of the animal, the cost and ease of maintenance, the phylogenetic similarity of the animal to the species from which the immunogen is recovered, and the absence of native materials antigenically similar to the immunogen (11). Goats are the best host for the production of large volumes of polyclonal antibodies because of the hardiness of the goat and the ease and relatively low cost of maintaining goat herds for long periods of time (11). Rabbits are an excellent source of quality antibodies, but their smaller size and higher maintenance cost are disadvantages compared with goats. Guinea pigs have been used for the production of antiinsulin antibodies because their native insulin is significantly different from human insulin, whereas fowl, fish, and reptiles produce antibodies that do not fix complement (11). The immunogen dosage and schedule of administration are important considerations in determining the quality and potency of antibody produced (11). The immunogen is commonly injected at 1- to 2-week intervals for periods of several weeks, followed by a challenge dose after a month's rest; however, each investigator uses a different schedule. If a good response is obtained, antibody titers may continue to rise with additional immunizing doses for 6 months or more. The animal is bled on a regular schedule until the antibody titer begins to drop; at that time, reimmunization is attempted. Some immunogens are given with an adjuvant (e.g., Freund's adjuvant, mineral oil) that enhances antibody production by producing an inflammatory response at the site of injection. The dose of the immunogen is the most difficult factor to control because a subliminal dose may cause tolerance, whereas an excessive dose can produce immune paralysis (11). Administration of the immunogen over a wide range (5 log or greater) may be required to produce immunization.

After immunization has taken place, the antibody concentration, reactivity, and other properties are measured (12). Undiluted serum can generally be stored at -20°C to -70°C for several years without loss of potency. However, in many cases, the immunoglobulin fraction of the serum is isolated and further purified. Purified antibody preparations can also be lyophilized and stored for years at room temperature. Immunoglobulin of the IgG class is usually preferred over IgM for immunoassays because it is more stable, easier to purify, and produced in the highest concentration (11). Antibody fragments such as Fab and $F(ab')_2$ are also gaining favor (13).

The major disadvantage of polyclonal antibodies for diagnostic testing is lot-to-lot variation in specificity and avidity, so that quality control becomes difficult. In addition, the antigen must exist in a highly purified state to generate a polyclonal antibody preparation that does not cross-react with other substances, and many antigens are poorly immunogenic or nonimmunogenic as well. Nevertheless, until recently, polyclonal antibody preparations were the only ones available, and they have been applied very successfully for diagnostic purposes, especially in radioimmunoassay (RIA) techniques and immunoperoxidase staining.

Monoclonal Antibodies

A clone of plasma cells derived from a single B lymphocyte produces immunoglobulin molecules that are specific for a certain antigen (monoclonal antibodies). Plasma cells represent the end stage of the B-cell maturation process and cannot be stimulated to divide further. Therefore, until the mid-1970s, the only source of monoclonal antibodies in large quantity was body fluids and effusions from patients with malignant plasma cell tumors. Monoclonal antibodies from this natural source were essential for the elucidation of antibody structure and function but were of no value for other research or clinical applications. However, in 1975 Kohler and Milstein (14) devised a practical method of immune manipulation that allowed the production of virtually unlimited quantities of monoclonal antibodies specific for any immunogenic substance. This discovery was one of the most important in the history of science because it has led to innumerable discoveries in the immunologic sciences, to numerous improvements in medical diagnosis, and to a large industry devoted to the commercial production of monoclonal antibodies.

Production of Monoclonal Antibodies

Kohler and Milstein (14) devised a technique to fuse nonimmortalized antibody-producing plasma cells from an immunized animal with immortalized myeloma tumor cells that produced either no antibody or antibody with an irrelevant specificity. The resulting hybridoma rapidly reproduced in culture while retaining the ability to synthesize immunoglobulin with the specificity of the original plasma cell. The hybridoma cells could be grown in culture or in the peritoneal cavity of an animal to obtain large quantities of monoclonal antibody.

Hybridoma production begins with the immunization of a mouse with a specific antigen (15). Plasma cells from the spleen of the mouse are then isolated and brought into contact with cells from a nonantibody-producing mutant murine malignant myeloma cell line that is deficient in the enzyme hypoxanthine guanine phosphoribosyl transferase (HPRT). Cellular interaction takes place in the presence of polyethylene glycol or another fusing (surface active) agent such as the Sendai virus. Under these conditions, fusion of the cell membranes of some of the cells occurs, and later the nuclear membranes fuse together as well. Many hybrid cells lose chromosomes in the process of fusion, but some retain most of the genetic information from both parents. The next step is to separate proliferating cells that produce antibody of the desired specificity from nonantibody-producing cells, from hybridomas producing specificity of undesired specificity, and from parental myeloma and plasma cells. A special culture medium containing hypoxanthine, aminopterin,

and thymidine (HAT medium) is used to separate hybridomas from the cell mixture. The plasma cells cannot divide, and the HAT medium prevents growth of the mutant myeloma cells because they are deficient in HPRT and cannot use the hypoxanthine and thymidine in the culture media. Therefore, only the hybridoma cells survive and grow in the culture. Each hybrid, antibody-producing cell is isolated and cultured, and the antibody produced is studied for its specificity. If a desired antibody is found, large quantities can be produced by growth of the hybrid cells in conventional tissue culture systems, in the peritoneal cavities of immunodeficient mice or nude rats, or in hollow fiber culture systems. The hybridoma cells continuously produce monoclonal antibody as long as they retain the necessary genes (Fig. 61.2). The recent application of genetic engineering techniques to monoclonal antibody technology has resulted in the production of recombinant antibodies, which are partly of human origin, for therapeutic purposes (16,17,18,19,20,21 and 22).

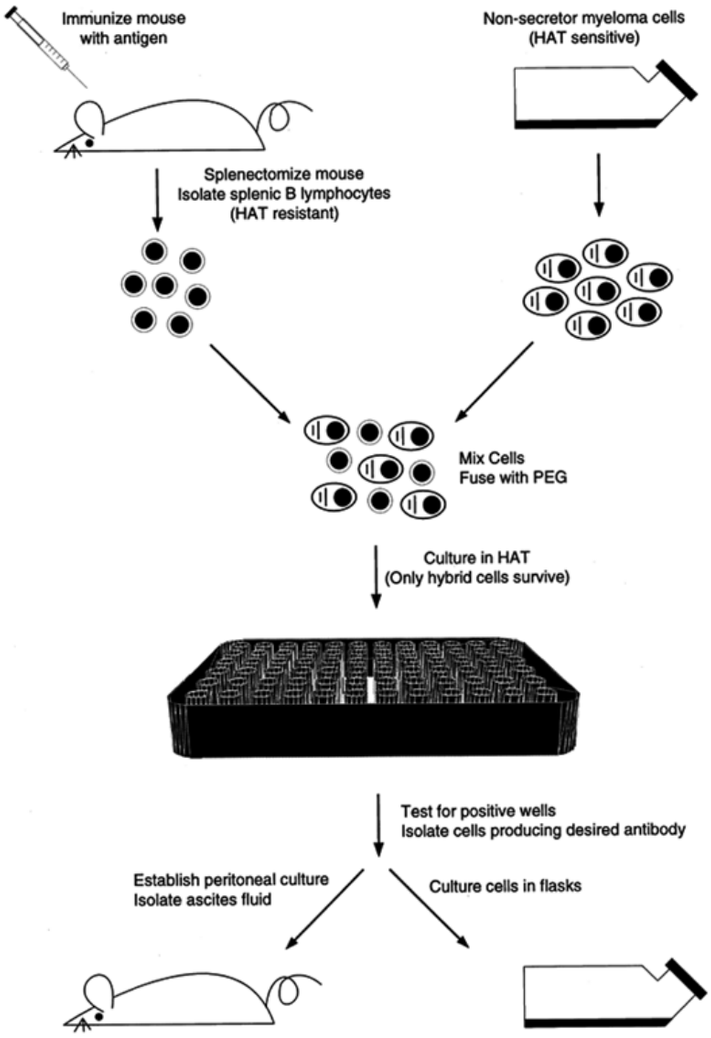


FIGURE 61.2. Production of a monoclonal antibody.

Clinical Applications of Monoclonal Antibodies

Monoclonal antibodies are highly specific for a desired antigen and can be produced in essentially unlimited quantities. In addition, monoclonal antibodies can be generated from relatively impure antigen preparation and show consistent sensitivity and specificity from lot to lot. For *in vivo* applications, even though relatively small amounts of monoclonal antibodies are utilized, they often lead to the development of heterophile (interspecies specific) antibodies in the host (23). This can lead to interferences

that are discussed later. However, the production of a monoclonal antibody with the desired specificity does not always lead to a useful product because the avidity (combining ability) of the antibody may be low, the antigen may be sparse on some cells and difficult to detect, or the monoclonal antibody may even be too specific, such as failing to detect variously glycosylated forms of human luteinizing hormone, even though they all have the same biological activity (24). For these and other reasons, polyclonal antibodies are still preferred over monoclonal antibodies for some applications.

Monoclonal antibodies have become essential reagents for many *in vivo* and *in vitro* diagnostic and therapeutic purposes (25,26,27 and 28). These applications can be classified as followed:

- *In vitro* diagnosis

  Monoclonal antibodies specific for leukocyte antigens

  Monoclonal antibodies specific for nonleukocyte antigens

- *In vivo* imaging

- Immunotherapy

  Organ transplantation

  Malignant neoplasms

Monoclonal Antibodies Specific for Leukocyte Antigens

Monoclonal antibodies specific for lymphoid antigens have been the most widely utilized of all monoclonal antibodies. Because the differentiation of T and B lymphocytes by rosetting and other techniques was tedious, time-consuming, and inaccurate, the development of a series of monoclonal antibodies specific for T-lymphocyte antigens in the late 1970s was a revolutionary event. These monoclonal antibodies (the OKT series) were rapidly applied in the diagnostic immunology laboratory for studying the phenotype of lymphoproliferative disorders and for the care of patients infected with the human immunodeficiency virus (HIV). Unfortunately, the rapid availability of numerous monoclonal antibodies, often with similar or identical reactivity, leads to much confusion among immunologists. Fortunately, the World Health Organization sponsored a series of workshops to compare and classify the reactivity of antileukocyte antibodies supplied from laboratories around the world. It was decided to group antileukocyte antibodies with similar reactivity into clusters. In addition, a cluster designation (CD) number was assigned to each of these groups.

Cellular identification and quantitation with a monoclonal antibody require that the antibody be labeled or tagged by some means so that the antibody-cell complex can be visualized. Fluorescent dyes such as fluorescein isothiocyanate (FITC) have been utilized most extensively for this purpose, and the labeled cells can be detected by flow cytometry or fluorescence microscopy. For the identification of cells in tissue sections, immunoperoxidase staining using horseradish peroxidase or another enzyme to produce a colored reaction product is used most.

The major diagnostic applications of the antileukocyte monoclonal antibodies include the diagnosis and monitoring of patients with congenital or acquired immunodeficiency disease (including HIV infection), the classification and monitoring of tumors of the hemopoietic system, and monitoring patients undergoing *in vivo* therapy with monoclonal antibodies (see below). In addition, antileukocyte monoclonal antibodies have been used as reagents for automated hematology counters and have great potential to assist in the diagnosis and monitoring of allograft rejection, graft-versus-host disease, autoimmune disease, and various infectious diseases. Therapeutically, monoclonal antibodies specific for leukocyte cell surface antigens are under evaluation for the treatment of hematologic malignancies, autoimmune disease, graft rejection, sickle cell disease, and even myocardial infarction (29,30,31,32,33,34,35,36,37,38 and 39).

Monoclonal Antibodies Specific for Nonleukocyte Antigens

The availability of monoclonal antibodies specific for a multitude of nonleukocyte antigens has changed both anatomic and clinical pathology. First, hybridoma technology has provided a means to detect substances that previously could not be evaluated. With the technique of immunoperoxidase staining and an armamentarium of monoclonal antibodies, the surgical pathologist can differentiate between lesions that were impossible to diagnose precisely only a decade ago. In addition, small numbers of human cancer cells can often be detected in body fluids and effusions with the use of monoclonal antibodies. Examples of some of the nonleukocyte monoclonal antibodies used in surgical pathology include those directed against epithelial cell antigens (epithelial membrane antigen), intermediate filaments (keratins, desmin, vimentin, neurofilament, glial fibrillary acidic protein), hormones (serotonin, synaptophysin), infectious agents (cytomegalovirus), and tumor-specific antigens (prostatic-specific antigen). The *in vivo* potential of some of these monoclonal antibody antibodies for cancer diagnosis and therapy is under evaluation (26,40,41,42 and 43).

In the clinical pathology laboratory, monoclonal antibodies have also provided a means to quantitate accurately many soluble biological products such as antibodies, hormones, enzymes, and drugs. In addition, monoclonal antibodies have improved the specificity of assays that were previously performed with polyclonal antibodies or colorimetric methods (27). Therefore, fewer false-positives occur due to inflammatory bowel disease or liver disease.

GENERAL PRINCIPLES OF IMMUNODIAGNOSIS

Part of "61 - Basic Principles of Immunodiagnosis"

Immunoglobulin Structure

Antibodies are protein molecules with a specific structure and function. Each molecule consists of two types of chains (heavy and light). The primary amino acid sequence of each chain is specified by the genetic code. Sequence analysis of the heavy and light chains from many immunoglobulin molecules revealed regions with similar (homologous) composition (constant domains), and other regions with little homology (variable domains). In addition, selective areas within the variable regions showed extreme variability (hypervariable regions) (44) (Fig. 61.3). Five immunoglobulin classes (IgA, IgG, IgM, IgE, and IgD) have been defined based on the sequence identities in the constant domains.

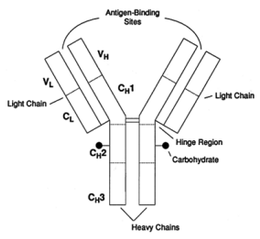


FIGURE 61.3. Basic structure of the immunoglobulin molecule.

The spatial conformation of the heavy and light chains of the immunoglobulin molecule is determined by hydrogen bonding between portions of the amino acid chains (secondary structure). Hydrogen and hydrophobic bonds between the polypeptide chains results in a complex globular configuration of the

molecule (tertiary structure), whereas associations between the heavy and light chains define the quaternary structure.

The specific biological function of each immunoglobulin class is determined by the tertiary structure of the constant domains (Fc region), whereas the antigen-combining regions (Fab regions) of the molecule are formed by spatial associations between the variable regions of the heavy and light chains. In each antigen-combining site, the hypervariable regions of one heavy chain and one light chain are brought in close proximity by tertiary folding to define an area complementary to the structure of an antigen. The number of antigen-combining sites determines the valence of the immunoglobulin molecule. IgG, IgD, and IgE exist as monomers, with a valence of 2, whereas the IgM molecule is a pentamer, with a valence of 10. Both monomeric and dimeric forms of IgA have been defined, with respective valences of 2 and 4. Binding sites for complement, the phagocyte and bacterial Fc receptor, staphylococcal protein A, and rheumatoid factor are present in the Fc receptor of the molecule (45).

Antigen-Antibody Interactions

Antigen-antibody interactions can be classified as primary, secondary, and tertiary. Primary reactions involve antigen binding and the formation of the antigen-antibody (immune) complex, secondary reactions involve interactions between antigen-antibody complexes (e.g., agglutination, precipitation, complement fixation), and tertiary reactions arise as *in vivo* consequence of the interaction of immune complexes with immune cells and other immune products (e.g., opsonization, chemotaxis).

The antigen-binding region is a shallow cleft (approximately 30Å× 14Å × 6Å) at the distal end of the Fab region. The binding of an antigen to an antibody is an event determined by the complementary geometric fit of the antigen in the binding groove, and by other factors, such as the temperature, pH, and ionic strength. The antigen-binding site accommodates approximately four to five peptides of a protein antigen or three to five residues of other types of antigens. Binding involves noncovalent forces, including hydrogen bonds, as well as electrostatic, hydrophobic, and Van der Waals forces. In chemical terms, the interaction of an antigen and antibody is a reversible reaction governed by the law of mass action, such that:

$$\text{Free Ag} + \text{Free Ab} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{Ag} \cdot \text{Ab Complex}$$

where k_1 and k_2 are the respective rates for the formation and dissociation of the antigen-antibody complex. At equilibrium, the formation of a stable immune complex determined by the ratio between the rate constants:

$$K = \frac{k_1}{k_2} = \frac{(\text{Ag} \cdot \text{Ab})}{(\text{Ag})(\text{Ab})}$$

where K is the equilibrium constant (affinity constant) of the reaction. The association constant of an antigen and antibody is also influenced by other factors, including the valence of the antibody and the valence and epitopic density of the antigen. The term avidity is used in reference to the overall stability of an antigen-antibody complex. Although the kinetics of the antigen-antibody reaction can be precisely defined for reactions involving monoclonal antibodies, polyclonal antibody preparations are more difficult to characterize because they consist of heterogeneous mixtures of antibodies of different class, valence, specificity, and affinity. In this circumstance, the average affinity (K_0) is defined by the reciprocal of the free antigen concentration when 50% of the binding sites are occupied:

$$K_0 = \frac{1}{(\text{Ag Free})}$$

Cross-reactivity occurs when antibodies react with heterologous antigens that are similar in structure but not identical to the original (homologous) antigen. Cross-reactivity is partly owing to the heterogeneous nature of the humoral immune response against an immunogenic substance and by the presence of multiple antigenic determinants on most complex natural antigens. In addition, an antibody will bind substances that are structurally similar to the homologous antigen, although usually with less affinity. The phenomenon of cross-reactivity is an important practical consideration in the design and clinical utilization of immunoassays because the sensitivity of the assay can be reduced, false interpretations can be made, and inappropriate therapy initiated.

Several different consequences can result from the interaction of antigen-antibody complexes with each other and with other immune substances such as complement. These secondary immune reactions are influenced by the nature, valence, and concentration of the antigen and by the valence, concentration, avidity, and immunoglobulin class of the antibody. Generally, multivalent antigens and antibodies are required for the induction of a secondary immune reaction. In the past, detection of the secondary immune reaction was the only means available for the detection and analysis of immune components. As a result, antisera were often classified by their ability to cause precipitation (precipitins), the hemolysis of red blood cells (RBCs) (hemolysins), the agglutination or precipitation of particulate antigens (agglutinins, precipitins), and other reactions. In addition,

the suffix *-ogen* was attached to an antigen that elicited a secondary immune response (e.g., agglutinogen, agglutinogen, precipitogen). The *in vivo* or tertiary consequences of the immune reaction are complex and generally involve the interaction of antigen-antibody complexes with cells and other immune products.

Clinical Utilization of the Antigen-Antibody Interaction

The humoral and cellular products of the immune system have achieved widespread utilization as reagents for medical diagnosis. In this regard, the unique specificity of the antibody has proven invaluable for the detection and quantitation of immune-related substances (other antibodies, immune products, and cells) as well as a wide variety of substances unrelated to the immune system (e.g., hormones, drugs, tumor antigens, RBC antigens). Therefore, antibodies and other immune products are standard reagents not only in the clinical immunology laboratory but also in surgical pathology, immunohematology, hematology, microbiology, clinical chemistry, and toxicology.

Classification of Immunoassays

Immunoassays are studies of the antigen-antibody reaction performed to detect the presence of an antigen, antibody, or immune complex (qualitative immunoassay) or to measure the concentration of one of the components of an antigen-antibody or hapten-antibody reaction (quantitative immunoassay).

Practical Considerations in Immunodiagnosis

The diagnostic relevancy of an immunoassay depends on the choice of the proper assay, adequate specimen collection and transportation, prompt and accurate laboratory analysis, and rapid reporting of the results of the assay in a format that is relevant to the clinician. For substances that are normally present in relatively large concentrations in the body (e.g., many serum proteins, some antibodies) or that undergo large changes in concentration in response to a disease process, relatively insensitive methods such as immunonephelometry or immunoturbidimetry may suffice for diagnosis. However, for the detection of small but diagnostically relevant changes in the concentration of an enzyme or to quantitate hormones, tumor markers, therapeutic drug levels, immunoregulatory substances, and other substances that are present in minute quantities, the most sensitive immunoassays available may be required.

In practice, an immunoassay is performed in a system of defined volume under controlled reaction conditions and with precisely measured volumes of test specimen and reagents. Under these conditions, the signal strength of the measured parameter is either directly (noncompetitive assay) or inversely (competitive assay) proportional to the concentration of the measured analyte. Dilutions of a purified preparation of the test substance are used to prepare a standard curve, and the concentration of the variable in an unknown specimen is determined from the standard curve. The results of these assays may be reported as quantitative, semiquantitative, or dichotomous (binary, positive/negative). In many immunoassay systems, especially for the presence of serum antibodies, serial dilutions of the specimen are tested, and the end result is taken as the highest specimen dilution that produces arbitrary defined signal strength (titer). In these circumstances, it is common to first assay a single fixed dilution of the specimen. If the signal strength of the screening assay is below a predefined limit, the assay is reported as negative and no additional testing is performed. However, if the signal strength of the screening assay is above the predefined limit (positive), additional dilutions are tested until the minimal signal strength is reached. The serum titer required to produce this result is reported as the end point.

Ultimately, immunoassay results are often used to determine whether an individual patient is diseased or nondiseased. Unfortunately, in any large population, there are diseased individuals with normal assay results (false-negative results), and nondiseased people with abnormal values (false-positive results). Causes of false-positive and false-negative immunoassays include cross-reactivity of the test reagents with an unrelated disease product, the presence of substances that interfere with the formation of the antigen-antibody complex, errors in test performance, and other factors. Under these circumstances, the correct clinical interpretation of the results of an immunoassay requires up-to-date information about the overall performance of the immunoassay procedure under consideration, such as that required by Clinical Laboratory Improvement Amendments 1988 (accuracy, precision, analytical sensitivity, analytical specificity, reportable range, reference interval, linear range) as well as knowledge of disease pathogenesis and the immune response to the disease process.

NONCELLULAR IMMUNOASSAYS

Part of "61 - Basic Principles of Immunodiagnosis"

Agglutination

Agglutination is a secondary immune phenomenon that occurs when insoluble or particulate antigens (cells or other particles) are cross-linked by an immune reaction (Fig. 61.4).

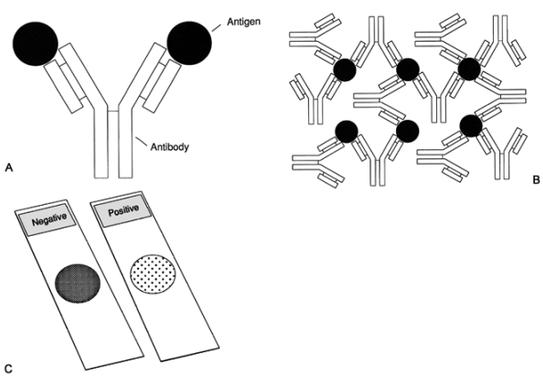


FIGURE 61.4. The agglutination reaction. Schematic illustration of the interaction of an antibody with an insoluble antigen, resulting in agglutination. **A:** The formation of antigen-antibody complexes. **B:** Cross-linking of the antigen-antibody complexes to form a lattice. **C:** Schematic illustration of the visual appearances of positive and negative agglutination reactions. The appearance of the reaction is greatly influenced by whether the assay is performed on a flat surface or in a vessel, such as a test tube or microtiter plate. The physical characteristics of the vessel (round bottom or V bottom) are also important.

Agglutination occurs because antibodies have two or more antigen recognition sites (bi- or multivalency). If multiple antigenic recognition sites are present on a particle, lattices can be formed that grow in size and eventually become a mass that is macroscopically visible. The major factors affecting the agglutination reaction include the class, affinity and avidity of the antibody, the proximity and number of binding sites on the particle, the relative concentrations of antibody and particles, electrostatic interactions (zeta potential), and the viscosity of the medium. Antibodies of the IgM class, with 10 antigen combining sites, are usually the best agglutinins and are more efficient than IgG in agglutination.

Agglutination assays are most often utilized for the detection of antibodies in serum or other body fluids but can also be used to detect soluble antigens. Agglutination assays are classified as direct or indirect, depending on whether the analyte is present in its native state or linked to a particle (carrier) to allow detection of the antigen-antibody reaction. Carriers vary in size from approximately 0.05 μ to 7 μ and may be RBCs, latex particles, liposomes, microcapsules, or other particles (46).

Because particles in suspension often possess negative charges and repel each other, it may be necessary to overcome the zeta potential by the addition of protein or inorganic salts to permit particle cross-linking. Antibody excess may prevent agglutination by coating all the binding sites, whereas inadequate antibody may not provide sufficient cross-linking for agglutination. In practice, this "prozone" phenomenon may require adjustment of particle or antibody concentration or performing an assay at a number of different concentrations. Interference with agglutination by any of these factors after a primary immune reaction is termed incomplete agglutination.

Direct Agglutination

Direct agglutination occurs when a suspension of particles is mixed with a specific antibody solution. If specific antibody with two or more antigen recognition sites is present, the particles may rapidly link together to form visible clumps (46) (Fig. 61.5). The classic example of the application of direct agglutination reaction is the Widal test for the diagnosis of typhoid fever. In this test, specific anti-*Salmonella* antibodies are detected by adding a constant amount of a *S. typhi* cell suspension to serum that has been serially diluted. After appropriate incubation, the tubes are examined for visible agglutination. The highest dilution of serum showing agglutination is the titer of the antibody.

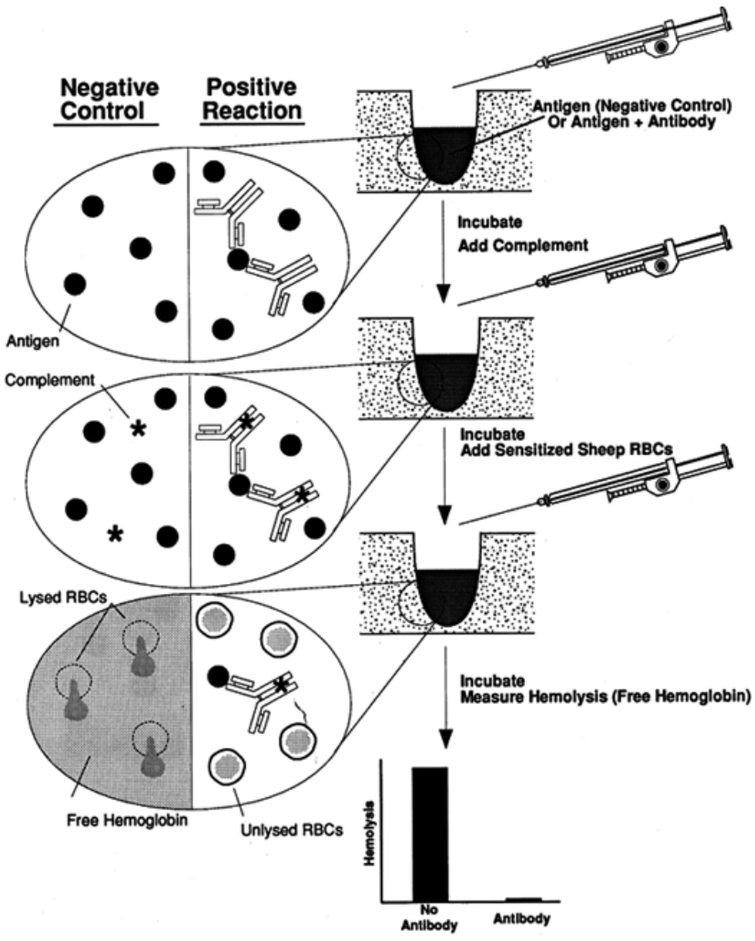


FIGURE 61.5. Principle of the complement-fixation assay. In this assay, antigen and antibody are permitted to interact in solution, and complement is then added. In the presence of an antigen-antibody complex, complement binds to the Fc portion of the antibody molecule. If there is no antigen-antibody complex, complement exists free and unbound in the reaction solution. Sensitized sheep red blood cells are added as an indicator for the presence of unbound complement. The sheep red blood cells are lysed only if there is free complement and an absence of antigen-antibody-complement complexes.

The agglutination of RBCs by antibodies directed against natural components of the cell surface is a type of direct agglutination termed hemagglutination. In immunohematology, direct hemagglutination is the principal method for the detection of antibodies specific for erythrocytic antigens.

Interspecies hemagglutinins (heterophile antibodies) can lead to false-positive results in agglutination tests that utilize nonhuman erythrocytes. Heterophile antibodies can be produced in the sera of some normal individuals as well as patients with serum sickness, autoimmune diseases, and some infectious diseases. However, in infectious mononucleosis the presence of sheep RBC (SRBC) agglutinins in high titer is highly characteristic of the disease.

Indirect Agglutination

Indirect agglutination techniques utilize particles as passive carriers of an antigen (direct passive agglutination) or antibody (reverse passive agglutination) (46). Although coated RBCs (passive hemagglutination) were initially utilized in the clinical applications of the indirect agglutination principle (47), inert particles (charcoal, bentonite clay, inorganic colloidal particles) have proven more versatile for many applications (Fig. 61.6). More recently, the use of coated latex microspheres, gelatin particles, colloidal gold, dye particles, and other substances has permitted the development of simple, cost-efficient assays for a wide range of clinically important substances. The sensitivity of these assays does not approach that of enzyme immunoassay (EIA), RIA, and other techniques but is adequate for many purposes. Particle inhibition

tests have also been described in which the presence of an antigen in the serum prevents the attachment of an antibody to an antigen on the surface of a particle. The term particle immunoassay has been applied to agglutination tests in which particles are used as a label for an antigen or antibody (48).

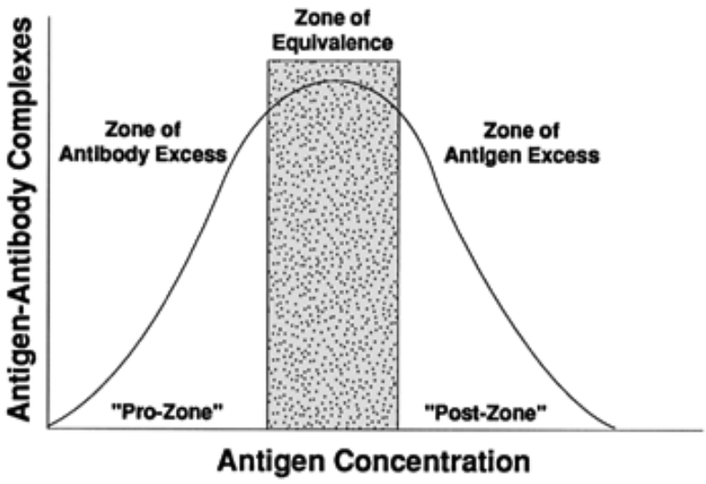


FIGURE 61.6. The precipitation curve and a schematic depiction of the phenomenon of immunoprecipitation.

The coating or sensitization of a particle for use in an immunoassay is accomplished by three techniques. Untreated erythrocytes spontaneously adsorb many proteins to their surface (thyroglobulin and other hormones, purified protein derivative, serum proteins, bacterial products). Other substances are adsorbed after treatment of the erythrocyte with tannic acid (“tanned” RBCs) or enzymes (49). Covalent attachment has been used for substances that cannot be adsorbed by these techniques. Sensitized erythrocytes are usually fixed with formaldehyde, glutaraldehyde, or another compound to increase the shelf life. The use of latex particles for immunoassay was first reported by Singer and Plotz (50) in 1956. Latex particles are usually coated by passive means, with the quantity of the adsorbed protein adjusted to provide agglutination of the analyte in its biological range. In addition, the use of latex particles avoids much of the variability encountered with RBCs. Even so, the prozone phenomenon can still be significant, and careful adherence to the manufacturer's instructions is necessary during the performance of clinical assays utilizing coated microspheres.

Examples of substances routinely detected by passive agglutination include rheumatoid factor, antinuclear antibodies, C-reactive protein, human chorionic gonadotrophin (hCG), parasitic, fungal, and bacterial antigens, and antibodies specific for

some infectious agents. Special variations of the passive agglutination technique include microhemagglutination, the Rose-Waaler test, the VDRL and rapid plasma reagin (RPR) test, and the Paul-Bunnell test.

Microhemagglutination techniques are used in a popular confirmatory test for syphilis [microhemagglutination assay—*Treponema pallidum* (MHA-TP)], and in the detection of antithyroid antibodies specific for microsomal and thyroglobulin antigens. These assays utilize small quantities of tanned, coated RBCs that are reacted with the test serum in a special plastic microtiter tray. The reaction is read with a magnifying mirror or microscope with a low-power objective. Agglutination results in the formation of a diffuse matrix coating the bottom of the tube, whereas nonagglutinated cells settle to form a compact button at the bottom of the tube. The Rose-Waaler test is an assay for rheumatoid factor that utilizes SRBCs sensitized with subagglutinating amounts of rabbit anti-sheep erythrocyte IgG. Agglutination results when these cells are mixed with serum containing anti-IgG (rheumatoid factor). The VDRL and RPR are tests for antisyphilitic antibodies that utilize the natural cross-reactivity of these antibodies with a lipid (cardiolipin) found in high concentrations in neural tissue and cardiac muscle. In the VDRL, particles of purified cardiolipin are mixed with a dilution of the test specimen [serum or cerebrospinal fluid (CSF)] under defined conditions and examined for macroscopically visible aggregates. The RPR test is an improved modification of the VDRL test that utilizes charcoal particles coated with a stabilized mixture of cholesterol and cardiolipin. Serum dilutions are used to titer positive specimens. However, all states have enacted laws requiring confirmation of a positive RPR or VDRL test by a confirmatory test utilizing antigens isolated from *T. pallidum* (MHA-TP or fluorescent treponemal antibody-absorption test).

Several techniques have been used to increase the sensitivity of agglutination assays. Immunoassay by particle counting has a sensitivity of 10 to 1,000 times that of slide-based latex agglutination assays (46). In this technique, a standard number of coated latex particles are incubated with the specimen and the reaction mixture is passed through an automated particle counter. Agglutinated and unagglutinated particles are differentiated based on light scatter and the number of unagglutinated particles is used to construct a dose-response curve. The concentration of an unknown specimen is determined from the dose-response curve. Although these assays are rapid and fully automated and do not require a separation step, they require expensive equipment and do not have the sensitivity of RIA or EIA. Latex agglutination on microtiter plates has also been described, with the quantification of agglutination by turbidimetric measurement (latex turbidimetric assay) (46,51).

Agglutination Inhibition

Agglutination inhibition reactions are based on the competition for antibody-binding sites between haptens (i.e., antigens) in solution and the same or similar haptens localized on sensitized indicator cells or particles. In the presence of a limiting amount of antibody, agglutination is inhibited by hapten present in the solution (46). Therefore, lack of agglutination is indicative of a positive test. These assays are normally performed by reacting the test solution with an antibody preparation, followed by reaction with the indicator cells or particles. Antigen in the test solution, if present, reacts with the antibody and subsequently prevents or decreases agglutination of the indicator. Agglutination inhibition assays are more sensitive than passive agglutination assays but are more difficult to develop and control.

Agglutination inhibition is widely utilized for the detection of hCG in the urine. In this assay, test urine is reacted with anti-hCG antibody and then with hCG-coated latex particles. Agglutination inhibition has also been used for the detection of drugs of abuse in the urine and antiplatelet glycophorin and antimicrobial antibodies in the serum (52,53,54,55,56 and 57). An adaptation of agglutination inhibition, turbidimetric latex agglutination inhibition, utilizes spectrophotometric measurement of the turbidity of the agglutination suspension as a detection method. This technique was recently adapted for automated therapeutic drug monitoring (58,59).

Cell Lysis

An antigen antibody reaction on a cell membrane can result in lysis of the cell if complement is present. If the cells are RBCs, hemolysis takes place. Cell lysis forms the base for the complement-fixation test, which is a sensitive method for the detection and quantitation of antigens or antibodies. The serum specimen is first heated to 56°C to inactivate native complement. Measured amounts of antigen and complements are then added. If antibody specific for the added antigen is present in the serum, antigen-antibody complexes will form, and these will bind or fix all the complement. The reaction mixture is the checked for the presence of free complement. The indicator system for complement consists of SRBCs plus an antibody specific for SRBCs. If all the complement has been fixed, none will be free to lyse the SRBCs and the test is positive. If no antibody was present in the patient's serum, then the complement is not fixed and is free to interact with the indicator system and lyse the red cells. In this case, the test is negative (Fig. 61.5). Properly conducted complement-fixation tests require the use of appropriate controls to ensure that the results will not be adversely affected by the presence of anticomplement compounds, such as denatured immunoglobulins,

heparin, chelating agents, or microbial contaminants. The complement-fixation test was widely used in the past to measure antibodies against viruses and other microbial agents but has recently been replaced by EIA to avoid anticomplement interference and other problems (60).

A different procedure using lytic reactions is the viral neutralization test. This method is based on the fact that some viruses, such as herpes simplex, produce effects (cytopathic effects) when added to target cells growing in tissue cultures. This phenomenon can be used to search for virus-neutralizing antibodies in a serum sample. This is done by adding the serum suspected of containing the antibody to a virus suspension and then adding this mixture to a susceptible cell culture. If the cells fail to show any change, then antibodies present in the serum sample neutralized the cytopathic effect of the virus. If cytopathologic effects develop, then neutralizing antibodies were not present.

Immunoprecipitation

Immunoprecipitation is a secondary immunologic phenomenon that results from the reaction of an antibody and soluble antigen in the proper proportions, with the formation of an insoluble immune complex. Immunoprecipitation has been used since the 1920s for the quantitation of antigens and antibodies.

Basic Principles of Immunoprecipitation

Heidelberger and Kendall first described immunoprecipitation as a manifestation of the antigen-antibody reaction, in 1935. Immunoprecipitation requires the presence of bi- or multivalent antibodies and an antigen with at least two antigenic determinants per molecule. Cross-linking of antigen molecules results in the formation of a lattice that gradually grows in size. Eventually, the lattice may become insoluble and precipitate from the solution. If immunoprecipitation is studied as a function of antigen concentration, with the same concentration of antibody, three phases can be delineated. During the initial phase of antibody excess (prozone), the number of antigenic determinants is small relative to the antibody concentration, and lattice formation is prohibited. As the amount of antigen is increased, the zone of equivalence is reached, where the reactants are at their optimal proportions and maximal precipitation occurs. At this point, a hydrophobic and ionic interaction between smaller complexes increases the size of the aggregates. The final phase is one of antigen excess (postzone), in which all the antibody-binding sites are saturated, cross-linking of antibody molecules cannot occur, and lattice formation is prohibited (Fig. 61.6).

The time course of an immunoprecipitation reaction can vary from hours to days, depending on the affinity, avidity, and valence of the antibody; the size and number of antigenic determinants of the antigen; the reaction temperature, ionic strength, and viscosity of the medium; the interactions between the medium and reactants; and other factors. The time course can be greatly accelerated by assisting the migration of the reactants through the use of an electric field. However, immunoprecipitation reactions generally lack the speed and sensitivity of many other immunoassay techniques, and the use of this type of analysis in the clinical laboratory is declining.

Immunoprecipitation reactions can be classified according to whether the interaction of the antigen and antibody is allowed to occur passively or actively (under the influence of an electric field) (61). Passive immunoprecipitation can be further classified according to the nature of the reaction medium (fluid or gel) (Table 61.1).

TABLE 61.1. CLASSIFICATION OF IMMUNOPRECIPITATION ASSAYS

Passive immunodiffusion assays
Ouchterlony double diffusion
Radial immunodiffusion
Active immunodiffusion assays
Countercurrent immunoelectrophoresis
Rocket immunoelectrophoresis
Combined immunodiffusion assays
Immunoelectrophoresis
Immunofixation electrophoresis

Support media composed of a substance extracted from seaweed (agarose) have been found to provide much better resolution than the gels used in the past. In passive immunodiffusion techniques utilizing gels, antigen and antibody solutions are usually placed in separate regions of the gel and the molecules diffuse radially outward through pores in the gel. When equilibrium is reached, a concentration gradient is formed for each reactant. The concentration is highest near the well and decreases geometrically with distance from the well. At the point in the region between the reactants where their concentrations fall into the zone of equivalence, an immunoprecipitate appears as a narrow band (precipitin band). Multiple precipitin lines are formed if more than a single antigen and/or antibody with different rates of diffusion is present. An alternative method of passive immunoprecipitation uses gels in which one of the reactants (antibody or antigen) is dissolved during the preparation of the gel. The rate of diffusion depends on the molecular weight and size of the antigen(s) and antibody(s); the pore size; degree of hydration and other properties of the support medium; the ionic strength and composition of buffer; and the interaction between the reactants and the support medium. In the preparation of a gel for passive immunoprecipitation, agarose (usually 1% wt/vol) is completely dissolved in distilled water or saline by heating with frequent stirring at temperatures of 100° to 110°C. If antisera are added, the solution is first cooled to approximately 50°C to prevent protein denaturation (61). Gels are cast in Petri dishes or on glass plates or Mylar sheets using gel-molding frames. Wells are added at the time the gel is cast, or at a later time, using a variety of templates. Precipitin lines may be directly visible if the concentration of antigen and antibody was high enough. Alternately, it may be necessary to stain the gel with dyes such as Coomassie brilliant blue R-250 or amido black to visualize the faint precipitin lines or clarify the pattern of precipitation. Gels may be stored in a preservative, or dried and stored.

Passive Immunodiffusion Immunoassays

Ouchterlony Double Diffusion

The two basic passive immunodiffusion techniques are double (Ouchterlony) immunodiffusion and single diffusion radial immunodiffusion

(RID). In the Ouchterlony procedure, antigen and antibody preparations are placed in separate wells cut in a thin layer of agar in a Petri dish. Typically, an antibody solution is placed in a center well, and antigen controls and test solutions are placed in wells around the outside of the gel. The reactants diffuse through the agar until they meet at optimal proportion, where bands of precipitation are formed. The antigenic relationships of the reactants are reflected in the precipitation pattern. The three basic patterns of reaction are identity, nonidentity, and partial identity. In reactions of identity, the two precipitation bands merge into a solid line. In reactions of nonidentity, the lines of precipitation cross. Partial identity is indicated by spur formation, indicating that one of the antigens is cross-reactive with, but not identical to, the other (Fig. 61.7).

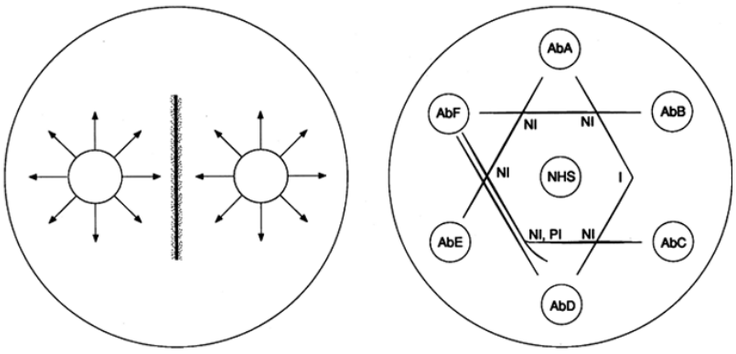


FIGURE 61.7. Principle of double (Ouchterlony) immunodiffusion. In this adaptation of the immunoprecipitation reaction, antigen and antibody are placed in wells punched in agarose gel. The reactants diffuse through the gel, and a visible line forms at the point of equilibrium where an antigen-antibody reaction occurs. In this example, normal human serum was placed in the center well, and various antibody solutions in the six surrounding wells. If an antigen-antibody reaction occurs, the identity (I), nonidentity (NI), or partial identity (PI) of the reactants can be ascertained by the characteristics of the precipitin line. For example, the cross-lines (NI) indicate that antibody A (AbA) and antibody B (AbB) react with different antigenic determinants, whereas antibody B and antibody C (AbC) react with the same antigenic determinant (reaction of identity). Antibodies D (AbD) and antibody E (AbE) react with the same antigenic determinant, but the spur indicates the presence of an additional antigenic determinant recognized by AbD but not AbE. The double precipitin line indicates the presence of an antigen recognized by AbE but not by AbD or AbF.

Immunodiffusion assays are simple and inexpensive to perform and do not require elaborate equipment. However, the precipitin reactions require a day or more for completion, the results are qualitative and not quantitative, and the reaction patterns may be time-consuming and difficult to interpret. For these reasons, more rapid, quantitative methods of immunochemical analysis have replaced double immunodiffusion for most high-volume clinical applications. One exception is in the identification of patients with some autoimmune antibodies, such as those directed against the Smith (Sm) antigen and ribonucleoprotein (RNP). Problems can arise from the presence of irregular wells, overfilling of the wells, or incubation on a nonlevel surface or at an improper temperature. False-negative results can theoretically result from antigen or antibody excess, and the gels and antigen or antibody reagents may have a limited shelf life. Other potential problems include bacterial or fungal contamination and drying of the gels during incubation (61). Other antibody detection methods, such as immunoblot and EIA, are replacing immunodiffusion in the diagnostic laboratory because they avoid these problems and show improved sensitivity and specificity (62).

Radial Immunodiffusion

RID is a gel-based quantitative immunoprecipitation method. In this procedure, a monospecific antibody with high affinity and excellent precipitating ability is added to an agarose solution before gels are cast. Wells are punched into the gel, antigen solution is added to the wells, and the gels are incubated. Under these circumstances, the antigen diffuses radially outward, resulting in a ring of precipitation around each well at the zone of equivalence (63). The diameter of the ring is proportional to the concentration of antigen in the well (Fig. 61.8).

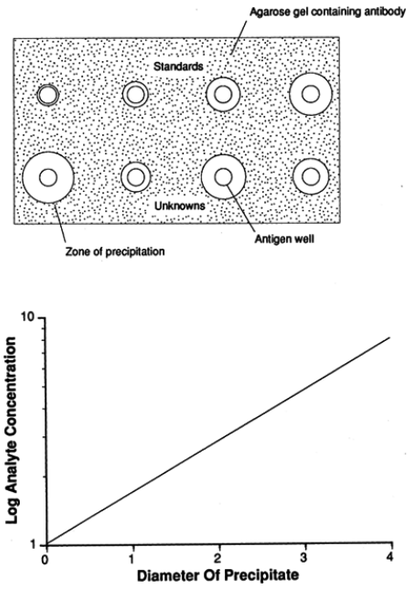


FIGURE 61.8. Radial immunodiffusion.

The molecular weight of the antigen is the major determinant of the time required for performance of a RID assay. Generally, this varies from 24 hours for small proteins like α_1-antitrypsin to as long as 72 hours for large molecular weight proteins such as IgM. In practice, at least three antigen solutions of known concentrations (standards) are included with each assay, and their ring diameter is used to construct a standard curve, from which the concentrations of unknown solutions are determined. Ring diameters can be measured using a ruler and magnifying glass, although more accurate measurements require a special reading device. To compensate for irregularities in the wells, it is customary to measure each ring diameter twice at right angles and to calculate the average of the measurement (44). Measurements

of ring diameter may be taken at the end point of the reaction (Mancini or end-point diffusion method), or at a fixed time before the end point (usually 18 hours) (Fahey or kinetic diffusion method). In the end-point diffusion method, the standard curve is prepared on linear graph paper by plotting the concentration of antigen on the *x* axis and the diameter squared of the precipitin ring diameter on the *y* axis. If the standard curve is nonlinear, a line of best fit is drawn. The antigen concentration of an unknown solution is usually determined directly from the standard curve, although mathematical interpolation can also be used. In the less precise kinetic diffusion method, precipitin ring diameter is plotted on the *x* axis of semilogarithmic graph paper and the concentration of the reference solutions is plotted on the *y* axis. Standard curves are not extended beyond the range of the standards. Specimens whose precipitin ring diameters exceed those of the reference solutions are diluted and reassayed, whereas specimens with precipitin rings smaller than the most dilute standard are reassayed with a more sensitive system or reported as being less than the value of the reference standard. Potential causes of inaccurate results include overfilling of the wells, damage to the wells during pipetting, inappropriate incubation time or temperature, and errors in reading precipitin diameters, preparing the standard cure, and determining unknown concentrations.

Active Immunodiffusion Immunoassays

The movement of a substance through a gel is greatly accelerated by an electric field. Active methods of immunodiffusion use this modification to decrease the time required for the formation of an immune complex. In addition to the properties of gels described previously, reverse buffer flow or electroendosmosis (EEO) is important in the consideration of gel electrophoresis. Gels with high reverse buffer flow are used in applications in which immunoglobulins move cathodically [i.e., countercurrent immunoelectrophoresis (CIEP)], whereas medium- or low-EEO gels are used for other applications (61).

Countercurrent Immunoelectrophoresis

In CIEP (electroimmunodiffusion), antigen and antibody solutions are located in parallel wells placed on opposite sides of an agarose gel. In the presence of an electric field at pH 8.6 (cathode on the side of antigen wells, anode on the side of antibody wells), the antibody molecules migrate toward the cathode, and the antigen toward the anode. After electrophoresis, precipitin lines will form between the wells at the zone of equivalence (61). The position of the precipitin line relative to the wells is dependent on the relative antigen and antibody concentrations, their ability to move through the pores of the gel, and the migration of these substances in an electric field. If the concentration of antibody is maintained constant, CIEP becomes semiquantitative because the migration distance of the precipitin line from the antigen well increases with the ratio of antigen/antibody concentration. In practice, serial dilutions of antigen are used as a reference for the concentration of an unknown solution. CIEP was first used clinically for the detection of hepatitis B antigen and has since been utilized for the detection of a variety of other microbial antigens, as well as autoantibodies and antibodies directed at infectious agents (i.e., influenza type A virus, adenovirus, cytomegalovirus, hydatid antigen, and *Haemophilus influenzae*). A recent study confirmed the diagnostic value of CIEP for the rapid diagnosis of typhoid fever (64).

Rocket IEP

Rocket IEP (electroimmunoassay) is a technique in which an electric force is used to drive an antigen into a gel containing an antibody. Immune complex formation results in the formation of a triangular or rocket-shaped precipitate (65). The total area, or even the height of the immunoprecipitate, is proportional to the concentration of antigen. Low-EEO agarose is utilized in rocket IEP, with the goal of minimizing migration of the antibody in the electric field. Because antigens may migrate either anodally or cathodally in the electric field, the wells must be placed to maximize use of the gel. Because the majority of plasma proteins migrate anodally in such a system, wells usually are placed near the cathode. Technical skill is required to perform rocket IEP because the samples must be applied to the gel rapidly to avoid lateral diffusion of the antigens, widening of the immunoprecipitin arcs, or even mixing of antigens in adjacent wells (61). A low-voltage field across the gel during sample application is used to minimize application artifacts. The voltage is then increased and maintained until the end of the reaction (usually several

hours). After electrophoresis, the gel is washed and stained, and the distance from the middle of each well to the top of the corresponding peak is measured with a precision of 0.1 mm. Standards spanning a wide range of concentrations are assayed with the test specimens, and a curve is plotted from the peak heights of the standard (*x* axis, peak height; *y* axis, concentration of standard). The concentration of the test specimens can be extrapolated from the standard curve, or calculated mathematically (Fig. 61.9). Rocket IEP is a cost-effective method of quantitative analysis that is frequently used in research and test development laboratories. In the clinical laboratory, its use is restricted to the analysis of substances in which high-volume methods of quantitation are not feasible, such as some clotting factors and α-fetoprotein.

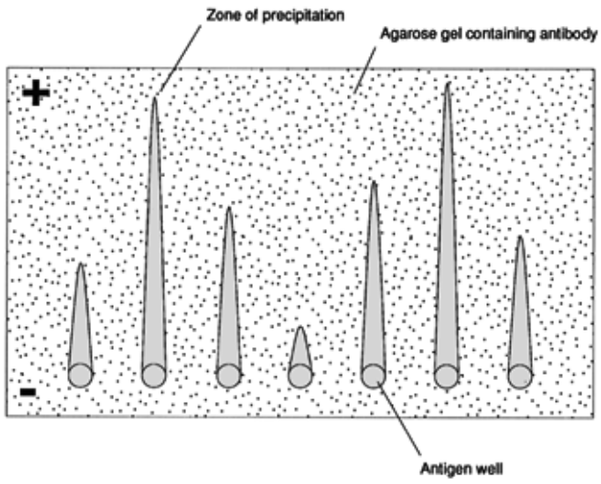


FIGURE 61.9. Principle of rocket immunoelectrophoresis.

Combined Immunodiffusion Techniques

Techniques in which active and passive immunodiffusion are combined [IEP, immunofixation electrophoresis (IFE)] are useful in resolving complex mixtures of antigens.

Immunoelectrophoresis

IEP is a technique for the separation and analysis of serum proteins that was first introduced by Grabber and Williams in 1953 and modified by Scheidegger in 1955. IEP is a two-stage procedure in which proteins are first separated electrophoretically and then identified by immunoprecipitation (66) (Fig. 61.10). IEP is performed as follows:

- The test solutions (e.g., serum, urine, CSF) are pipetted into cylindrical wells in a buffered agarose gel.
- Electrophoretic separation of the proteins is performed.
- Antisera are placed in troughs that run parallel to the path of electrophoretic migration and span most of the width of the gel.
- The gel is placed into a humidity chamber for 10 to 15 hours. During this time, proteins and antibodies diffuse into the gel and precipitin arcs form in the equivalence region between each well and trough.
- The gel is stained, dried, and then visually inspected and interpreted.

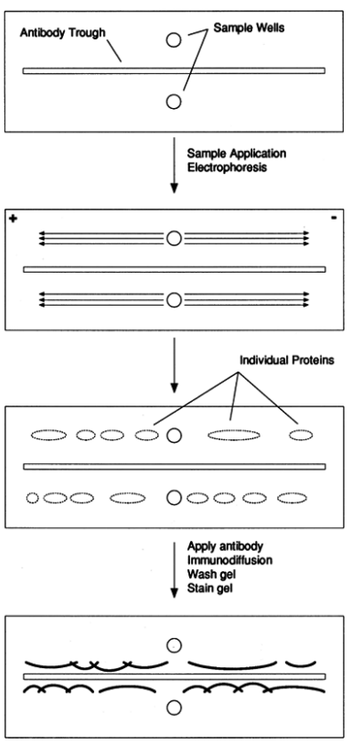


FIGURE 61.10. Theory and technique of immunoelectrophoresis. In immunoelectrophoresis, electrophoresis and immunodiffusion are sequentially applied to separate and identify protein antigens. The antigen is first applied to wells in an agarose or cellulose acetate gel, followed by electrophoresis. After electrophoretic separation of the proteins, an antibody solution is placed in the center trough and immunodiffusion is allowed to occur. Arcs of immunoprecipitation indicate an antigen-antibody reaction. The gel is then washed and stained with a protein-specific dye to clarify the reaction pattern. In clinical practice, antibodies against human immunoglobulin heavy- and light-chain immunoglobulins are commonly used with immunoelectrophoresis to identify abnormal proteins.

During IEP, a separate precipitin arc is formed for each antigenic component against which a corresponding antibody is present. Patterns of identity, nonidentity, and partial identity are formed that are identical to those formed during Ouchterlony

immunodiffusion and other immunodiffusion techniques. Although a complex pattern of arcs results if a polyspecific antiserum is utilized, monospecific antisera are usually employed for the identification of individual serum proteins.

The major advantages of IEP include small sample requirements, sensitivity, and specificity. However, the technique is time-consuming and labor-intense and requires skill and experience to perform and interpret. In addition, the results are only semiquantitative and are greatly affected by variations in the antigen/antibody ratio; antibody titer; and the specificity, avidity, and affinity of the antisera used. Problems of antigen or antibody excess are most frequently encountered. Under conditions of antigen excess, the precipitin arc moves closer to the antibody trough and becomes thickened and elongated. The precipitin arc often has a diffuse appearance on the trough side and can even disappear in situations of extreme antigen excess, in which the precipitin reaction is in the prozone. With antibody excess, the precipitin arc moves closer to the antigen well and assumes an increasingly diffuse appearance. A similar problem of indistinct precipitin arcs is encountered with a low antigen concentration and in the presence of a weak antigen-antibody reaction.

Immunofixation Electrophoresis

IFE was introduced by Alper and Johnson (67) for the identification of genetic variants of proteins, and was subsequently modified by Arnaud et al. (68) and Ritchie and Smith (69). However, because IFE has many advantages over IEP for the identification of monoclonal proteins, it has achieved increasing utilization in the clinical laboratory in recent years (71,72).

IFE, like IEP, is a two-step procedure requiring electrophoretic separation of proteins followed by immunoprecipitation. However, in IFE, the antibody is applied directly to the surface of the gel over the electrophoretically separated proteins rather than in a trough parallel to the axis of electrophoretic migration (Fig. 61.11). The technique is performed as follows:

- Samples are applied to agarose gel and separated electrophoretically, using conventional techniques. Protein solutions are diluted to a concentration of 50 to 100 mg/dL before application. During application, each specimen must be clearly separated from the adjoining specimens to avoid cross-contamination, and care must be taken in placing the electrodes so that the path of migration is parallel to the short axis of the gel. The gel must not be fixed after electrophoresis.
- Filter paper or cellulose acetate strips are cut to appropriate size, saturated with a dilute antibody solution, drained, blotted, and carefully applied over the gel. Separate strips are used for each sample. The strips must be applied without leaving air bubbles underneath and must be parallel to the path of electrophoretic migration.
- The gels are placed in a humidity chamber for approximately 1 hour. During this time, the antibody solution on each strip diffuses into the underlying gel, and immunoprecipitation occurs in the gel.
- The strips are removed, and the gel is pressed, washed, stained, and dried. Each immunoprecipitate appears as a band of varying width. Diffusion of antibody onto the application strips also occurs, so that the application strips can be stained.

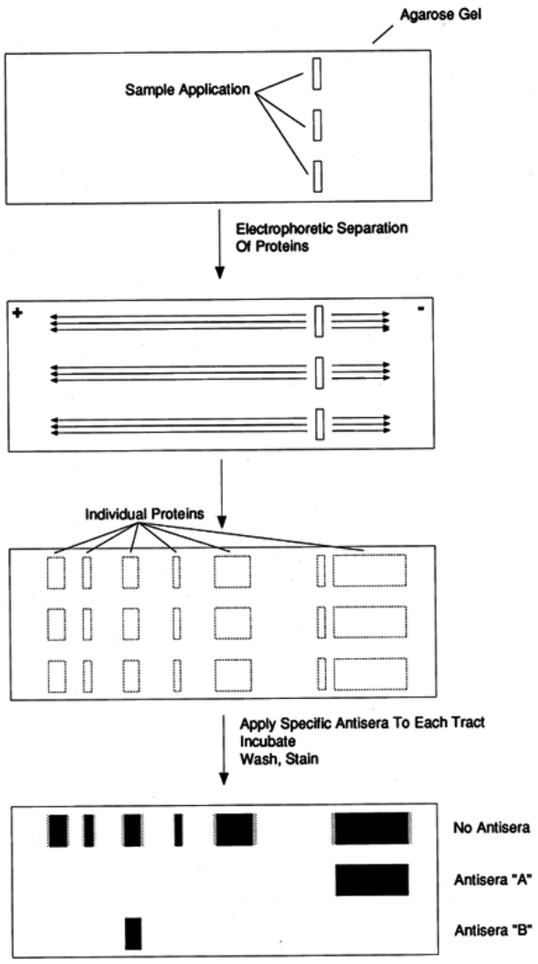


FIGURE 61.11. Technique of immunofixation electrophoresis. In this technique, strips of filter paper or cellulose acetate are soaked with antisera and placed on the surface of an agarose gel containing proteins first separated by electrophoresis. Immunodiffusion of the antisera into the agarose gel occurs, and a band of immunoprecipitation occurs at the site of each antigen-antibody reaction. The gel is washed and then stained with a protein-specific dye.

IFE has many advantages over IEP and is rapidly becoming the technique of choice for the identification of monoclonal proteins. The major advantages of IFE include its speed, high sensitivity, and relatively low cost. In conjunction with high-resolution electrophoresis and nephelometric analysis, IFE permits accurate, complete evaluation of protein abnormalities within hours instead of days. In addition, the precipitin bands obtained with IFE are directly comparable with high-resolution electrophoresis, which aids in the interpretation of electrophoretic patterns. The major disadvantage of IFE, in comparison with

IEP, is that the precipitin reaction is more dependent on the antigen/antibody ratio.

Turbidimetry and Nephelometry

Turbidimetry and nephelometry are techniques for the electrooptical quantitation of immune complex formation in solution. Because turbidimetry and nephelometry are precise, rapid, automated methods of immunoquantitation, they are extensively utilized in the clinical laboratory for the quantitation of a large number of medically important substances.

Light Scattering

The interaction of light with a particle can result in reflection, refraction, absorption, or scatter of the beam. When an immune complex is formed under carefully controlled conditions, measurement of these parameters can provide information regarding the quantity of analyte that is present (73). Turbidimetric techniques determine the reduction in the intensity of incident light from all interactions of an immune complex with a light beam, whereas nephelometric techniques measure light scattered at a specific angle to the incident beam. Although the ability of both systems to measure particles is highly dependent on the design and quality of the optical and electronic instrumentation utilized, nephelometry is more sensitive to small particles than turbidimetry, and turbidimetry more accurately quantitates large complexes (74).

Thus, for the interaction of light of wavelength λ and intensity I_0 with a particle of polarizability α, the intensity of emitted radiation ($I\theta$) at distance *r* from the particle is defined by the following relationship:

$$I_\theta = \frac{8\pi^4\alpha^2 I_0}{\lambda^4 r^2}$$

There are several consequences of this relationship (74). For example, because light scatter is proportional to $1/\lambda^4$, blue light (λ = 450 nm) is scattered more than red light (λ = 650 nm). In addition, white light scattered at an angle of 90 degrees appears blue, whereas transilluminated white light appears red because of removal of the blue wavelengths.

The formation of immunochemical aggregates, as depicted by the Heidelberger curve, involves the formation of small immune complexes through a primary immunochemical reaction, secondary interlinking of the complexes, and the tertiary formation of large aggregates through hydrophobic and charge-based interactions (75). Rayleigh scatter predominates during the primary and secondary immunochemical interactions, but large scatter complexes formed during the tertiary reaction may exceed the threshold for Rayleigh-Deybe scatter (Fig. 61.12).

Condition	In $Ag_x^{(m)}\ Ab_y^{(2)}$	Result
Extreme antibody excess	$\frac{y}{x} \geq m$	No steps (2), (3)
Antibody excess	$m > \frac{y}{x} > \frac{m}{2}$	Scatter↑ as $\frac{y}{x}$↓
Equivalence	$\frac{y}{x} = \frac{m}{2}$	Maximum scatter
Antigen excess	$\frac{m}{2} > \frac{y}{x} > \frac{1}{2}$	Scatter↓ as $\frac{y}{x}$↓
Extreme antigen excess	$\frac{y}{x} < \frac{1}{2}$	No steps (2), (3)

FIGURE 61.12. Light scatter during immunoprecipitation. The initial antigen concentration (x), antibody concentration (y), and antigen valency (m) are denoted as indices for various regions of the kinetic Heidelberger curve. The antibody concentration fixed, and the antibody is assumed to have a valence of 2. The regions of the kinetic Heidelberger curve include (1) primary immunochemical formation of antigen-antibody complexes, (2) secondary immunochemical linking of complexes, and (3) rapid build up of scattering centers owing to subsequent hydrophobic and charge-based interactions among complexes. (From Sternberg JC. Rate nephelometry. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:34, with permission.)

In the region of antibody excess, light-scattering immune complexes form most rapidly during the first 20 seconds of the reaction (74). However, the formation of additional light scattering complexes continues for several minutes, reaches a plateau (equilibrium phase), and then decreases as intraparticle destructive interactions occur, and an insoluble immunoprecipitate is formed. In conditions of antibody excess, both the peak rate of the formation of light-scattering complexes and the amount of complexes that exist at the plateau region depend on the antigen concentration. End-point nephelometric or turbidimetric systems measure the curve at the plateau phase to determine antigen concentration, whereas rate or kinetic immunoprecipitation methods use the portion of the curve where the peak rate of reaction is maximal.

A problem faced by all quantitative immunoassay systems, but particularly by turbidimetric and nephelometric systems, is the symmetric shape of the Heidelberger curve. Thus, a given measured signal can represent two distinct concentrations of antigen, one in the antibody excess zone and one in the region of antigen excess (75). In turbidimetry and nephelometry, measuring immune complex formation in the region of antibody excess minimizes this problem. This permits the formation of complexes of a constant size and provides a reproducible, stoichiometric relationship between the number of complexes formed at a given antigen concentration (75).

In addition to antigen concentration, antibody avidity, ionic strength, pH, the presence of molecules with hydrophilic or hydrophobic properties, and the nature and concentration of anion (60) influence the formation of immune complexes. Antichaotropic ions with a high charge density (PO_4^{3-}, HPO_3^{-}, SO_4^{2-}) decrease solvation effects and promote immune complex formation, whereas chaotropic ions (SCN^{-}, ClO^{4-}, NO^{3-}) increase the availability of solvent ions and reduce immune complex formation. The use of polymers, such as dextran and the polyethylene glycols (PEG), to enhance immunoprecipitation was first reported by Hellsing and Laurent in 1964 (76) and is commonly used in both turbidimetry and nephelometry. The immunoenhancing effect of polymers is probably due to steric exclusion, which has the effect of decreasing the solubility of protein molecules, driving the antigen-antibody interaction toward immune complex formation, increasing the slope of the antibody excess part of the precipitin curve, and displacing equivalence towards higher antigen concentrations (74). PEG with a molecular weight of 6000 daltons (PEG 6000) is one of the most effective immunoenhancers, leading to a markedly decreased reaction time and marked increases in peak rate (74). PEG also precipitates

lipoproteins, leading to increased optical clarity and decreased background scattering.

Turbidimetric Immunoassays

Turbidimetry has been used since the early 1950s as a method of quantitative analysis. In this technique, the reduction in light intensity caused by interaction of a light beam with a suspension of particles is determined spectrophotometrically. The ratio of the incident and transmitted light is determined, and expressed in absorbance units (*A*):

$A = 2 - \log_{10} \cdot T$

where *T* is the percentage of transmittance. In the clinical laboratory, the major advantage of turbidimetry is that it can be performed with automated spectrophotometric instruments. The major disadvantage of turbidimetry is the requirement of a relatively high particle concentration and the relatively small change in light intensity caused by particle absorbance. Consequently, optical quality and alignment are extremely important for accurate turbidimetric quantitation, and the results can be affected by extraneous factors, such as the presence of dust or dirt particles. Light with a wavelength in the near-ultraviolet (290–410 nm) is usually utilized for illumination, whereas the photodetector must be aligned with the incidence source and detects light over narrow angle (77). The optical clarity of the reagents and the reaction cuvettes or cells is important, and the system must be able to mix reagents rapidly for analysis. The antisera must have high avidity and specificity and show optical clarity. The IgG fraction is generally the most suitable for turbidimetric analysis. An immunoenhancer, such as dextran or PEG, is frequently added to the reaction mixture to facilitate immune complex formation, and this must not cause turbidity when mixed with the antisera. The presence of turbid substances (e.g., chylomicrons and other lipoproteins, endogenous immune complexes, monoclonal proteins) in the samples is problematic and difficult to compensate with reagent blanks or by other means.

Nephelometric Immunoassays

Nephelometric techniques have been applied very successfully to the immunochemical measurement of specific proteins, drugs and other substances (74,78). In nephelometry, a known amount of specific antibody is added to a solution containing the antigen being measured. The intensity of light scattered from the large antigen-antibody complexes formed during the reaction is measured, and the rate signal is transmitted to a microcomputer, where concentration units are determined.

Techniques of rate nephelometry are used almost exclusively in commercial nephelometers for clinical laboratory applications. In rate nephelometry, a fixed antibody concentration is utilized, and measurements of light scatter are made under conditions of antibody excess, where antigen concentration can be determined from measurements of light scatter. The time course of a nephelometric reaction has the complex shape shown in Fig. 61.13. A small amount of Rayleigh scatter is initially present, and this increases slightly when the sample and then the antisera are added. Scatter intensity then undergoes a sigmoidal increase with the progression of primary and secondary immunochemical reactions and the formation of large scattering centers. There is no real end point, but the scatter intensity gradually reaches a maximum and then decreases as large complexes settle out. The reaction conditions are adjusted so that the maximal rate (scatter versus time) occurs during the first minute of the immunochemical reaction.

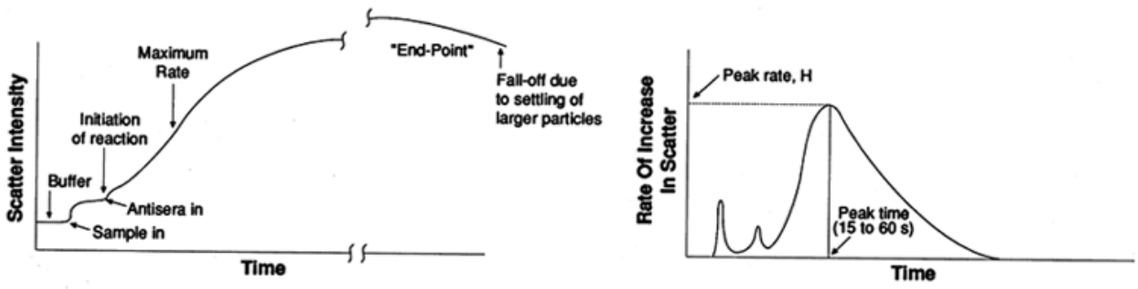


FIGURE 61.13. Time course of the nephelometric signal produced during an immunoprecipitation reaction. **A:** Variation in the t versus time. **B:** Rate of increase of scatter versus time. (From Sternberg JC. Rate nephelometry. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:34, with permission.)

Commercial nephelometric systems monitor scatter and the rate of change of scatter at a forward angle of 70 degrees in a stirred reaction cell. The reaction is performed in a phosphate-buffered saline solution, and PEG is used as an immunoenhancer. Other nephelometric systems, including a particle-enhanced turbidimetric inhibition immunoassay (PETINIA), have been described. PETINIA is a competitive binding assay using hapten-coated latex particles. Agglutination, measured by a change in turbidity at 340 nm, is inhibited by binding of the unlabeled hapten to the antibody (79).

Ligand Immunoassays

General Principles of Ligand Assays

A ligand is a substance that will bind to or complex with another substance. In a ligand assay, one component of the reaction (e.g., label or tracer) is tagged in some manner so that the reaction can be measured. The substance to be measured (ligand, analyte, or hapten) can be an immunoglobulin, drug, tumor marker, hormone,

infectious agent, autoantigen, lipoprotein, oncoprotein, or other substance. The substance that binds the ligand (binding substrate) is usually an antibody (immunoassay) but can be a receptor (receptor assay), carrier protein (protein-binding assay), or any substance that specifically binds the ligand with a high affinity. The labeled reagent can exist in a solubilized form, or it can be immobilized (bound) to a solid surface such as a test tube, the sides of the wells of a microtitration plates, a latex particle, a plastic bead, or a solution of microparticles (80). In addition to the analysis of soluble substances in body fluids, the principle of ligand analysis is also utilized in tissue-based assays such as direct and indirect immunofluorescence, flow cytometry, and immunoperoxidase staining.

Kinetic Aspects of Ligand Analysis

The amount of ligand that binds to a substrate is defined by the number of binding sites and by the affinity of the binding substrate. The affinity is the sum of the strength of the hydrogen binding, van der Waals forces, ionic interactions, and other forces that influence the reaction. Thus, by the law of mass action,

$$\mathrm{Ag} + \mathrm{Ab} \underset{K_2}{\overset{K_1}{\rightleftharpoons}} \mathrm{Ag} \cdot \mathrm{Ab}$$

where K_1 and K_2 are the equilibrium constants for the reaction.

At equilibrium, the rates of the forward and reverse reactions are equal ($K_1 = K_2$) and the proportions of bound and unbound ligand become constant. The equilibrium constant (K_a) for this reaction is represented by the equation:

$$K = \frac{k_1}{k_2} = \frac{(\mathrm{Ag} \cdot \mathrm{Ab})}{(\mathrm{Ag})(\mathrm{Ab})}$$

When 50% of the binding sites are occupied, K_a is equal to the reciprocal of the free ligand concentration. K_a is expressed in L/mole or moles/L^{-1} and has a magnitude in the range of 10^7 to 10^{11} L/mole in most clinical laboratory assays.

Many ligand assays are used in the clinical laboratory for the quantitation of a wide variety of substances of medical importance. Basically, these assays differ in the labeling substance (e.g., radioisotope, enzyme, fluorescent dye, chemiluminescence precursor, bioluminescent substance, metal atom, bacteriophage, liposome, metal sols, latex particles), binding substrate (e.g., antibody, receptor, transport protein), reagent concentration (antibody excess, competitive binding), and the requirement for separation of the free and bound label (heterogeneous or homogeneous).

Unfortunately, the diversity of methodology has led to practical problems in the classification and nomenclature of the ligand assay systems (81). By convention, the type of label appears first in competitive assays (i.e., EIA, RIA), whereas the binding substrate appears first in the name of a reagent excess assay (i.e., immunoenzymetric assay). All reagent excess assays using an antibody-binding substrate can be referred to as immunometric assays. The acronym ELISA (enzyme-linked immunosorbent assay) is commonly, and incorrectly, used to refer to all types of EIAs and immunoenzymometric assays. The classification system utilized in the following (Table 61.2) is based on the classic differentiation of heterogeneous and homogeneous assays and a classification system proposed by Gosling (81).

TABLE 61.2. CLASSIFICATION OF LIGAND IMMUNOASSAYS

Heterogeneous ligand assays
Competitive
Noncompetitive
Antigen capture
Antibody capture
Homogeneous ligand assays
Enzyme-multiplied immunoassays
Enzyme immunochromatography
Substrate-labeled fluorescent immunoassay
Fluorescence polarization immunoassay

Heterogeneous Ligand Assays

Heterogeneous ligand assays require physical separation of the free label from the label bound to the substrate. Techniques for this purpose rely on differences in the size or chemical composition of the bound and free label. The development of techniques that allow complete separation of the bound and free components has been one of the most challenging aspects of ligand assay techniques. In general, the relatively nonspecific separation methods used in the past (adsorption, chemical precipitation, immunoprecipitation) have been replaced by more specific separation methods, most notably solid phase techniques.

Adsorption is based on the difference in size between the free label and the label-substrate complex. Porous materials (e.g., charcoal, silica, ion exchange resins, Sephadex) adsorb small molecules (free label) but not larger particles (label-substrate complex). The adsorbent, with attached free label, can be removed from the reaction mixture by centrifugation or other techniques, leaving the substrate in the supernatant. In practice, a mixture of charcoal and cross-linked dextran was for a long time the most widely utilized adsorbent. Dextran prevents nonspecific adsorption of protein and determines the size of the free label that can be adsorbed. The properties of the adsorbing particles such as Sephadex can be also modified to maximize the specificity of the separation process. Chemical agents (e.g., ammonium sulfate, ethanol, sodium sulfate, PEG) alter the solubility of proteins in solution, causing precipitation. After centrifugation, the free label remains in the supernatant, where it can be measured. The action of these agents is nonspecific because both free and bound proteins are precipitated. Immunoprecipitation (second antibody, double antibody) methods use a precipitating antibody specific for the species in that the primary antibody is produced (e.g., anti-goat). Insoluble complexes are produced that can be removed by centrifugation, leaving the free label. These methods are similar to chemical precipitation in that all free and bound (labeled and unlabeled) substrate is removed from the reaction mixture.

In solid-phase separation, either the ligand or primary antibody is immobilized to an inert physical surface. Because the antigen-antibody reaction occurs on the surface, the free and bound ligand can easily be separated. Several imaginative solid-phase methods have been devised, first utilizing coated paper or

polystyrene (e.g., test tubes, beads, microtiter trays, paddles), followed by membranes (e.g., nitrocellulose, nylon, polymers) and more recently plastic and latex beads and microparticles, which may be magnetic. The major requirements for solid-phase immobilization is complete and uniform coating of the surface, linkage of the antigen or antibody to the solid phase by a technique that neither alters binding interactions nor permits spontaneous release during the reaction, and thorough washing to minimize nonspecific binding of the label. This analytical process has now been automated for a variety of analytes, such as therapeutic drugs, specific proteins, enzymes such as cardiac markers, hormones, antibodies to infectious agents, and drugs of abuse by many manufacturers. Automation has helped promote assay stability, reduced imprecision, speeded up the analytical process, and reduced labor requirements for performing these assays.

Heterogeneous ligand assays can be competitive or noncompetitive. In competitive assays, a constant amount of binding substrate and labeled ligand is used, and the amount of unlabeled ligand is the variable that is measured (Fig. 61.14). The labeled ligand is present in relatively large quantities, and a smaller amount of unlabeled ligand is allowed to compete for the available binding sites. The larger the quantity of unlabeled ligand present, the more labeled ligand will be displaced from the binding substrate. After separation, the amount of free or bound labeled ligand can be determined.

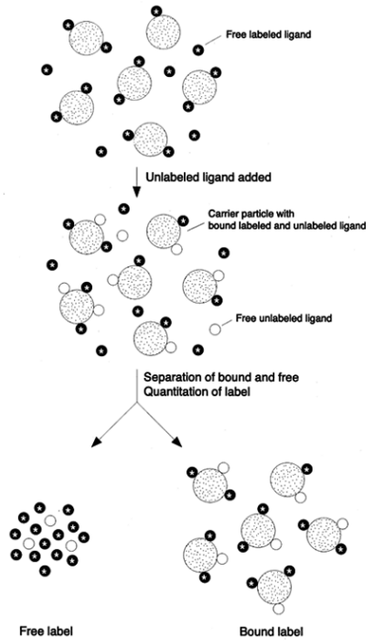


FIGURE 61.14. Principle of competitive binding immunoassays. An unlabeled ligand is added to a mixture containing measured amounts of free-labeled ligand and carrier particles with bound-labeled ligand. The label is usually a monoclonal antibody or a radioactive ligand. Displacement of some of the bound-labeled ligand by the unlabeled ligand results in a mixture of free-labeled ligand, free unlabeled ligand, and carrier particles with labeled and unlabeled ligand. The carrier particles are then separated from the mixture, and the amount of free or carrier-bound ligand is determined by the appropriate assay procedure. If standard curves are prepared using known amounts of unlabeled ligand, the amount of analyte in an unknown solution can be determined from the curves.

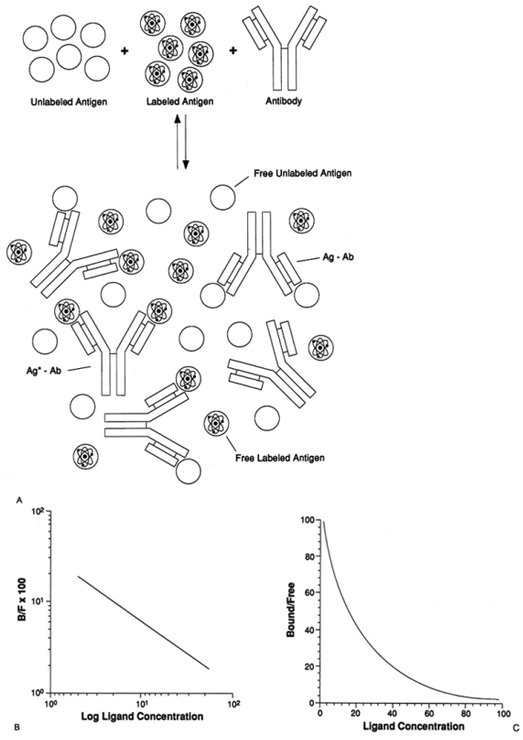


FIGURE 61.15. Radioimmunoassay by competitive binding. **A:** Schematic illustration of the theory of competitive binding. **B:** Graphic analysis by logit plot. **C:** Graphic analysis by a plot of bound/free versus ligand concentration.

The principle of competitive binding was first applied in an electrophoretic technique for the measurement of thyroxine in human plasma but quickly became the most widely used RIA technique (82). Competitive binding assays depend on the competition between radiolabeled and unlabeled antigen for the binding sites on a limited amount of antibody. In this procedure, the unlabeled analyte is allowed to compete with a highly purified radiolabeled antigen (radiolabeled ligand) for binding to a limited amount of antibody, under conditions of antigen excess. After incubation under appropriate conditions, the antibody-bound (B) and unbound (free, F) analytes are separated and the amount of radioactivity is determined in each. These values are used to calculate the ratio of bound to free analyte (B/F). The more unlabeled analyte present, the lower the B/F ratio because more labeled analyte will be displaced from the antibody. A sigmoidal curve results if the B/F ratio is plotted against the ligand concentration. These data are usually analyzed by a computer with a four-parameter logistic or spline function, avoiding the need to transform the data into some linearizing function such as logit/log of the ligand concentration, which was common when manual graphing of calibration curves was necessary. The logit is calculated by the following equation:

$$\frac{Y}{1-Y}$$

where,

$$Y = \frac{B/F}{B_0/F_0}$$

where B is the counts per minute of bound radioligand, F is the counts per minute of free radioligand, B_0 is the counts per minute of bound ligand determined in a system with no unlabeled ligand (zero value), and F_0 is the counts per minute of free radioligand in a system with no unlabeled ligand (Fig. 61.15).

In practice, the concentration of an unknown analyte is determined by reference to a standard curve prepared from the reaction of different concentration solutions of purified unlabeled antigen. It is necessary that the unknown specimen be analyzed under precisely the same conditions as the standard curve.

The labeled and unlabeled antigens are not required to be chemically or biologically identical but must follow the law of mass action.

A similar principle of competitive binding is also used in the radioreceptor assay. In the radioreceptor assay, the receptor is not an antibody but a purified preparation of a circulating binding protein (e.g., thyroid-binding globulin, cortisol-binding protein, intrinsic or for vitamin B_{12}), membrane receptor, or cytoplasmic receptor.

Noncompetitive (reagent excess) assays are usually of the "sandwich" type, in which a binding molecule is immobilized to an inert support. Antigen-capture assays require an immobilized antibody. The analyte or ligand control is added and binds to the available substrate sites. Labeled antibody is then added, and a sandwich is formed with the antigen in the middle. After a wash step to remove unbound labeled antibody, the bound antibody is measured and is directly proportional to the quantity of unlabeled antigen present. The antibody capture is similar but requires a ligand linked to the solid phase. After reaction with an antibody-containing specimen or control and a wash step to remove unbound antibody, the presence of bound antibody is detected by the use of labeled antihuman globulin. The amount of bound label is directly proportional to the quantity of specific antibody (Fig. 61.16). Note that these assays are generally applicable only to molecules that are large enough to have at least two distinct antibody-binding epitopes.

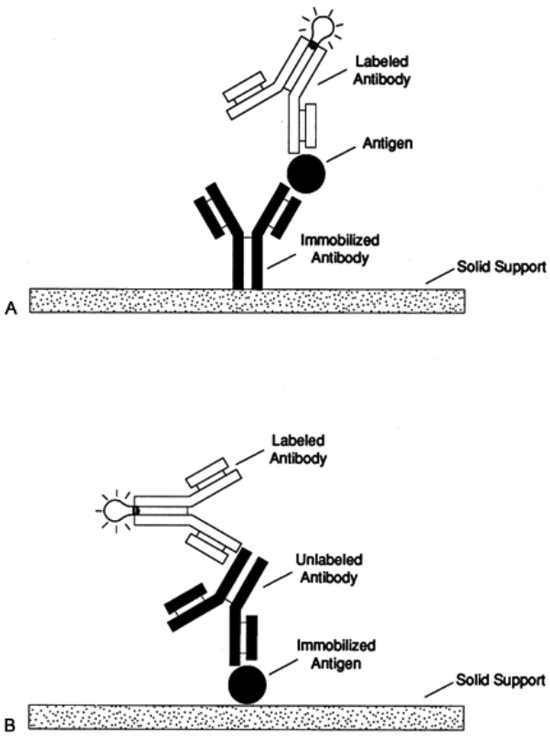


FIGURE 61.16. Principle of noncompetitive immunoassays. **A:** Antigen capture. **B:** Antibody capture. Antigen capture assays utilize an antibody immobilized to an inert support (usually a plastic test tube or plastic microtiter tray). Antigen (control ligand or analyte) is added, and an immune complex forms on the inert support. After incubation and washing, a labeled antibody specific for the antigen is added. The reaction mixture is incubated, washed, and the amount of bound antibody is measured. The concentration of analyte in an unknown solution can be determined from a standard curve prepared with known amounts of antigen. In antibody capture assays, a similar principle is used to measure the concentration of an antibody.

Homogeneous Ligand Assays

Homogeneous ligand assays do not require separation of the free and bound fractions. Many of these assays have grown in popularity in the clinical laboratory in recent years because of their speed, sensitivity, and overall technical simplicity. There are numerous types of homogeneous ligand assays using enzyme and fluorescent labels. Examples include the enzyme-multiplied immunoassay technique (EMIT, Syva Co., Palo Alto, CA, U.S.A.), enzyme immunochromatography, substrate-labeled fluorescent immunoassay, and fluorescent polarization immunoassay, and cloned enzyme donor immunoassay (CEDIA). These and other homogeneous ligand assays are described in the sections below.

Isotopic and Nonisotopic Labels

The choice of label is an important consideration in ligand assays. In this regard, the sensitivity of the assay largely depends on the specific activity (number of detectable events per label molecule per unit time) of the label and the completeness of the separation of bound label from unbound label. However, other factors, such as the ease and sensitivity of detection, freedom from background activity or environmental influences, ease of conjugation to a ligand, cost, and other factors are important.

Because the introduction of the RIA 40 years ago, iodine 125 (^{125}I) and other radioactive nuclides have been widely used as labels in immunoassay procedures. The popularity of radioisotopes as a label has been because of their sensitivity, measurement precision, lack of background, and freedom from environmental influences. However, these compounds are expensive, pose a slight health risk to the laboratory technologist, require special handling and disposal, and have a limited shelf life. In addition, specialized, expensive instrumentation is required for measurement of radioactivity, automation is difficult, licensing is required, and strict compliance with national, state, and local laws is required for their use (83). Fortunately, the sensitivity of the nonisotopic label immunoassays now exceeds that of RIA for many substances. With the recent emphasis on cost-effective laboratory diagnosis and the urgent need to decrease the use of radioisotopes and other substances that pose a health risk, the use of RIA techniques has decreased during the past two decades in clinical and research laboratories. By late 1999, the CAP Ligand survey (K/KN-C) indicated that less than 3% of laboratories used radiolabeled assays for prostate-specific antigen, 1% for thyroid-stimulating hormone (TSH), and approximately 2% for vitamin B_{12}. Because these represent assays for which radioactive labels had been thought necessary for sensitivity purposes, it is clear that this belief no longer holds. In contrast, in 1997, 35% of the articles in the journal *Clinical Chemistry* were devoted to some aspect of EIA (84).

The relative advantages and disadvantages of radioisotopes with nonisotopic labels (enzymes, fluorochromes, chemiluminescent substances) have received much consideration (85,86). For example, ^{125}I has a specific activity of one detectable event per second per 7.5×10^6 labeled molecules, whereas chemiluminescent

labels produce one detectable event per labeled molecule, and both enzymes and fluorescent labels can produce many detectable events per molecule. Thus, an enzyme label with a chemiluminescent or fluorescent substrate has in principle the lowest detection limit of commonly used labels. In practice, however, other issues such as nonspecific binding of labeled antibody to the solid phase clearly contribute to detection limit, as one commercially available immunoassay with a colorimetric substrate (in principle, several orders of magnitude less sensitive than chemiluminescence) has achieved a third-generation TSH assay. Moreover, various vendors' assays for TSH use a variety of label/separation/incubation combinations, often with very similar detection limit.

Radioimmunoassay

RIA techniques have been among the most sensitive laboratory procedures available. As a result, RIA was long the principal method for the detection of substances that are present in very low concentration in biological fluids (10^{-6} to 10^{-9} g/mL). In clinical medicine, the RIA was introduced in 1960 by Yalow and Berson (87) for the measurement of plasma insulin levels. It was quickly applied to the measurement of other peptide hormones and has subsequently been used for the determination of hundreds of substances, including drugs and toxins, infectious agents, tumor antigens, tissue peptides, and other substances.

The main requirements for a RIA are an antibody with a high specificity for the antigen being determined and a pure, radioactively labeled antigen of known concentration. In a RIA, one component of an immune complex reaction (antigen or antibody) is labeled with a radioactive isotope. After incubation, the immune complex that has formed is separated from the reaction mixture and the amount of bound (in the immune complex) and free radioactivity are determined. With appropriate standards, the amount of the substance under study (analyte) can be determined. A special type of RIA (radioallergosorbent technique) is used for the quantitation of antigens that are immunoglobulins. For nonimmunoglobulins, there are several RIA techniques, including the competitive binding assay and immunoradiometric assay.

RIA Techniques

Competitive binding assays and excess reagent (sandwich) [this is immunoradiometric assay (IRMA), discussed later] are the most common types of RIA. The sandwich technique requires a radiolabeled antibody that is utilized in excess. The amount of analyte is determined from the radioactivity in the immune complex. This type of RIA procedure is rarely used in the clinical laboratory. Competitive binding RIA techniques depend on the competition between radiolabeled and unlabeled antigen for a limited amount of antibody.

The excess reagent (two-site, double-antibody) RIA is a sandwich technique in that two antibodies are used in excess. One antibody, which is unlabeled, is attached to a solid phase (usually a polystyrene test tube) by electrostatic interaction. The test solution, containing unlabeled antigen (with two antibody-binding sites) is added to the tube, where an antigen-antibody complex forms. After excess antigen is removed by washing, the second (radiolabeled) antibody is added, and the formation of the sandwich is completed. After another wash step to remove unlabeled antibody, the amount of bound radioactivity is measured, and is proportional to the amount of antigen in the test solution (Fig. 61.17).

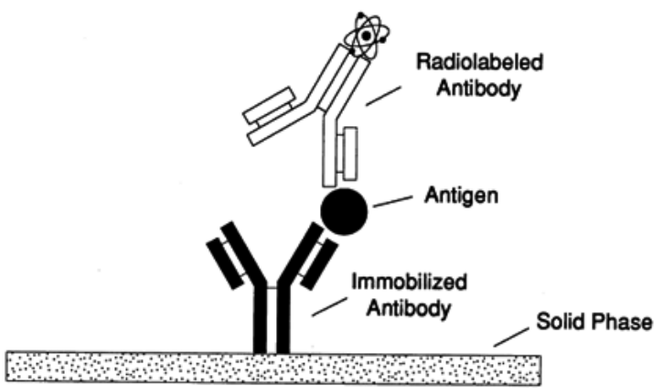


FIGURE 61.17. Principle of sandwich radioimmunoassay. This technique is an antigen capture immunoassay in that a radiolabeled antibody is used to measure the concentration of an antigen.

In the IRMA, the antigen is reacted with an excess of antibody, which is usually coupled to a solid phase. A wash step may or may not occur here, in two- and one-step assay designs, respectively. A second antibody, labeled with a radioactive tracer, is then added to the reaction mixture forming an Ab-Ag-Ab* sandwich. Finally, the solid phase is washed to remove nonbound radiolabeled antibody. The amount of antigen in the test solution is directly proportional to the radioactivity remaining.

The radioallergosorbent technique is a sensitive method for the detection of specific antibodies in biological fluids. It is a sandwich technique using a purified antigen bound to a solid support. The immobilized antigen is incubated with the test solution, resulting in the formation of an immune complex on the solid support. The formation of the sandwich is completed by the addition of radiolabeled antiglobulin, and the amount of bound radioactivity is determined after a final wash. Under these circumstances, the measured radioactivity is directly proportional to the amount of antibody present.

Practical Considerations in RIA Analysis

The major considerations in a RIA are an antibody with a high specificity for the analyte being determined, a highly purified, radioactively labeled preparation of the analyte, and an efficient method for separating antibody-bound and free components after the reaction. Other considerations include specimen adequacy, measurement of radioactivity, and the appropriate utilization of standards.

- Specimen
- Antibody
- Labeled analyte
- Separation of bound and free antibodies
- Measurement of radioactivity
- Standards

The specimen for analysis must be representative and obtained in the correct manner. In addition, care must be taken to

maintain the integrity of the analyte until the analysis is performed. Because labile protein substances are frequently analyzed, specimens for RIA are often centrifuged and maintained in the frozen state. Specimens sent to another laboratory for analysis are often kept frozen using dry ice.

Polyclonal and monoclonal antibodies have been successfully used in RIA procedures. The preparation of an antibody for an RIA procedure is performed by the standard techniques described above. Because many substances assayed by RIA have small molecular weights, carrier proteins are commonly employed during antibody production, and some antibodies recognize the analyte in addition to a portion of the carrier molecule. The specificity, avidity, and specificity of the antibody are critical factors in RIA. The avidity (affinity) of the antibody is a major determinant of the sensitivity of the assay and is determined using Scatchard or Wolf plots. Generally, antisera for RIA procedures must have K_a values between 10^{-2} and 10^{-9} mol/L (88). The sensitivity of the antibody is determined by binding studies utilizing the purified analyte and potential cross-reacting substances at a wide range of concentrations. The specificity of a polyclonal antibody preparation can usually be improved by the absorption of cross-reacting antibodies. In addition, steps can be taken to remove cross-reacting substances from the test solution before analysis. An antisera titer is chosen (usually $1:10^5$ to $1:10^6$) that will bind 30% to 50% of the labeled analyte (88).

The preparation of a labeled analyte involves the choice of a radioactive isotope, physical attachment of the isotope to the analyte, and isolation of the labeled analyte in a highly purified form. The choice of radiolabel depends on its ease of conjugation, specific activity, radioactive half-life, and emission characteristics, with gamma emission being preferred when possible because of the ease of counting compared with beta emission. The specific activity (activity per unit weight of isotope) determines the amount of radiolabeled ligand that must be present to produce a adequate sensitivity. The half-life of the radioisotope must be sufficient to provide adequate shelf life. Generally, ^{125}I is used as a label for proteins, peptides, and hormones because it has a relatively long half-life (60 days), emits a large number of disintegrations per minute, and emits gamma radiation of relatively low energy. Hydrogen 3 (^3H) is often used to label steroids, but its long half-life, beta emission, and relatively low number of disintegrations per minutes are disadvantages. Other radioisotopes used in RIA include carbon 14 and iodine 131. For the measurement of vitamin B_{12}, cobalt 57 or cobalt 60 are used to substitute for the cobalt that is naturally present. The characteristics of these radioisotopes are listed in Table 61.3.

TABLE 61.3. PROPERTIES OF SELECTED RADIOISOTOPES USED IN PATIENT CARE

			Beta Emission			Gamma Emission	
				Energy (MeV)			
Element	Isotope	Half-life	Abundance %	Max	Mean	Abundance %	Energy (MeV)
Hydrogen	^3H	12.3 a	100%	0.0186	0.0057	None	—
Carbon	^{14}C	5,730 a	100%	0.1561	0.0493	None	—
Phosphorus	^{32}P	14.3 d	100%	1.710	0.6948	None	—
Chromium	^{51}C	27.8 d	ec-1	—	—	9	0.3198
			ec-2		—		
Iron	^{59}Fe	45 d	53	0.4750	0.1527	56	1.0950
			45	0.02730	0.0808	43	1.2920
			1.1	0.1300	0.0355	2.8	0.1925
Cobalt	^{57}Co	270 d	ec-1	—	—	89.2	0.0144
			ec-2	—	—	89.0	0.1219
						11.0	0.1363
Cobalt	^{60}Co	5.26 d	99.8%	0.313	0.0941	100	1.3325
						99.8	1.1732
Technetium	^{99m}Tc	6.0 hr		Isometric level decay		98.6	0.0022
						98.6	0.1405
						1.4	0.1427
Iodine	^{125}I	60 d	ec	—	—	100	0.0355
Iodine	^{131}I	8.05 d	90.4	0.0606	0.1917	85.3	0.3645
			6.9	0.33	0.0955	6.9	0.6370
			1.6	0.25	0.0701	5.1	0.0802
						5.1	0.2843
						1.6	0.7229
Cesium	^{137}Cs	30.0 a	93.5	0.514	0.1749	93.5	0.6616
			6.5	1.176	0.4272		

From Holden NE, Walker FW. *Chart of the nuclides*, 10th ed. U.S. Atomic Energy Commission, 1968.

Chemical methods are used to attach the radiolabel to the analyte. It is essential that this be done in such a way as to achieve a uniformly labeled product with maximal utilization of the reactants and with minimal chemical damage to the analyte. For peptides and proteins containing tyrosine residues and labeled with ^{125}I, an iodination reaction catabolized by lactoperoxidase or chloramine-T is often used (88). Lactoperoxidase catalyzes the reaction of H_2O_2 and $Na^{125}I$ to form active iodine (I^+), which in turn reacts with the tyrosyl residue of proteins. Chloramine-T is a derivative of ρ-toluene sulfonamide that produces a similar incorporation of iodine into the aromatic ring of tyrosine. Enzymatic iodination with lactoperoxidase is preferable because it

causes fewer reactions with other amino acid residues than chloramine-T. A conjugation method using an iodinated ester is employed to label proteins lacking tyrosine residues.

When the labeling process has taken place, the labeled analyte is separated from free label and damaged analyte by chromatographic or electrophoretic procedures. The product is thoroughly evaluated for nonspecific binding and immunoreactivity under several conditions (88). In addition, a high specific activity (large amount of radioisotope per quantity of analyte) is desirable. The specific activity is commonly expressed as microcuries per microgram (μCi/μg) or disintegrations per micromole (dpm/μmol). For routine RIA procedures, an amount of label producing 10,000 to 25,000 dpm per assay tube (generally 0.01 μCi) is used, although maximal sensitivity is achieved with a high-affinity antibody and minimal quantities of label (often producing only 1,000 to 2,000 dpm/tube) (88).

Separation of the free and bound analyte is the least precise part of the RIA. Because the antigen-antibody complex is larger than the free antigen, this size difference is the basis of most separation methods. Although several separation methods have been employed, solid-phase separation, immunoprecipitation, and chemical precipitation have been most widely described. Adsorption with dextran-coated charcoal particles is a solid-phase separation technique that is employed in the RIA analysis of steroids and many other substances (89). In this technique, the smaller unbound antigen particles are bound to the charcoal, whereas the larger complexes remain in the supernatant. Unfortunately, it is difficult to prepare a uniform suspension of charcoal particles for pipetting, and ionic factors, protein concentrations, and many other factors can interfere with the adsorption process. Adsorption with silicates and chalk has also been used. Chemical substances such as saturated ammonium sulfate solution, ethanol, or PEG (Carbowax) precipitate the antigen-antibody complex, which can be removed from the reaction mixture by centrifugation (90). Immunoprecipitation (double-antibody precipitation) methods use an antibody (secondary antibody) from a different animal species, which is directed against epitopes of the antibody (primary antibody) binding the analyte. Under the proper reaction conditions and antibody concentrations, precipitates are formed that can be separated by centrifugation, electrophoresis, or chromatography (Sephadex, polyacrylamide, ion exchange, and paper). Although simple in theory, in practice, it is difficult to achieve precision with immunoprecipitation. A combination of immunoprecipitation and chemical separation with PEG has been used to shorten turnaround time in RIA testing. Recently, coated magnetized particles have been used successfully in some assays.

Enzyme Immunoassay

The EIA is a type of nonisotopic immunoassay in which enzymes, coenzymes, fluorigenic substrates, or enzyme inhibitors are used as labels (91).

Enzymes are utilized as labels in both heterogeneous and homogeneous assay systems and are used to assay both antigens and antibodies. The major prerequisite is that the antigen or antibody must be linked to an enzyme without destroying the immunologic or enzymatic activity of the antigen-antibody complex. In addition, in heterogeneous EIAs requiring a solid phase, the antigen or antibody under study must be attached to the solid-phase support in a manner that does not alter immunologic activity.

Enzymes used in immunoassay systems must be stable, available in a highly purified state, have a high turnover rate, undergo minimal interference by substances likely to be in the test solution, and be specific for the substrate. The final reaction product should be detected by a convenient means with a low detection limit. The most widely used enzyme in EIA is horseradish peroxidase (HRP). The substrate of HRP is hydrogen peroxide (H_2O_2) and the product is oxygen. This oxygen produced during the reaction is used to oxidize a reduced, colorless chromagen (usually reduced orthophenylenediamine). The final product, oxidized orthophenylenediame, has a brown color. Most EIAs utilize HRP or alkaline phosphatase as labels, although glucose oxidase, ß-D-galactosidase, and many other enzymes have also been used. Utilizing fluorometric techniques, the respective detection limits for HRP, ß galactosidase, and alkaline phosphatase are 5, 0.2, and 10 attomol (81). The practical detection limit of the EIA is approximately 0.01 to 0.02 attomol of ligand (92,93).

The coupling of an enzyme to an antibody or protein antigen must be performed in a manner that minimally alters immunoreactivity and enzyme activity. In addition, the coupling reaction must result in efficient incorporation of the label, and the conjugate should be stable with a long shelf life (94). Chemical reactions resulting in the activation of a carboxyl group are commonly utilized for the formation of hapten-enzyme conjugates. The mixed anhydride, carbodiimide, and *N*-hydroxysuccinimide ester coupling methods utilize this reaction. Other methods of hapten-enzyme coupling include periodate cleavage, or the use of heterobifunctional active esters such as m-maleimidobenzoic acid *N*-hydroxysuccinimide ester, or related compounds (95). Methods for protein-protein coupling include glutaraldehyde cross-linking, periodate coupling, dimaleimide coupling, the use of homobifunctional reagents such as 4,4′-difluoro-2,2′-dinitrophenyl sulfone, toluene-2,4-diisocyanate, or benzoquinone, and the use of m-maleimidobenzoic acid *N*-hydroxysuccinimide ester. Although coupling reactions that involve sites of antigen-antibody combination usually result in a reduction in avidity, coupling can result in an increased affinity and avidity in some circumstances. After the coupling reaction, the conjugate must be purified and then characterized in terms of the hapten-to-enzyme ratio, enzyme activity, and immunoreactivity (95).

Classification of EIAs

EIAs are usually classified based on the requirement of separation of the free and bound reactants (Table 61.4). In heterogeneous EIAs, enzyme activity is not affected by the interaction of antigen and antibody, and a separation step is required. The term ELISA is widely used as a general term for the heterogeneous EIA. In homogeneous EIA, there is no need for a separation step because enzyme activity is modulated by the antigen-antibody interaction.

TABLE 61.4. CLASSIFICATION OF ENZYME IMMUNOASSAYS

Heterogeneous enzyme immunoassays
Competitive heterogeneous enzyme immunoassays
Noncompetitive heterogeneous enzyme immunoassays
Two-site immunoenzymometric assay
Double antibody immunoenzymometric assay
Homogeneous enzyme immunoassays
Enzyme-multiplied immunoassay technique
Substrate-labeled fluorescent immunoassay
Coenzyme-labeled immunoassays
Immunoassays with ligand-labeled prosthetic groups and ligand-labeled modulators
Enzyme-enhancement immunoassay

The separation techniques that can be used in heterogeneous EIAs are limited by the large size of most enzyme molecules. As a result, solid-phase techniques (e.g., antigen- or antibody-coated

polystyrene test tubes or microtiter trays; polystyrene, latex, or agarose beads or particles, magnetized beads) have been principally utilized for separation of the bound and free enzyme conjugate. Immunoprecipitation with PEG, a second antibody, or a preprecipitated complex of two antibodies has also been utilized in some cases.

Competitive EIA techniques are similar to RIA, with the exception that they may be configured to be homogeneous as well as heterogeneous formats. One widely used adaptation of this technique, EMIT, uses a limited amount specific antibody that is reacted with unlabeled ligand (unknown sample, control, or standard) and a known amount of labeled antigen in addition to the enzyme substrate. In this case, binding of the ligand-enzyme conjugate to antibody greatly reduces the enzyme activity, so activity is proportional to ligand concentration. The end product is quantitated by spectrophotometry. The major disadvantages of this method include the need for relatively large amounts of pure antigen and the requirement that each different antigen be coupled to an enzyme.

EMIT was the first homogeneous EIA to be developed and has since been extensively been used for the quantitation of numerous chemotherapeutic substances (e.g., drugs of abuse, antiepileptic drugs, cardioactive drugs, antineoplastic drugs, antimicrobial drugs, antihistaminic drugs), hormones, and other substances. In the EMIT assay system, the ligand (usually a hapten) is covalently linked to an enzyme (usually nicotinamide adenine dinucleotide (NAD)–dependent glucose-6-phosphate dehydrogenase). In most EMIT assays, the binding of an anti-hapten antibody inhibits enzyme activity, probably by interfering with the molecular conformational changes that take place during the enzyme reaction or by direct steric exclusion of substrate binding (96). Because free hapten competes with the antibody for binding to the enzyme-immobilized hapten, the quantity of end product is directly proportional to the concentration of free hapten. In a few EMIT assays, the binding of antibody to the bound hapten increases enzyme activity. For example, in the EMIT assay for thyroxine, the enzyme-hapten (malate dehydrogenase-thyroxine) complex is inactive until antibody binding occurs. It is believed that the bound thyroxine occupies the active site of the enzyme but is removed from this area during the steric rearrangements that take place upon antibody binding (97). In the standard EMIT assay, 50 mL of the specimen is diluted with buffer, and the enzyme substrate is added, followed by the enzyme-hapten conjugate. The reaction mixture is aspirated into the flow cell of a spectrophotometer that is maintained at 30°C, and the change in absorbance is determined over a 30-second interval. The assay for some substances requires different specimen volumes, reading times, or pretreatment of the specimen to improve sensitivity. Although the basic equipment for an EMIT assay consists of a spectrophotometer and pipetter-diluter, numerous adaptations have been made for use in automated analyzers such as continuous flow systems, discrete analyzers, centrifugal analyzers, and the Du Pont ACA. The NAD-dependent reaction can be measured fluorometrically as well, and this detection method has been used to increase the sensitivity of the reaction.

Ligand-labeled enzyme cofactors have been effectively utilized in EIA systems (98,99 and 100). Most commonly, these systems are competitive in nature, with a specific antibody against the ligand, preventing the coenzyme-ligand conjugate from binding to the apo-enzyme and reducing enzyme activity. Because unlabeled ligand competes with the coenzyme-ligand conjugate for binding to the antibody, enzyme activity is directly proportional to the amount of unlabeled antigen. NAD is a cofactor for many enzymes (e.g., lactate dehydrogenase, malate dehydrogenase) and has been used as the cofactor in several of these assays. The major disadvantage of coenzyme-labeled immunoassays is their susceptibility to endogenous cofactors and other enzymes.

Immunoassays using ligand-labeled prosthetic groups are similar to those using enzyme cofactors. The ligand-prosthetic group (usually flavin adenine dinucleotide) competes with unlabeled ligand for a limited amount of specific antibody. The free, but not the antibody-bound, conjugate can interact with an apo-enzyme, such as apo-glucose oxidase, to generate the active enzyme (98). Similar assays with ligand-labeled modulators (e.g., antibody, inhibitor, receptor) have been reported (101).

The enzyme-enhancement immunoassay is a method in which enzyme-labeled antibodies are used (102). In this system, an enzyme-labeled antibody (E-AB1) and a succinylated antibody (AB2) are utilized. These antibodies form a complex with a polyvalent antigen in the reaction mixture to form a negatively charged microenvironment. Under these circumstances, the substrate is converted into a product (P2) that alters the light-scattering properties of the complex. The free-enzyme conjugate forms a product (P1) that remains soluble. The reaction is measured by turbidimetry.

Immunoenzymometric assays are noncompetitive (i.e., excess reagents are used). The antigen is first reacted with an excess of enzyme-labeled antibody, and a solid-phase antigen is then added in excess to remove unreacted enzyme-labeled antibody. After washing, the substrate is added. The amount of final product is inversely proportional to the concentration of free antigen. The major advantages of this technique include no need for a purified antigen and the ability to detect small haptens that are difficult to quantitate by other means.

Several sandwich modifications of the EIA to quantitate antigen have been described. The two-site immunoenzymometric assay utilizes a solid phase coated with specific antibody in excess. The test solution containing antigen is added to the solid phase and incubated, permitting the antigen to bind to the immobilized antibody. After washing to remove unreacted antigen,

the solid phase is incubated with an excess amount of enzyme-labeled antibody in the liquid phase. The enzyme substrate is added after a second wash, and the amount of end product is directly proportional to the amount of antigen in the test solution. Because the antigen must be able to bind to two different antibodies simultaneously, this procedure is best utilized to detect large, complex antigens with two or more antigenic sites. A competitive sandwich assay for antigen detection (double-antibody immunoenzymometric assay) uses solid-phase antigen. The solid phase is allowed to compete with free antigen for binding to a specific antibody in the liquid phase. After incubation and washing, an enzyme-labeled antibody against the first antibody is used to determine the amount of bound complex. This method is economical because a single enzyme-labeled antibody can be used in assays for many different antigens.

Recent Advances in EIA

Numerous innovative modifications of EIA techniques have been described utilizing microparticles, liposomes, recombinant enzyme fragments, biosensors, flow cytometry, and other materials and techniques. For example, Wilkins et al. (103) developed a highly sensitive homogeneous assay for thyrotropin using 800-nm particles and a new modular immunoassay system (Multipact). The assay was also the first to use fragmented monoclonal antibodies to avoid serum interference. The accuracy and precision of the assay were comparable with a commercial, highly sensitive IRMA for TSH. Two novel phase-separation immunoassays were described by Auditore et al. (104). In one system, the solid phase is generated *in situ*, after specific binding has occurred. This technique (*de novo* polymerization) to minimize nonspecific binding and enhance reaction kinetics. Thermal precipitation was used in the second system, which incorporated water-soluble polymers and temperature-dependent solubility.

The use of liposomes in immunoassay systems has also been described. For example, in the technique reported by Litchfield et al. (105), alkaline phosphatase was incorporated into liposomes. In a competitive assay, lysis of the liposomes and enzyme release by a hapten-cytolysin conjugate was prevented by conjugation with an anti-hapten antibody.

An EIA system using recombinant enzyme fragments has been described (106,107). In this system, termed CEDIA, two separate *Escherichia coli* genes for ß-galactosidase are engineered to produce separate, enzymatically inactive fragments that can spontaneously recombine to form an active ß-galactosidase enzyme. An anti-ligand antibody is attached to one fragment (enzyme acceptor), whereas the second fragment (enzyme donor) is labeled with ligand conjugate. The binding of antibody to the enzyme donor–ligand conjugate regulates the degree of recombination of the two fragments to form an active enzyme. In practice, a competitive system is used in which there is competition between the ligand in the specimen and the enzyme donor–ligand conjugate for a limited amount of antibody. Variations of the assay include a three-reagent system with enhanced sensitivity, and a single-reagent system suitable for on-site testing. The advantages of CEDIA compared with other homogeneous immunoassays include high sensitivity, linear standard curves, and low endogenous enzyme activity. The assay has been adapted for use in several commercial automated clinical laboratory analyzers.

Immunosensors have been utilized for a variety of *in vivo* and *in vitro* applications. Although most immunosensors are based on electrochemical detection, potentiometric, voltamperometric, thermometric, optoelectronics, acoustoelectronic and other methodologies have been described (108,109,110,111,112,113,114 and 115). Another immunoassay detection system monitors antigen or antibody binding on a surface by surface plasmon oscillations (116).

Several new membranes, fibers, and other materials are being used as solid phases in immunoassays. In radial partition immunoassays, antibodies are immobilized to glass-fiber paper, and enzyme activity is measured fluorometrically (117). In one assay for the detection of hCG, the capture antibody is immobilized to a nylon membrane in contact with an absorbent material (118). Both reference and test zones are included. Another popular assay for pregnancy detection uses the novel approach of microparticle-capture on glass-fiber membranes. This technique is based on the principle that coated polystyrene microparticles irreversibly adhere to the surface of glass fibers or glass-fiber membranes. Microparticles coated with anti-hCG antibodies are applied to a glass filter along a narrow line (test line), whereas hCG-coated microparticles are fixed along an intersecting horizontal line at right angles to the first (reference line). The sample is first passed through the filter, followed by the enzyme-antibody conjugate and substrate. If the reagents are functional, color is formed along the reference line to give a - sign. If hCG is present in the sample, the test line is also colored, giving a + sign.

Fluorescence and Chemiluminescence

Coon and associates were the first to use fluorescent dyes (fluorochromes) as a label for an antibody molecule. Since that time, they have proven to be sensitive and versatile labels for the detection and quantitation of antigens and antibodies in body fluids and tissues. Fluorochromes are easily conjugated to most substances of biological importance and are easily measured in the laboratory. In addition, some fluorescent probes can be coupled to other excited states or chemical species by energy transfer or chemical reactions.

Fluorescence and Phosphorescence

The emission of light is one possible outcome of the interaction of a molecule with light or other electromagnetic radiation. In response to the absorption of energy, fluorochromes enter an excited state, during which changes in electron configuration occur. However, the molecules are unstable in the excited state and undergo a rapid return to the stable or resting state. During this transition, most of the absorbed energy is dissipated in the form of heat, but some is released (emitted) as a photon of light (luminescence). There are two types of luminescence, depending on whether the time between excitation and emission is short (less than 10^{-8} seconds, fluorescence) or long (more than10^{-4} seconds, phosphorescence). Because some energy is lost as heat, the wavelength of the emitted energy is longer (lower energy) than that of the wavelength of the excitation source (higher energy) (Fig. 61.18).

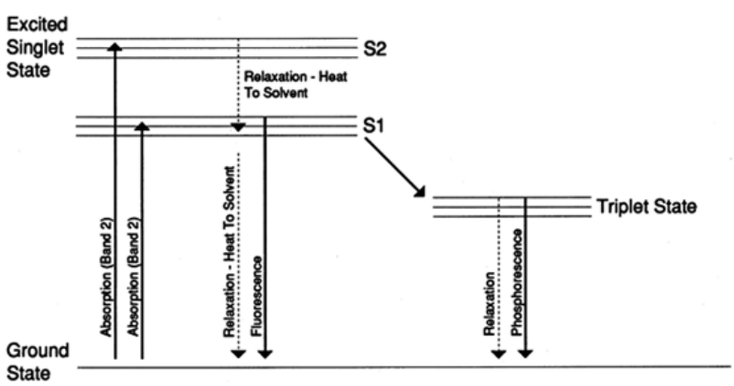


FIGURE 61.18. Energy transitions during fluorescence and phosphorescence. The absorption of light causes a transition of the molecule from its ground state singlet (S0) to one of a number of excited singlet states (e.g., S1, S2). The molecule is electronically unstable in the excited state, and the excess energy is dissipated by nonradiative means (heat) or radiative energy transfer (luminescence) during a return to the ground state. Fluorescence occurs if the molecule undergoes a direct radiative transition to the ground state, whereas phosphorescence results in cases in which the transition involves a series of semistable triplet states (e.g., T1, T2). Nonradiative transitions are indicated by straight arrows and radiative transitions by wavy arrows. (From Hemmila, I. Fluoroimmunoassays and immunofluorometric assays. *Clin Chem* 1985;31:360, with permission.)

Each fluorochrome is characterized by a number of parameters,

including the maximum wavelengths of maximal excitation (λ^{Abs}_{Max}) and emission (λ^{Fluor}_{Max}), the Stokes' shift, the extinction coefficient (ε) in the region of excitation, the quantum yield (Q), and the decay rate of the excited state (τ). The Stokes' shift is the difference between λ^{Abs}_{Max} and λ^{Fluor}_{Max}. It is usually 20 to 50 nm for most fluorescent compounds but may be as great as 200 nm for some phosphorescent compounds (Table 61.5). The maximal intensity of fluorescent emission occurs when a fluorochrome is excited with light at the wavelength of the excitation maximum. Decreases in fluorescent intensity occur at wavelengths above and below the maximum excitation wavelength. In practice, the excitation maximum cannot be utilized under two circumstances: (a) the light source does not provide an emission line of the proper wavelength and (b) two or more fluorochromes with different excitation maxima are simultaneously used with a single monochromatic excitation source. Under these circumstances, compromises can be made as long as the intensity of the signal emitted is adequate for particle detection. Fortunately, the intensity of the fluorescent emission is always proportional to the amount of fluorochrome present at a fixed wavelength. Optical

TABLE 61.5. PROPERTIES OF FLUOROCHROMES

Type of Probe	Name of Probe	Absorption Maximum (nm)	Emission Maximum (nm)	Extinction Maximum ($M^{-1}cm^{-1}$)	Quantum Yield
Covalent labels					
	FITC-NH-CH3	490	520	67	0.71
	FITC-NH-Ab	490	520	67	0.1-0.4
	Phycoerythrin-R	480-565	578	1960	0.68
	Allophycocyanin	650	660	700	0.68
	Peridinin-chlorophyll	470	677	—	—
	TRITC-amines	554	573	85	0.28
	XRITC-NH-CH3	582	601	79	0.26
	XRITC-NH-Ab	580	604		0.08
	Texas Red-amines	596	620	85	0.51
	Texas Red-NH-Ab	596	620	85	0.01
Extrinsic RNA/DNA probes					
	Propidium iodide	536	623	6.4	0.09
	Ethidium bromide	510	595	5.2	—
	Hoechst 33342	340	450	120	0.83
	DAPI	350	470	—	—
	Acridine orange (DNA)	480	520	—	—
	Acridine orange (RNA)	440-470	650	—	—
	Pyronine Y (dsDNA)	549-561	567-574	67-84	0.04-0.26
	Pyronine Y (dsRNA)	560-562	565-574	70-90	0.05-0.21

FITC, fluorescein isothiocyanate.

filters are utilized in laboratory instruments to maximize parameters for a particular fluorochrome or combination of fluorochromes.

The chemical and biological properties of a fluorochrome are influenced by the chemical composition of the substance, by the physical and chemical environment in which the substance is placed, and by a variety of other factors. In addition to a large extinction coefficient (ε) in the region of excitation, high quantum yield, an optimal excitation wavelength, and a proper excited state lifetime, fluorochromes must be photostable, biologically inert, and undergo minimal interference by other substances that may be present (quenching) (119). The compound must also possess functional groups for binding to an antigen or antibody. Isothiocyanates, chlorotriazine derivatives, or hydroxysuccinimido active esters covalently bind to the primary amino group of proteins, whereas iodoacetamido and maleimido groups attach to sulfhydryl groups (119). In conjugates of a fluorochrome with a protein antigen or antibody, the molar fluorescein-to-protein ratio (F/P ratio) is important. The F/P ratio is an indication of the relative number of fluorochrome molecules per protein molecule. Staining sensitivity is reduced if the F/P ratio is too small, whereas nonspecific fluorescence can be caused by an F/P ratio that is too high.

Although no single substance has all of the properties of an ideal fluorochrome, FITC, rhodamine, and phycoerythrin are employed for most applications at present. The properties of these and other fluorochromes of biological interest are summarized in Table 61.5.

FITC is the most widely used fluorochrome. FITC is excited at a wavelength of 492 nm and emits light at a wavelength of 510 nm (Fig. 61.19). This substance binds neutral amino acids and has been used as a tag for antibodies, hormones, lipids, lectins, and a wide variety of other biological molecules. The advantages of FITC are its solubility in water, high quantum efficiency, and large extinction coefficient, whereas its disadvantages include a pH dependence of fluorescent emission, moderate photoinstability, and an excitation wavelength less than 500 nm (119). Between two and five fluorescein molecules can usually be attached to most biological molecules before quenching limits fluorescent intensity. Covalent binding requires an alkaline pH (pH 8.5 to 9.0) at which the pH amino groups are in a reactive form.

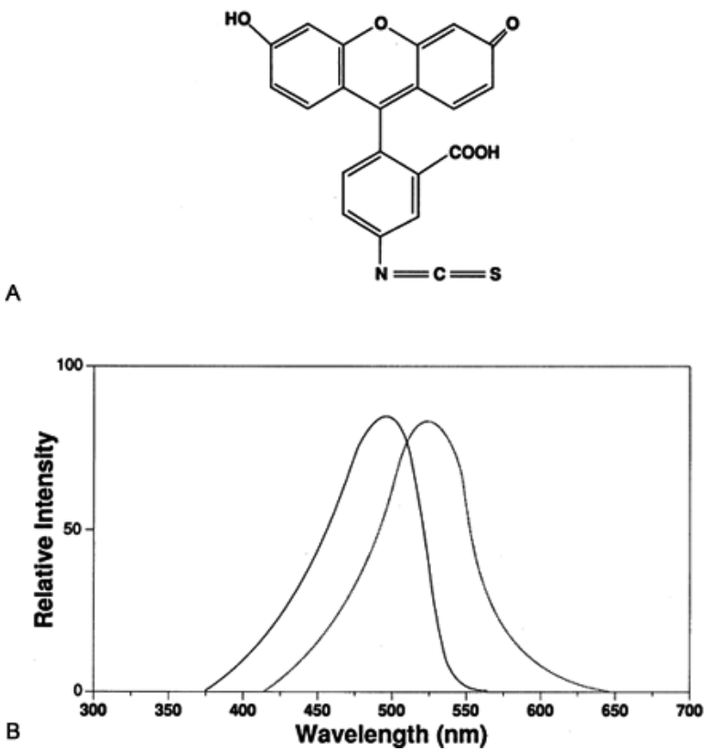


FIGURE 61.19. Molecular structure **(a)** and fluorescence spectra **(B)** of fluorescein isothiocyanate.

Rhodamine derivatives (e.g., rhodamine isothiocyanate, Texas red) have higher fluorescent excitation and emission wavelengths than FITC and are often used in conjunction with FITC for multiparametric analysis. Most of these compounds are less water soluble than FITC but are photostable and have fluorescent properties that are independent of pH (119).

Research in the early 1980s led to the development of a new class of fluorochromes that have been termed phycofluors (119,120,121,122 and 123). The major phycobiliproteins include the phycoerythrins (R-phycoerythrin, B-phycoerythrin), R-phycocyanin, and allophycocyanin. The advantages of these compounds include extremely high absorbance coefficients over a wide spectral range, very high quantum yields, constant fluorescence over a broad pH range, solubility in aqueous solution, and environmental stability. In addition, phycobiliproteins contain numerous lysyl side chains that permit easy conjugation to biological molecules that can confer binding specificity (immunoglobulins, lectins, protein A, avidin) (121,122).

Fluorescence Immunoassays

The use of fluorochromes for the identification and localization of antigen, antibodies, and immune complexes in histochemical, cytochemical, and cellular assays is described below. More recently, fluorochromes have been used in the clinical laboratory to quantitate antigens and antibodies in a manner analogous to RIA and EIA (124,125). Unfortunately, these assays have achieved limited application because of the need for specialized instrumentation, as well as extensively purified and well-characterized reagents. In addition, inherent limitations in the detection of fluorescent emissions limit the sensitivity of conventional fluorometric methods to the nanomolar range and above (83). These limitations include background fluorescence (particularly from autofluorescent substances normally present in serum), nonspecific binding of the reagents, fluorescence quenching, difficulties in differentiating the excitation and emission signals, and light scattering phenomenon. Fortunately, new fluorescent reagents and new techniques of applying these labels have alleviated many of these problems. Particularly important has been the use of some heavy metal compounds (i.e., lanthanide complexes) in conjunction with techniques such as time-resolved fluorescence immunoassay. In addition, the ability to chemically alter the properties of fluorochromes has resulted in great interest in the adaptation of fluorescent compounds for clinical laboratory diagnosis. Various biological dyes and even dyes utilized in the textile industry have been targets for these investigations as the parent compounds for synthesis of new fluorochromes. Because fluorochromes can act as substrates in biochemical reactions, these fluorogenic substrates have been extremely useful in

the study or quantitation of cellular properties (e.g., pH, membrane potential) or biological activity (e.g., enzyme activity, activation). The utilization of a compound as a fluorogenic substrate requires that it undergo a change in fluorescent properties (e.g., gain or loss of fluorescence, change in) with a change in the biological property or substance being studied.

Analytical fluorescence assays are most commonly classified as heterogeneous or homogeneous, based on the need for separation of the bound and free reagent. Other considerations include the type of reaction mechanism (competitive or noncompetitive), labeled component (ligand or antibody).

A solid phase is utilized in most heterogeneous fluorescent immunoassays for separation of the labeled and unlabeled reagents. Both competitive and noncompetitive reaction mechanisms have been described. Quantitation of the reaction mechanism is usually performed with a fluorometer.

Homogeneous immunoassays are rapid and easy to perform, although their sensitivity is limited compared with heterogeneous assays unless special instrumentation is utilized (typically 10^{-10} molar of analyte per liter) (126). In addition, these assays require highly purified, labeled antigen and specific antibody.

The quenching or enhancement of fluorescence of a ligand-labeled chromophobe caused by the binding of an antibody has been used to quantitate some substances. The enhancement fluoroimmunoassay for thyroxine described by Smith (127) utilizes a fluorescent derivative of thyroxine that undergoes an enhancement in fluorescent intensity on the binding of an antithyroxine antibody. In a competitive binding assay, unlabeled thyroxine in the reaction mixture competes for the antibody. With a sensitive fluorometer (Aminco Bowman SPF fluorometer), the assay has sensitivity equivalent to the standard RIA for thyroxine.

The phenomenon of energy transfer between fluorochromes has been incorporated into clinical analytical assays for multivalent antigens (128). Most commonly, fluorescein is used as the donor fluorochrome and rhodamine is the acceptor. Because the maximal fluorescence emission of fluorescein occurs at a wavelength of 525 nm, and tetraethyl- and tetramethyl-rhodamine have a strong absorption line at this wavelength, energy transfer can occur from fluorescein to rhodamine. Energy transfer occurs through dipole-dipole interaction at a rate that is inversely proportional to the sixth power of the distance between the molecules. Two types of fluorescence excitation transfer immunoassays have been described. The antibody is labeled with rhodamine in both types of assays. However, in one type, the antigen is directly labeled with fluorescein, whereas in the other type an indirect label (FITC-labeled antibody) is used. In the direct assay the fluorescein-labeled antigen and rhodamine-labeled antibodies are mixed together, causing a quenching of fluorescein fluorescence. When the unlabeled antigen is added to the reaction mixture, competition for binding sites on the antibody occurs, resulting in liberation of some of the fluorescein-labeled antigen, decreased quenching, and increased fluorescein fluorescence. The principle of the indirect assay is similar.

A fluorescent protection assay has been described for the analysis of protein antigens (e.g., immunoglobulin, hormones, serum proteins) or antibodies. In the assay for an antigen, the reactants include a fluorescein-labeled protein antigen, an antibody specific for the antigen, and a second antibody specific for fluorescein. The principle of the assay is that (a) binding of the anti-fluorescein antibody to the labeled protein results in the quenching of fluorescence, (b) the antigen-specific antibody sterically inhibits binding of the anti-fluorescein antibody, and (c) unlabeled antigen competes with the fluorescein-labeled antigen for antibody binding. Thus, fluorescence intensity is proportional to the quantity of unlabeled antigen.

The substrate-labeled fluorescent immunoassay has been used for the quantitation of haptens (including many therapeutic drugs), as well as immunoglobulins and other proteins. A complex of antigen and a fluorogenic substrate of an enzyme forms the basis of this assay. The principle is as follows: (a) the antigen is cleaved from the substrate complex by an appropriate enzyme, resulting in the formation of a fluorescent product, (b) the presence of an antibody to the antigen inhibits enzymatic cleavage, (c) free antigen inhibits antibody binding, (d) the rate of production of fluorescence is proportional to the free antigen concentration. In practice, umbelliferyl-ß-galactoside is usually utilized as the labeled enzyme substrate, and ß galactosidase is the enzyme. For many antigens, the sensitivity of this assay is comparable with that of RIA.

Particle concentration fluorescence immunoassays use small polystyrene beads as the solid phase. The reaction is performed in microtitration plates in which a 2-mm diameter porous membrane forms the bottom of each well. Specimens and reagents are removed from the wells by using suction. Total particle fluorescence is determined by front surface fluorescence. The utilization of this system for the quantitation of serum immunoglobulins and antibodies has been reported (129).

Fluorescence Polarization Immunoassays

Fluorescence polarization is a special property of fluorochromes that was first applied to an analytical assay by Dandliker et al. (130,131). Light is comprised of electric and magnetic fields (vectors), and polarization is defined by the orientation in time and space of the electric vector (132). If the electric vector can be considered to consist of two orthogonal components, the state of polarization of a light beam is defined by the relative amplitudes and phases of the orthogonal components. The orthogonal components are of equal magnitude and have no fixed relationship in the case of unpolarized (incoherent) light. In contrast, perfectly polarized light is coherent, and the orthogonal components are fixed in their relationship to each other (132).

Light produced from fluorescence emission is noncoherent because it arises from the summation of uncorrelated individual emission events (132). However, polarized light is absorbed by a chromophore in the same manner as nonpolarized light, provided that the absorption transition moment of the chromophore is in the same direction as the plane of polarization. The rate of depolarization can be determined by measuring the sample with detectors parallel and perpendicular to the plane of polarization (132,133).

The degree of polarization increases with restricted motion of a chromophore owing to increased viscosity of the medium or decreased fluidity of the membrane of other structural components. Therefore, small fluorochrome-labeled molecules in an aqueous solution have unrestricted motion and exhibit low polarization. However, when combined with an antibody, the

molecular motion of the fluorochrome is decreased, and the degree of fluorescence polarization increases.

Fluorescence polarization assays have been primarily utilized in the clinical laboratory for the analysis of small molecular weight substances, such as drugs and hormones. In one technique, polarization fluoroimmunoassay, the increased signal associated with binding of an antibody to a fluorophore-labeled antigen is measured (131). Assays for gentamicin, phenytoin, amphetamine, and other substances have been described (134,135,136 and 137). The recent advent of simpler, less expensive instrumentation for the laboratory may increase the use of fluorescence polarization.

The excited-state lifetime of a fluorescent compound has been applied in a type of fluorescent immunoassay termed time-resolved fluorescence immunoassay. These assays use fluorochromes with relatively long decay times, especially the rare earth metal chelates (10 to 10,000 milliseconds). In time-resolved fluorescence immunoassay, a rapidly pulsing excitation source is used, and fluorescence intensity is measured at a fixed interval after excitation. Because the decay time of substances that may cause background fluorescence is usually less than 10 nanoseconds, the inherent sensitivity of these methods is much greater than conventional fluorescent immunoassays (83,138). The rare earth metal europium has been most extensively used as a label in time-resolved immunoassays, although other trivalent lanthanide ions, such as terbium and samarium, have been used as well (138). Fluorometers for time-resolved fluorescence immunoassay are modified to measure only a portion of the emission cycle.

Particle-based flow cytometric fluorescence immunoassays have been described. For example, fluorescence quantitation by flow cytometry has been applied in an equilibrium-type competitive-binding fluorescence immunoassay (139). In this assay, relatively large (10 μm diameter) antibody-coated nonfluorescent particles were used with very small (0.10 μm diameter) antigen-coated fluorescent latex particles. During incubation with the sample, soluble unlabeled antigen competes with the small, antigen-labeled particles for binding to the larger spheres. The fluorescence distribution of 5,000 of the large spheres is determined by flow cytometry and read against a standard curve prepared with known concentrations of antigen. The sensitivity of this technique was 10^{-12} mol/L for quantitation of the antigen HRP, and the sensitivity was increased to 10^{-12} mol/L in a double-antibody sandwich assay. Lindmo et al. (140) described a different flow cytometric technique using two particle types coated with antibody of the same specificity but different affinity.

Chemiluminescent Immunoassays

Chemiluminescence is the emission of light after energy is transferred to the chemiluminescent molecule via a chemical reaction (141). Chemiluminescence and chemiluminescence energy transfer have been used for the detection of a variety of substances in the 10^{-18} to 10^{-21} mol/L range (108,142,143,144,145,146 and 147). Most of these assays utilize synthetic chemiluminescent compounds such as luminol (3-amino-phthalhydrazine), luminol derivatives such as isoluminol, or acridinium ester derivatives as labels for antigens or antibodies. Photon emission is measured in a luminometer or liquid scintillation counter after the addition of an oxidizing agent (e.g., H_2O_2) and a catalyst (e.g., hemin, lactoperoxidase, copper or cobalt ions). Although the low efficiency (less than 1%) of photon emission limits the sensitivity of assays based on chemiluminescence, the stability of the reagents and the low background activity are advantages. Chemiluminescence energy transfer has been used to increase the sensitivity of the assay (145).

Interferences in Immunoassays

Several molecules can cause false-positive or false-negative reactions in immunoassays (23,148,149). Commonly considered is the phenomenon known as cross-reactivity, in which an epitope or epitopes on one molecule have enough similarity with the analyte that they are recognized by the antibody or antibodies used in the assay. Early assays for TSH, for example, had significant cross-reactivity with similar hormones with identical α and similar β chains, such as luteinizing hormone, follicle-stimulating hormone, and hCG. More specific antibodies utilizing two-site principles have largely eliminated this problem for these hormones, but for some drugs and steroid hormones this is still an important consideration.

High-Dose Hook Effect

For immunometric assays, when an analyte is present in extraordinarily high concentration, it is possible to exceed the binding capacity of the immobilized capture antibodies. When this occurs in a one-step assay, the analyte can then react with the labeled antibody, preventing the formation of some of the sandwich complexes that would usually form, thus decreasing the signal measured and the apparent analyte concentration. A signal versus concentration plot reaching such high analyte levels actually begins to trail off or “hook” when high enough concentrations are reached (150). The high-dose hook effect is named for this characteristic curve. Manufacturers of diagnostic kits typically generate this graph, then make a claim in the package insert that samples with analyte concentrations up to X will not have the high-dose hook effect (Fig. 61.20). There is usually an error code generated by the instrument when a value is above the upper limit of reportability but below the level at which the hook effect will not interfere with the assay, alerting the operator that the sample needs to be diluted and rerun. Note that in a two-step assay, the signal reaches a maximal level but never decreases with increasing concentration. Thus, this type of assay format does not exhibit this hook effect.

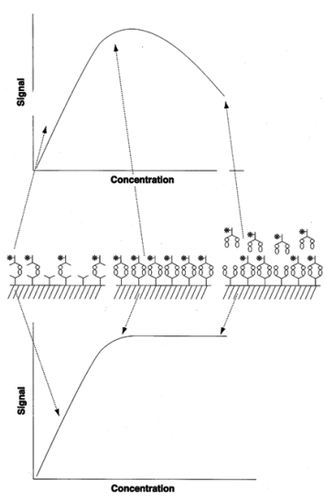


FIGURE 61.20. Principle of high-dose hook effect.

Heterophile Interference

This topic has been very well presented by several authors in great detail. The basic issue is that antibodies can be formed by humans, for a variety of reasons, that interact with the antibodies that are used as reagents in various assays. The most commonly described variant of this is referred to as HAMA or human anti-mouse antibodies. These HAMA can react with mouse monoclonals used in many assays and, depending on the assay configuration, can cause a false-positive or a false-negative result. Although HAMA is generally illustrated and thought of as a positive interference, an interesting report describing the development of a HAMA-resistant carcinoembryonic antigen assay found that both positive and negative interferences could be found in different patient specimens in the same assay (151).

Automation of Immunoassays

The most significant advance in immunoassay progress during the past decade has undoubtedly been in the area of automation, specifically random access automation. Before the 1990s, instrumentation had been available to perform assays in a batch mode, i.e., the instrument could analyze samples for only one analyte at a time and did not allow samples to be added after the run had begun. Random access automation, however, allows the analysis of multiple analytes on multiple consecutive samples without the need to pause between analytes or samples. This has allowed a tremendous increase in productivity for technologists, greatly reduced turnaround times for many analytes, and improved stability as well. Calibration and on-board reagent stability of greater than 1 month are now commonly achieved. Assays adapted to such platforms include drugs, hormones, specific proteins, tumor markers, and antibodies to infectious agents. The emphasis on random access immunoassay systems is owing to manufacturers' solving the problem of automating an effective wash step and configuring the assays on their instruments in such a way that incubation times were either constant for most assays or that variable incubation times did not interfere with throughput.

This emphasis on automation has been accompanied by the emphasis on enzymes as labels for immunoassays. Although fluorescent or colorimetric substrates were often used in early systems, chemiluminescent substrates are commonly employed with new instruments today.

CELLULAR IMMUNOASSAYS

Part of "61 - Basic Principles of Immunodiagnosis"

The development of assays for the identification and enumeration of immune cells and for the assessment of cell function has been essential for the understanding of the immune system. In addition to applications in organ transplantation, these assays have been useful in the diagnosis and treatment of patients with malignant neoplasms, autoimmune diseases, and infectious diseases. The major technical developments in cellular immunology have included:

- *In vitro* cytotoxicity assays
- The discovery and utilization of anti-HLA antisera from multiparous women in histocompatibility testing
- Cryopreservation and thawing of lymphocytes
- Assays of cell-mediated immune function (mixed lymphocyte culture, antibody-dependent cell-mediated cytotoxicity)
- Cellular typing using cell-mediated lympholysis (CML), homozygous typing cells, and primed lymphocyte typing
- Miniaturization, automation, standardization, and improved quality control
- Monoclonal antibody technology, flow cytometry, and immunochemical analysis
- Gene cloning, DNA sequence analysis, and recombinant DNA technology

Basic Techniques in Cellular Immunoassay

Specimen Collection and Cell Preparation

Specimen Collection and Transportation

A successful cellular assay begins with the procurement of the proper specimen in the correct manner with prompt transportation of the specimen to the clinical laboratory. Care must be taken that the person obtaining the specimen is correctly informed about specimen requirements for the assay that has been ordered and that any special instructions for obtaining the specimen are followed. For most cellular immunoassays, 5 to 20 mL of anticoagulated venous blood is required, although cellular immunoassays may require serum or plasma, other body fluids, or lymphoid tissue. Because immune cells and many immune substances are labile, the specimen should be received within the laboratory and processed within 24 hours. Most specimens are maintained at room temperature during transportation.

Cell Separation and Purification

The peripheral blood, lymph nodes, spleen, and other components of the immune system consist of a heterogeneous mixture of cell types (lymphocytes, monocytes, granulocytes, macrophages, RBCs, and platelets). However, most cellular assays assess the properties or function of one particular cell type and cannot be performed in the presence of large numbers of contaminating cells. Purified lymphocyte preparations are preferred

in many cellular assays because they strongly express HLA class I and II antigens, whereas monocyte cell preparations are utilized for many cellular assays, but other assays require isolated granulocytes, platelets, or even a purified subset of a particular cell type. Some techniques for the isolation and purification of cells are based on differences in basic physical properties (e.g., size, density, granularity), and others rely on functional differences (e.g., adherence, phagocytosis), or the presence of a particular cell surface antigen (152).

Whole blood leukocyte preparations are adequate for many purposes, including the enumeration of lymphocyte subsets by flow cytometric immunophenotypic analysis. The direct removal of RBCs from peripheral blood by ammonium chloride lysis is being increasingly utilized for routine lymphocyte enumeration studies. This technique is efficient, rapidly performed, and minimizes specimen handling. In addition, the selective loss of some lymphocyte subsets associated with density gradient centrifugation is avoided.

Cell density is the conventional property used for the separation of mononuclear cells (lymphocytes and monocytes) from other cell types. In cell isolation procedures using differences in cell density, the unpurified cell suspension is centrifuged in a solution containing a single or multilayer density gradient. After centrifugation, cells are distributed in the solution in layers based on differences in their density. A Ficoll-Hypaque gradient is most commonly used for this purpose (152,153,154,155,156,157 and 158). Ficoll-Hypaque is comprised of sodium diatrizoate (3,5-bis acetylamino-2,4,6 tri-iodo benzoic acid, hypo-opaque, iso-opaque) and Ficoll. Ficoll is a high molecular weight sucrose polymer (specific gravity, 1.076 to 1.078), whereas Hypaque is a dense iodinated organic compound also used as a radiographic contrast medium. Ficoll contributes viscosity to the solution and promotes rouleaux formation of the RBCs, whereas Hypaque increases the viscosity of the solution. When properly prepared, a Ficoll-Hypaque solution has a specific gravity of 1.077 at room temperature and is more dense than lymphocytes, monocytes, and platelets but less dense than granulocytes and RBCs.

In practice, blood for mononuclear cell isolation is collected aseptically into a heparinized container, a buffy coat is prepared and layered over or under Ficoll-Hypaque. Mononuclear cell preparations are prepared from tissue specimens by mincing the tissue into a tissue culture media [usually Roswell Park Memorial Institute (RPMI) medium containing 5% fetal calf serum, RPMI/5% fetal calf serum] or by gently forcing the tissue through nylon mesh. Wire or nylon mesh of various grades is also useful in removing cell aggregates and removing stroma and extraneous pieces of tissue from a cell preparation. Cell suspensions prepared from lymphoid tissue usually contains few erythrocytes and can be directly placed on Ficoll-Hypaque. After a one-step centrifugation with Ficoll-Hypaque, platelets (specific gravity, 1.040) and plasma (specific gravity, 1.025 to 1.029) are located above the Ficoll-Hypaque, lymphocytes (specific gravity, 1.070) and some platelets are found at the plasma-Ficoll-Hypaque interface, and granulocytes (specific gravity, 1.087 to 1.092) and RBCs (specific gravity, 1.093 to 1.096) form a pellet at the bottom of the tube (Fig. 61.21). Residual platelets are removed by washing and a low-speed spin. Incubation of the buffy coat preparation at 37°C with carbonyl iron can also increase the purity of the final cell preparation because neutrophils and monocytes phagocytize the carbonyl iron, thus increasing their density and allowing better gradient separation. Isolated lymphocyte preparations prepared from fresh peripheral blood by Ficoll-Hypaque isolation should provide 1×10^6 lymphocytes per milliliter that are 90% to 100% viable and contain fewer than 5% granulocytes and monocytes. Granulocytes can also be obtained from a Ficoll-Hypaque separation by removing the supernatant, suspending the pelleted cells, and removing RBCs by the addition of an ammonium chloride solution.

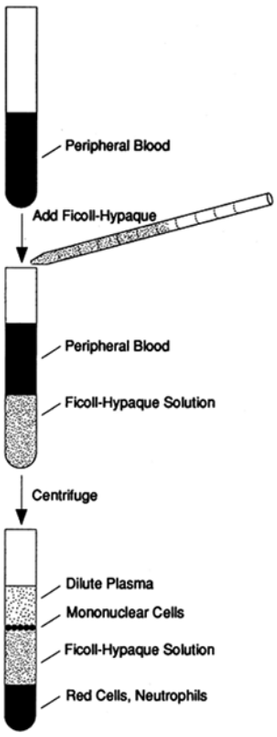


FIGURE 61.21. Technique of Ficoll-Hypaque density isolation of mononuclear cells. Whole blood is layered unto a Ficoll-Hypaque mixture with a specific gravity of 1.007, and the test tube is centrifuged. During centrifugation, the lymphocytes are separated from the other formed elements of the blood and form a discrete layer at the plasma-Ficoll-Hypaque interface, which can easily be removed with a pipette. (From from Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

Rosetting with SRBCs is an older method for the separation of T lymphocytes from other mononuclear cells (159,160). In this procedure, a monocyte-depleted peripheral blood mononuclear cell preparation is prepared by Ficoll-Hypaque or another technique. After incubation with SRBCs, the rosetted T cells are

separated from the B lymphocyte–enriched medium by density gradient separation. T cells can be isolated from the rosettes by ammonium chloride lysis of the RBCs. Pretreatment of the SRBCs with neuraminidase, 2-S-aminoethyl-isothiouromium bromide, or papain can be used to increase the purity of the final cell preparations. B lymphocytes can be rosetted with anti-human Fab-coated SRBC or ox erythrocytes coated with anti-B cell monoclonal antibodies (160).

The anti-$F(ab')_2$ monolayer technique for the isolation of B lymphocytes utilizes polystyrene flasks absorbed with affinity-purified anti-human $F(ab')_2$. Surface immunoglobulin-positive B lymphocytes in a monocyte-depleted peripheral blood mononuclear cell preparation adhere to the surface of the flask. Nonadherent T lymphocytes are washed away, whereas the adherent B lymphocytes are eluted through the use of immunoglobulin-rich serum (161,162). This technique results in highly purified T and B lymphocyte preparations but requires more technical expertise to establish and quality control.

Immunomagnetic separation techniques are increasingly replacing older cell separation techniques, particularly when a cell population with a selectively enriched cell subpopulation is needed (163,164). This technique utilizes detachable immunomagnetic microspheres labeled with monoclonal antibody(s). The initial cell suspension is first incubated with the microspheres, then a magnetic field is applied to separate the microspheres with attached cells. The cells are removed from the microspheres by washing.

Assessment of Cell Purity and Viability

Adequate cell purity and viability are essential for most cellular assays. In this regard, dead cells can cause false-negative reactions by nonspecifically absorbing antibodies and lowering antibody titer. In addition, false-positive reactions may also occur in assays using cell death as the end point (i.e., microlymphocytotoxicity) because dead cells previously present in the assay cannot be differentiated from cells killed during a microlymphocytotoxicity assay. For histocompatibility assays, 80% or more of the cells must be viable for accurate results (165). Assays using purified cell preparations require fewer than 20% contaminating cells.

Generally, cell preparations of adequate viability can be obtained from peripheral blood specimens if they are processed within 8 to 10 hours after drawing. Bone marrow, body fluids and effusions, and tissue specimens must be processed as soon as possible. Cryopreservation or storage at reduced temperature in culture media is best if the specimen cannot be processed immediately. Enzymatic treatment can also be utilized to remove nonviable cells. DNase is commonly used for this purpose because it enzymatically digests dead cells but is harmless to living cells at the concentration used.

Cell viability is usually determined by light microscopy and the use of a vital dye, such as trypan blue or eosin, which is incorporated into dead cells but excluded from living cells (165,166). In this technique, an aliquot of the cell suspension is mixed with the dye solution and placed onto a hemocytometer. After settling, the nonstained viable cells are rounded and refractile, whereas dead cells are round and flattened. A minimum of 100 cells is examined and the proportion of viable cells is determined and expressed as the percentage of viability. Cell concentration can be adjusted at the same time. The use of fluorescent dyes provides a more accurate alternative to vital dyes. In the most recent variation of the fluorescent dye technique, dual-color fluorescence is measured by flow cytometry, using the dyes fluorescein diacetate and propidium iodide. Fluorescein diacetate is a nonpolar, nonfluorescent compound that readily permeates the cell membrane. However, in viable cells, it is rapidly converted to a polar, highly fluorescent (green emission) substance that is trapped within the cell. Propidium iodide is a highly polar compound that emits light in the red region of the spectrum and can only permeate cells with damaged cell membranes. In practice, the cell suspension is stained with solutions of both dyes, and the proportion of cells with high green fluorescence and low red fluorescence (viable cells) or low green fluorescence and high red fluorescence (nonviable cells) is determined by flow cytometry (167,168).

Purity assessment can be performed by morphologic examination with light microscopy or phase contrast microscopy or by immunophenotypic analysis with flow cytometry or fluorescent microscopy (165).

Cryopreservation and Thawing

Long-term storage of cells (months to years) is required for histocompatibility analysis and may be desirable for the storage of functionally intact cells from patients with malignant tumors or immunologic diseases. Controlled-rate freezing and preservation in the vapor phase of liquid nitrogen are most commonly used for long-term cell preservation (166,168,169,170,171,172 and 173). Dimethyl sulfoxide is usually used to prevent damage during the process. If carefully frozen and thawed, cells remain viable for indefinite periods of time (174).

Immunofluorescent Staining

Monoclonal antibodies specific for cell surface antigens are commonly utilized for the enumeration of mononuclear cell populations and subpopulations. FITC is the most popular fluorescent probe for labeling monoclonal antibodies, although rhodamine, phycoerythrin, Texas red, and other fluorochromes are also used. Of the five classes of immunoglobulin (IgG, IgM, IgA, IgD, IgE), monoclonal antibodies of the IgG class are most often used in clinical flow cytometric studies. In addition, the Fc portion of the antibody molecule is usually removed by pepsin treatment to reduce binding to Fc receptors, which are present in great numbers on many types of cells. Removal of the Fc portion of the molecule leaves the F(ab)2 fragment, which consists of the two F(ab) antigen-binding sites. Two different immunofluorescent staining techniques (direct and indirect) are utilized, depending on the monoclonal antibodies available and the goal of the analysis.

Direct immunofluorescent staining is used when large quantities of a monoclonal antibody can be obtained and easily labeled with a fluorochrome. In this technique, the labeled monoclonal antibody is applied to the cell in a one-step staining process requiring the following steps (Fig. 61.22).

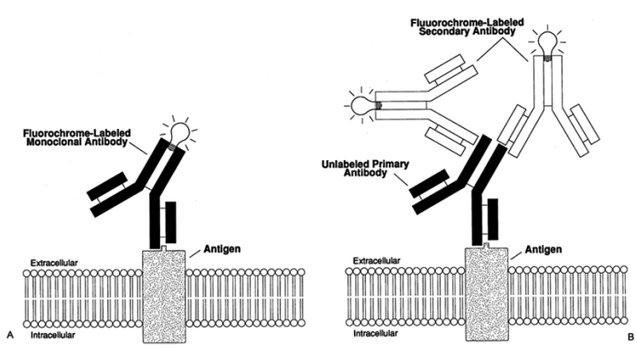


FIGURE 61.22. Principle of direct and indirect immunofluorescent staining. (From Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

- Mix cell suspension with FITC-labeled antibody solution
- Incubate
- Wash cells
- Analyze by flow cytometry

Nonspecific binding is not a problem if monospecific antibodies are used in a direct staining procedure. In addition, direct staining is rapid and easily performed and permits simultaneous staining with two or more labeled monoclonal antibodies (multiparametric analysis). Multiparametric analysis is more economical than analysis using a single labeled marker per cell (single color immunofluorescent analysis) and has the additional advantage of providing a means to assess the interrelationship between cell populations and analyze subpopulations of cells. Most commonly, two monoclonal antibodies, one labeled with FITC and the second with rhodamine or phycoerythrin, are applied to the cell suspension at the same time (two-color or dual-color immunofluorescent analysis). More recently, the simultaneous use of three monoclonal antibodies has become popular, in part because of the development of a fluorochrome consisting of a tandem conjugate of R-phycoerythrin plus Texas red. The excitation of this compound results in energy transfer from phycoerythrin to Texas red and the emission of light at a wavelength distinct from that of phycoerythrin. Specifically, the phycoerythrin component of the conjugate excites at 488 nm and emits light with a maximum of 575 nm. The phycoerythrin emission in turn excites adjacent Texas-red molecules, which emit red light with a maximum of 618 nm. In three-color analysis, one monoclonal antibody is labeled with FITC, a second with phycoerythrin, and the third with the tandem conjugate of Texas red and phycoerythrin. Unfortunately, the sensitivity of direct immunofluorescent staining is limited in circumstances in which a small number of surface markers are present. In addition, cell autofluorescence can interfere with the detection of the fluorescence emission from the fluorochrome, and labeled monoclonal antibodies are not yet available for every substance of biological interest. In systems in which two or more fluorochromes are simultaneously analyzed, differentiation of the separate emission signals requires more complex optical and electronic systems and the need for additional quality control.

In indirect immunofluorescent staining procedures, an unlabeled antibody (primary antibody) is first attached to the cell and then developed in a second step, utilizing a fluorochrome-labeled monoclonal antibody directed against the first monoclonal antibody (secondary antibody). Polyclonal FITC-labeled goat anti-mouse antibody is often used as the secondary antibody (Fig. 61.22). The following steps are required.

- Mix cell suspension with unlabeled monoclonal antibody
- Incubate
- Wash
- Apply FITC-labeled anti-mouse antibody
- Incubate
- Wash
- Analyze by flow cytometry

Indirect immunofluorescent staining is economical because a single FITC-labeled secondary antibody can be used with many primary monoclonal antibodies. However, it is technically more complex and time-consuming than direct staining and produces a relatively higher degree of background staining. Indirect staining is primarily used when a limited quantity of a monoclonal antibodies is available to avoid the loss of antibody associated with purification and conjugation and for the detection of cell surface antigens present at a low density. Multiparametric analysis using indirect immunofluorescent staining is impractical because

it requires the use of antibodies obtained from different species.

Removal of Autoantibodies

Autoreactive antibodies are commonly present in patients with connective tissue diseases, chronic liver disease, some renal diseases, chronic viral infections, and other diseases. In histocompatibility testing, the presence of autoantibodies (especially those directed against HLA class I and II antigens) may interfere with the detection and analysis of alloreactive antibodies. Autoreactive antibodies are detected by the inclusion of an autocontrol (patient's serum with patient's lymphocytes) in every histocompatibility assay. If the autocontrol is positive, attempted removal of the autoreactive antibodies by absorption with lymphocytes from the same individual is attempted and the assay is repeated. Another technique [dithiothreitol (DTT) treatment] is utilized when autoreactive antibodies cannot be removed by absorption.

Autoantibodies and the circulating pan–T-cell antibodies found in lupus patients are usually of the IgM class and are inactivated during treatment with DTT. The anti-HLA antibodies resulting from previous transplantation and/or multiple blood transfusions are usually of the IgG class and are not affected by DTT. DTT reduces the IgM molecule to its constituent units, leaving the IgG molecule intact. The finding of a positive HLA cross-match with both untreated and DTT-treated serum is consistent with the presence of anti-HLA antibodies. A positive cross-match with untreated serum, but not with DTT-treated serum, suggests the presence of autoreactive antibodies and is presumptive evidence for the absence of anti-HLA antibodies.

Cell Identification and Enumeration

Lymphocytes can be subdivided into stages of differentiation and into functionally active subsets, based on the cell surface antigens that are present. A revolution in immunology was brought about by the subclassification of lymphocytes into T and B cells based on the cell surface antigens and has now expanded to the analysis of other cells, such as monocytes, macrophages, myeloid stem cells, tumor cells. Clinically, cell surface analysis has been useful both in the detection of abnormal cells and in the analysis of normal cells that vary in different diseases. Specific clinical applications of cell surface marker analysis (immunophenotypic analysis) include:

- Determining the origin and stage of differentiation of lymphomas and leukemias
- Detecting early recurrence of hematologic malignancies
- Diagnosing and monitoring inherited and congenitally immunodeficient patients
- Chemotherapeutic monitoring
- Pre- and posttransplantation monitoring and evaluation

Rosetting

Rosetting techniques were originally used in the differentiation of T and B lymphocytes (159,160). Because T lymphocytes have a cell surface receptor (CD2) that binds SRBCs, aggregates of T lymphocytes and SRBC (E rosettes) are formed when the two cell types are incubated together. The presence of three or more SRBCs adhering to a lymphocyte is considered a rosette and evidence for identifying the cell as a T lymphocyte. Rosettes are visualized and counted under the light microscope, with the percentage of rosetting lymphocytes taken as the percentage of T cells. B lymphocytes can be enumerated with rosetting with antibody-coated SRBCs (EA rosette) or complement-coated SRBCs (EAC rosette). Unfortunately, these techniques are subjective, labor intensive, and inaccurate. In addition, the immature lymphocytes found in many types of acute leukemia and other hematologic malignancies may not possess the SRBC receptor or Fc receptors. Because of these problems, flow cytometry has become the method of choice for the quantitation of cell populations.

Assays of Cell Function

Lymphocyte Transformation Immunoassays

Lymphocyte transformation assays directly measure the ability of lymphocytes to respond to a stimulus. Because lymphocyte functional abnormalities can occur in the absence of changes in morphology or relative or absolute cell numbers, these assays are important in the evaluation of patients with congenital immunodeficiency or acquired immune defects as a result of viral infections or immunosuppressive therapy. In addition, the principle of lymphocyte transformation is used in different ways, including identifying lymphokines, detecting antigen sensitization, and detecting transplantation antigens.

Principles of Lymphocyte Transformation

The basis of all lymphocyte transformation assays is the detection of metabolic changes that occur in the lymphocyte during activation. In this regard, virgin or resting lymphocytes that encounter their specific foreign antigen or a specific chemical signal are triggered to proliferate and differentiate into an activated form that can carry out the predestined effector or regulatory functions of the cell. During this transformation process, T cells synthesize DNA, RNA, and protein, express HLA class II antigens, and synthesize and release lymphokines and other products. Morphologically, these changes are reflected by an increase in cell size, the acquisition of abundant basophilic cytoplasm, the appearance of multiple nucleoli, and changes in chromatin characteristics (lymphoblasts). Repeated mitotic divisions of the transformed cell occur, the progeny become progressively smaller, and the cell eventually resumes the size and appearance of the original unactivated lymphocyte.

The incorporation of tritiated thymidine (^{3}H-thymidine) has been the standard method for the quantitation of cellular proliferation (175). However, techniques utilizing fluorescent dyes have been found to provide equivalent sensitivity without the requirement of radioisotopes. In addition, flow cytometric analysis permits multiparametric quantitation of the cell populations under study and is more rapidly performed than radioisotopic determination. In one type of flow cytometric analysis, the mitogenic response of lymphoid cells, cellular activity in mixed lymphocyte cultures, and related parameters have been performed by measuring cell cycle progression with a DNA-specific

fluorochrome, such as propidium iodide (176). The use of a fluorescent-labeled monoclonal antibody against Ki67, a nuclear antigen that appears in proliferating cells, has been used for measuring lymphocyte proliferation (177), together with measuring the incorporation of an analog of uridine [bromodeoxyuridine (BrdUrd)] (178) and PKH26 dye, a fluorescent dye directly incorporated into the cell membranes of living cells (179).

Mitogen-Induced Blastogenesis

Mitogens or interleukin-2 are used to induce lymphocyte transformation. Of the mitogens, phytohemagglutinin and concanavalin A induce T blastogenesis, whereas pokeweed mitogen and staphylococcal protein A trigger B cell activation. In the past, the incorporation of tritiated thymidine has been the standard method for determining lymphocyte transformation. In this procedure, lymphocytes are isolated with Ficoll-Hypaque density gradient centrifugation and cultured for 2 to 4 days with a mitogen (or 5 to 7 days with an antigen or mixed allogeneic lymphocyte preparation) in a media enriched with human AB serum. Tritiated thymidine is added to the culture medium, and the plates are incubated for an additional 18 hours. Cells are isolated (harvested) from the plates using a special device (cell harvester) and transferred to disks of filter paper (Fig. 61.23). The disks are placed in scintillation fluid and counted for tritium in a liquid scintillation counter.

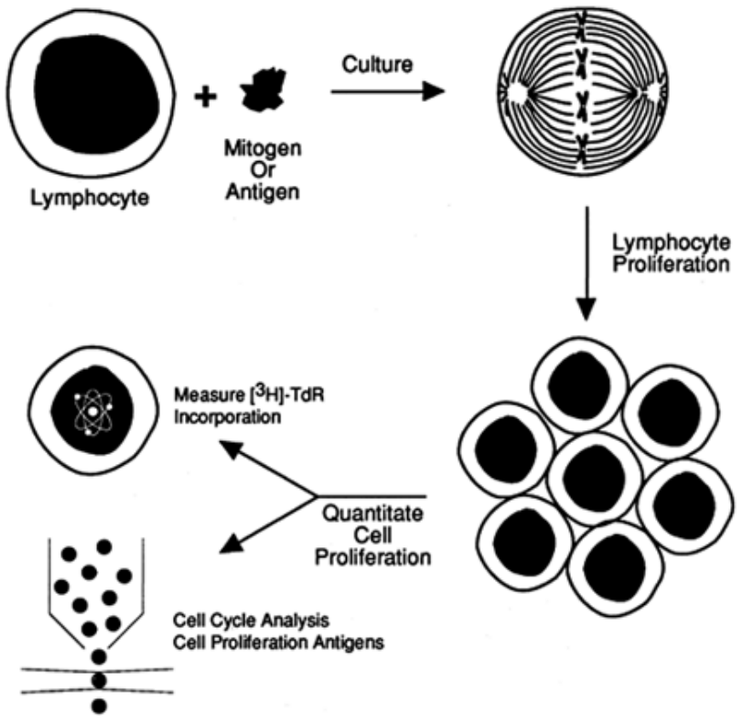


FIGURE 61.23. Principle of lymphocyte stimulation assay. Lymphocytes are cultured with antigen or mitogen. If stimulation occurs, a proliferative reaction takes place, with the formation of daughter cells. The amount of proliferation can be determined from tritiated thymidine (^{3}H-TdR) incorporation, by flow cytometric cell cycle kinetic analysis, or by measuring a cell proliferation antigen such as Ki67. (From Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

Several controls are necessary for lymphocyte proliferation assays because the assay may be influenced by the number of cultured cells, incubation time and temperature, mitogen concentration, and various inhibitory factors that may be present. The results are usually expressed as a stimulation index, that is, the ratio of radioisotope incorporation into the test and control cells. Alternatively, the relative proliferation index (RPI) can be calculated (175,180). The RPI is the ratio of the Dcpm (counts per minute of stimulated cells minus counts per minute of unstimulated control cells) of the test subject to the Dcpm of a panel of normal individuals that are tested simultaneously. The RPI is theoretically more accurate because it corrects for the large individual variations in blastogenic response. Specific soluble antigens may be used in place of a mitogen. Flow cytometric assays of blastogenesis have also been reported (177,181).

Mixed lymphocyte Culture

The mixed lymphocyte culture assay (MLC) is a special type of lymphocyte stimulation assay based on the ability of histoincompatible lymphocytes from one individual to stimulate the lymphocytes of another individual (mixed lymphocyte reaction) (182,183,184 and 185). Although complex in etiology, the MLC phenomenon is determined by the D locus of the HLA system. When two cells share common D loci, they are not able to stimulate each other, but when the D loci are different, the cells are stimulated. The MLC is unilateral (one way) when one group of cells is made incapable of responding (by treatment with radiation or mithramycin) or bidirectional (two way) when no radiation or mithramycin is utilized. The one-way MLC is almost exclusively used in clinical immunology. In the one-way MLC, the untreated cells are termed the responder population, and the treated cells are the stimulator population.

In the MLC, purified lymphocyte suspensions are prepared from the individuals to be tested as well as from controls. Responder cells are distributed into the wells of round-bottom microtiter plates. Stimulator cells are first irradiated with 3,000 rad using a cesium irradiator or other appropriate source of radiation and then added to appropriate wells of the microtiter tray. After an incubation period of 4 to 6 days in a humidified atmosphere of 5% CO_2 in air, ^{3}H-thymidine (usually 1 μCi) is added to each well, and the trays are incubated for an additional 18 hours. The

cells are then harvested onto glass or paper fiber filters, using a cell harvester with distilled water. After drying, the radioactivity present on each filter is counted (cpm or dpm) in a scintillation counter (Fig. 61.24).

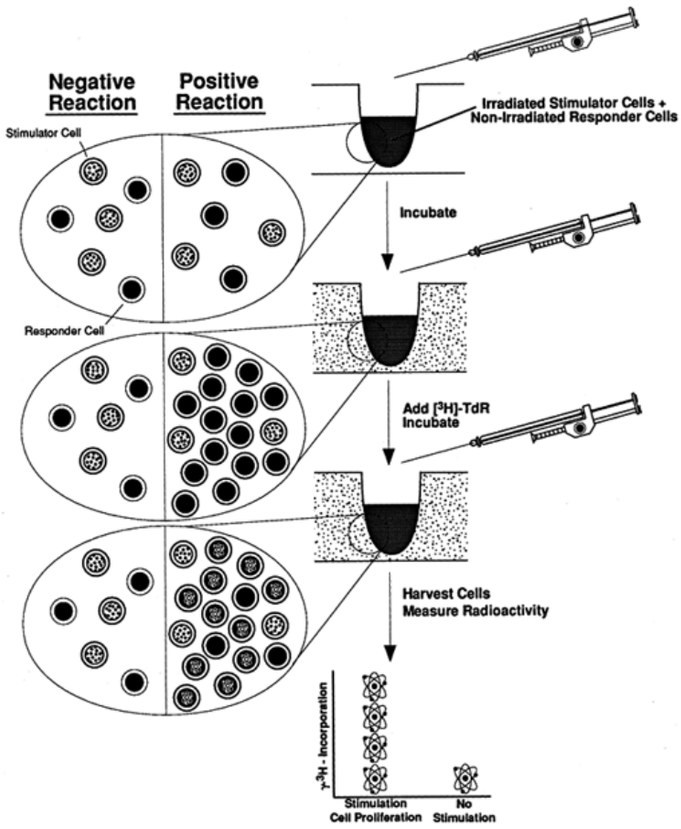


FIGURE 61.24. The one-way mixed lymphocyte culture assay. Irradiated stimulator cells, which are incapable of cell division, are mixed with nonirradiated responder cells, which can proliferate and divide. After incubation under cell culture conditions for approximately 1 week, the responder cells proliferate in the presence of a disparity in the HLA-D locus **(right)**, whereas no proliferation occurs if the responder and stimulator cells are HLA-D compatible **(left)**. Tritiated thymidine is added to the reaction mixture, and the degree of proliferation is determined by liquid scintillation counting. (From Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

Numerous factors must be considered in the performance and interpretation of the MLC. Because the reaction is complex, a lack of stimulation between two cell types cannot be considered conclusive evidence of histocompatibility at the HLA-D locus. Other minor histocompatibility loci may also be involved as well as some non-HLA loci. In addition, specific and nonspecific serum factors can inhibit the assay, and lymphocytes from diseased individuals may not perform well. Numerous controls (including autologous controls) must be included and must performed as expected for the results to be accepted as valid.

The primary use of the MLC assay in patient care is to assist in the selection of a compatible donor for a kidney or bone marrow transplant. The MLC assay is a predictor of host response to a transplanted organ. To prevent graft rejection or graft-versus-host disease, the donor and recipient cells must be mutually nonstimulatory. Practically, the use of the MLC is restricted to the selection of living related donors because the performance time of the assay is far longer than the organ preservation time limit. Another use of the MLC is to determine the HLA-D type. For this purpose, irradiated stimulator cells that are homozygous at

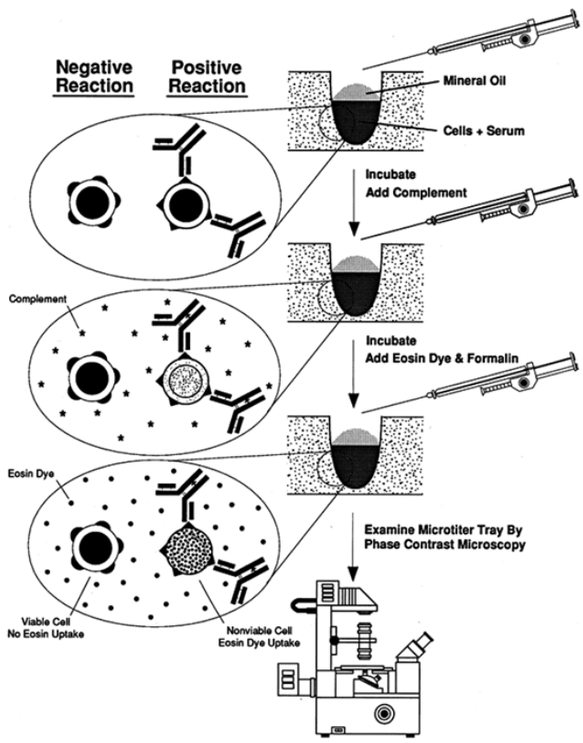


FIGURE 61.25. Principle of microlymphocytotoxicity by dye exclusion. Cells and serum are incubated together in the wells of a microtiter tray. Mineral oil at the top of the wells prevents fluid evaporation. Complement is added, and a second incubation is performed. The cells are washed, eosin is added, and the cells are fixed with formalin. Cells that fixed antibody and were killed through the action of complement take up of eosin and appear dark and nonrefractive under a phase-contrast microscope. In contrast, living cells are refractive and easily identified by phase contrast microscopy. Variation of the microlymphocytotoxicity assay are used for HLA typing, the HLA cross-match, and detection of anti-HLA antibodies. (From Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

the HLA-D locus are used in a one-way MLC. Stimulation suggests nonidentity of the HLA-D antigens.

Flow cytometric quantitation of the mixed lymphocyte reaction has been reported. In the method reported by Cram et al. (181) and by Kanda et al. (186), nuclear DNA content during lymphocyte blastogenesis was determined with a DNA-specific fluorochrome (propidium iodide). In this technique, the proliferative index (percentage of cells in the S phase and G_2M phase of the cell cycle) is a direct measurement of *de novo* DNA synthesis. A good correlation (r = 0.9493) was observed between the flow cytometric measurement of the MLR and the results using a traditional MLC with ^{3}H-thymidine incorporation. Although the flow cytometric determination is potentially more accurate than the MLC and avoids the use of radioisotopes, the flow cytometric technique has not achieved routine use in histocompatibility testing at present. Bontadini et al. (178) measured BrdUrd incorporation by flow cytometry to measure the mixed lymphocyte reaction. BrdUrd is an analog of thymidine against which a monoclonal antibody is available. BrdUrd incorporation is an extremely sensitive method for the detection of cycling cells because the incorporation of a few molecules of BrdUrd can be detected. This sensitivity permitted the detection of cell proliferation 24 to 48 hours before tritiated thymidine (^{3}H-TdR) incorporation.

Spontaneous Blastogenesis

Spontaneous blastogenesis assays determine the ability of cultured, unstimulated lymphocytes to undergo activation. The assay is usually performed by measuring the rate of ^{3}H-thymidine incorporation. Increased spontaneous blastogenesis in transplant recipients has been associated with allograft rejection.

Cytotoxicity Assays

Assays with cell death (cytotoxicity) as the end point are commonly used in cellular immunology. These assays are used in various ways to measure cell function activity and to detect cell surface antigens. In these assays, cytotoxicity may occur as the result of complement activity (complement-mediated cytotoxicity) or the direct effect of one cell on another (cell-mediated cytotoxicity). Until recently, target cell lysis was determined either by the release of a substance such as chromium 51 (^{51}Cr) from the cell on death or by the incorporation of a vital dye such as eosin or trypan blue. However, the accuracy and reproducibility of these assays have been compromised by problems such as poor uptake or nonspecific release of the marker, the length and technical complexity of the assays, or the need for the subjective interpretation of the results. The advent of new markers, in combination with the quantitative and multiparametric abilities of the flow cytometer, may make cellular assays much more practical for routine use in patient care. One example of these new markers is PKH-1, a fluorochrome that emits light in the green region of the spectrum, binds avidly to the cytoplasmic membrane, and is not transferred to other cells. In one application, PKH-1 is used to label the target cells, and a second fluorochrome with a red emission signal (propidium iodide) is used to detect nonviable cells (187). This combination of markers allows statistically valid quantitation of both dead and live effectors and targets, permitting the simultaneous determination of the percentage of target lysis, effector-to-target ratio, viability of the effector cells at the termination of the assay, and viable effector-to-target cell ratios. The use of the fluorescent dyes carboxyfluorescein diacetate and/or propidium iodide for the determination of lymphocytotoxicity has been described (187,188 and 189).

Microlymphocytotoxicity Assay

One type of cytotoxicity assay, the dye exclusion lymphocytotoxicity assay, is a commonly used procedure for the detection of an antibody-antigen interaction on a cell surface. The lymphocytotoxicity assay was introduced by Terasaki and McClelland in 1964 and was later accepted as the National Institutes of Health (NIH) standard procedure for histocompatibility testing. In the histocompatibility laboratory, variations of the lymphocytotoxicity assay are used for HLA typing, the detection of anti-HLA antibodies, and cross-match testing.

In the lymphocytotoxicity procedure, viable cells (usually lymphocytes) are incubated with serum containing antisera. If a cell surface antigen is present that antibodies in the sera recognize, an antigen-antibody complex will form on the surface. These complexes are detected by the sequential addition of rabbit complement and a vital dye, such as eosin, to the reaction mixture. The occurrence of complement fixation on the cell membrane leads to activation of the terminal complement components and eventually to cell lysis and death. Dead cells are detected and counted by phase microscopy after differential uptake of the eosin dye and fixation with formalin. Antibody-bound lymphocytes will die, take up the eosin dye, and give a positive reaction, whereas unbound lymphocytes will remain viable, exclude the eosin dye, and give a negative reaction (dye exclusion) (Fig. 61.25).

Microtiter trays are read under a properly adjusted inverted phase contrast microscope in a serpentine fashion, beginning at well 1A through 1F, continuing with 2F through 2A, 3A through 3F,…, 12F through 12A for a 72-well tray. Cells that are not injured appear small, bright, and refractile, whereas injured cells that have taken up the eosin dye are larger, darker, and nonrefractive and have a slightly granular surface. The result of the analysis is expressed as an estimated percentage of cells killed, taking background killing into consideration. To facilitate and standardize the evaluation process, a standardized system approved by the NIH is used. With this system, each well is assigned a score based on the number of cells killed (Table 61.6).

TABLE 61.6. SCORING SYSTEM FOR HISTOCOMPATIBILITY TESTING

Cell Death	Interpretation	Score
Unreadable well	—	0
0–10%	Negative	1
11–20%	Weak negative	2
21–50%	Weak positive	4
51–80%	Positive	6
81–100%	Strong positive	8

HLA typing is utilized to define the HLA-A, B, C, DR, and DQ locus antigens present on human T lymphocytes. In addition to selecting donors for organ transplantation, HLA typing is important in disease association, parentage testing, and the selection of donors for platelet or leukocyte transfusions. HLA typing is performed by lymphocytotoxicity using viable cells from the patient, and serum containing antisera of known specificity. Whole lymphocyte preparations or purified T-lymphocytes are commonly used for HLA class I antigen (HLA-A, B, C) determinations, whereas enriched B-lymphocyte suspensions are required for the determination of HLA-DR and HLA-DQ alleles. Currently, clinical laboratories do not routinely perform typing

for HLA-DP antigens because anti–HLA-DP antibodies are not generally available for microlymphocytotoxicity assays. The antisera used in HLA typing are commonly obtained from multiparous women and are commercially available on preplated trays. Antisera are also obtained through exchanges with other laboratories, regional sharing programs, and the NIH serum bank. Because some antisera are monospecific and others are polyspecific, it is necessary that each HLA antigen be defined by at least two monospecific antisera or by three antisera that are operationally polyspecific. Fluorescent procedures for HLA-DR

typing are preferred in some laboratories because of their reproducibility and technical simplicity. In the two-color fluorescent technique, mononuclear cells from an unseparated Ficoll-Hypaque purified mononuclear cell preparation are mixed with FITC-labeled anti-immunoglobulin antibody. After incubation for 5 minutes at 37°C, T lymphocytes are unstained, monocytes show speckled fluorescence, and B lymphocytes develop a fluorescent cap at one end of cell. After incubation of the labeled cells with anti-HLA serum and complement, the number of viable and nonviable cells in each population can be counted under an inverted phase fluorescent microscope (190,191). A simplified single-color version of this technique has been described that greatly reduces the time required to perform the assay (192).

The microlymphocytotoxicity assay is also used for the detection of anti-HLA antibodies and HLA cross-matching. The development of antibodies against HLA class I and II antigens is common in patients awaiting transplantation, who can be immunized by blood transfusions, previous organ transplantation, pregnancy, or a combination of these factors. Antibody screening is helpful in determining the cause of a positive cross-match and avoiding a similar incompatibility in the future. It is also helpful in identifying highly sensitized patients who are difficult candidates for transplantation. Sensitization against the HLA antigens also occurs in patients receiving multiple transfusions of blood products. The detection of anti-HLA antibodies (antibody screening) is performed with a panel of lymphocytes of known HLA type. Serum is added, and the lymphocytotoxicity assay is performed. Each antigen must be represented on at least two or three different panel lymphocytes because a technical problem with one cell could result in antiserum specificity being missed or misinterpreted. The results of the assay are expressed as the percentage of cells in the panel that give a positive reaction (percent panel reactive antibody, PRA). If anti-HLA antibodies are present, their specificity can usually be determined from the pattern of reactivity with the panel lymphocytes.

The third application of the microlymphocytotoxicity assay in the histocompatibility laboratory is in the HLA cross-match.

Lymphocyte-Mediated Cytotoxicity

The detection of cytotoxic effector killer cells generated *in vivo* is commonly performed by lymphocyte-mediated cytotoxicity (LMC). In this assay, ^{51}Cr-labeled target cells are incubated with a purified mononuclear cell preparation from the test subject. Killing of the target cells occurs by a direct, complement-independent mechanism, which results in the release of the ^{51}Cr into the medium (193).

The LMC has been utilized in histocompatibility testing and research studies of third-population cells. In organ transplantation, the LMC is used to detect the presence of preformed recipient killer cells specific for the potential donor. In this application, target cells are obtained from the peripheral blood of a potential living-related donor or from the lymph node and/or spleen cells of a cadaveric donor (194). Four-, 12-, and 16-hour LMC assays have been described.

Cell-Mediated Lympholysis

The LMC assay detects the *in vitro* generation of cytotoxic T lymphocytes against a histoincompatible-immunizing cell. In CML, isolated peripheral blood lymphocytes are incubated with stimulator cells in a 5- to 8-day mixed lymphocyte culture. The effector cells are then isolated and incubated with phytohemagglutinin-activated, nonirradiated stimulator cells that have been labeled with ^{51}Cr. Target cell lysis and ^{51}Cr release is proportional to the number of cytotoxic T lymphocytes (195,196). Appropriate controls must be utilized, and the data must be corrected for spontaneous ^{51}Cr release. Generally, different dilutions of effector cells are used to yield effector/target cell ratios of 100, 50, 25, 12.5, and 6.25. In addition, triplicate determinations of each effector-target cell combination are used. The results are expressed as the percentage of cytotoxicity or in lytic units (LU). The LU is the number of effector cells required to kill a given number of target cells and is usually expressed as LU/107 cells. Because the CML assay duplicates the *in vivo* process of tissue destruction through cellular interactions, it has been of great value in studies of allograft rejection (194).

Spontaneous Cell-Mediated Cytotoxicity

Natural killer (NK) cells are defined by their capability to lyse target cells without prior sensitization.

The release of ^{51}Cr from labeled target cells is the standard method for the characterization of NK cells. The chronic myeloid leukemia cell line K562 has been used as the standard target for these assays. Various cell separation techniques have been studied for isolation of a purified fraction of effector cells, and a whole blood assay has also been described (197). However, the ^{51}Cr-release technique has many disadvantages, including the use of radioisotopes, low sensitivity, and the spontaneous release of ^{51}Cr. Several other techniques have been developed for the assessment of NK cell function, but none has achieved widespread utilization. These methods include the incorporation of radioactive nucleotides during cell proliferation (198) dye exclusion (199,200), the measurement of target cell enzyme release (201), and flow cytometry (202,203). Although determinations of single-cell cytotoxicity by dye exclusion has some advantages over ^{51}Cr release for the measurement of NK-cell function, this technique requires laborious, subjective evaluation by microscopy. Therefore, flow cytometric techniques have received considerable attention. In this regard, McGinnes et al. (202) labeled the target cells with carboxyfluorescein diacetate (CFDA) and measured the decrease in fluorescence during lysis, whereas Zarcone et al. (189) quantitated nonviable cells by the use of propidium iodide dye exclusion, and Vitale et al. (204) utilized light scattering parameters. Shi et al. (203) used two-parameter analysis with fluorescein diacetate and propidium iodide and were able to study the interaction between live effector and live target cells, in addition to quantitating lysis of target cells.

Antibody-Dependent Cell-Mediated Cytotoxicity

Antibody-dependent cytotoxicity (ADCC) is mediated largely by a subpopulation of large granular lymphocytes (K cells), although certain other cell types with Fc receptors, including macrophages and granulocytes, also have this ability. ADCC involves fixation of IgG to the surface of a target cell, followed by binding of the Fc receptors of the effector cells to the IgG, and lysis of the target cells (Fig. 61.26). The reaction is usually measured by the use of ^{51}C-labeled target cells (205).

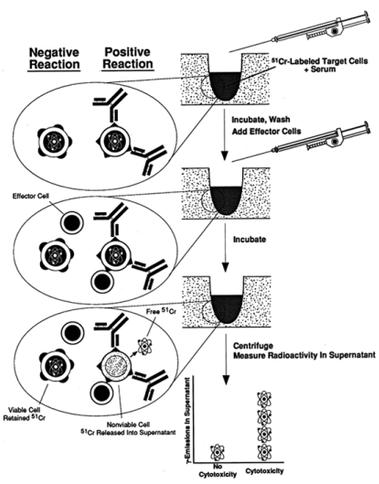


FIGURE 61.26. Principle of the antibody-dependent cell-mediated cytotoxicity assay. This assay is performed with mononuclear cells from a kidney organ donor or from several unrelated individuals (panel cells) that are stimulated with phytohemagglutinin (PHA) and labeled with ^{51}Cr. After incubation with eluate from another nephrectomy specimen or normal human serum, the labeled cells are washed and incubated with effector cells from a normal unrelated individual. Radioactivity (γ emissions) in the cells at the end of the experiment represents unlysed cells. The results are expressed as the percentage of cytotoxicity. (From Riley RS, Mahin EJ, Ross W. *Clinical applications of flow cytometry*. New York: Igaku-Shoin Medical Publishers, 1993, with permission.)

Clinically, this assay has been used to determine the functional activity of cells capable of mediating ADCC and to detect specific antibodies against target cell antigens. It is believed that *in vivo* ADCC is an important mechanism for the destruction of virally infected cells and tumor cells. In addition, antibodies capable of initiating ADCC have been reported to be responsible for some cases of hyperacute allograft rejection.

Assays of Nonspecific Immunity

The activity of neutrophils and macrophages is complex and involves the movement of the cell toward an inflammatory stimulus (chemotaxis), adhesion, opsonization, phagocytosis of foreign particles, and intracellular killing of microorganisms. Some patients with immune function defects have abnormalities of cell movement, phagocytosis, or intracellular killing. The laboratory analysis of these nonspecific immune parameters have been difficult to measure, and many of the assays now in use are subjective and lack standardized methodology and quality control. However, several flow cytometric assays of phagocytic function measure the phagocytosis of microorganisms by cells. In contrast to the subjective assays, flow cytometric studies are rapid, reproducible, and statistically valid. In addition, the multiparametric capability of the flow cytometer allows the simultaneous determination

of several cell functions and an assessment of the interrelationship between cell functions. Flow cytometric assays are currently used for the measurement of phagocytosis of microorganisms, degranulation and enzymatic activity, intracellular killing and degradation of microorganisms, exocytosis, opsonization, oxidative metabolism, membrane potential, phagosomal pH, intracellular calcium, cell volume, and other parameters (206,207,208,209 and 210).

Assays of Cell Movement

Unstimulated neutrophils exhibit random motility. However, in the presence of some substances (chemotaxins), movement becomes nonrandom and unidirectional in the direction of the chemotaxin. Chemotaxis is induced by certain complement activation fragments (C5a and Ba), some coagulation factors (kallekrein and plasminogen activator) substances produced by leukocytes, some byproducts of arachidonic acid metabolism (prostaglandins and leukotrienes), some substances produced by microorganisms, and some synthetic peptides.

The random locomotion of neutrophils can be demonstrated by the capillary tube method. In this procedure, a thin layer of cells from a buffy coat preparation is carefully layered in a capillary tube and observed on an hourly basis. Movement of the cell layer occurs, and the distance from the starting point is recorded after several hours.

The micropore filter technique is the most sensitive means of evaluating chemotaxis and cell migration. Several types of chemotactic chambers are available for this purpose. These consist of two chambers separated by a micropore filter with holes large enough to permit active locomotion but not passive transit (5-mm porosity). A solution containing a chemotactic substance is placed in the lower chamber and a cell suspension in the upper chamber. The chemotactic factors commonly used in the assay include supernatant from a culture of *E. coli*, serum containing activated C5 components, or a synthetic chemotactic peptide (*N*-formylmethionylleucylphenylanaline) (211). After incubation of the chamber, the filter paper is removed and stained. The presence of neutrophils on the underside of the filter is indicative of migration through the paper. However, more accurate assessments of cell migration can be obtained by determination of the average distance migrated per cell (microsectioning method). This leukotactic index is measured by counting the number of cells at 10-mm intervals from the proximal surface of the filter to the distal surface. The number of cells at each interval is multiplied by the distance from the proximal surface and the sum of the products is divided by the total number of cells counted (211). The micropore filter technique has been used to determine the chemotactic ability of different substances, in addition to the evaluation of patients with suspected congenital or acquired disorders of neutrophil migration.

An alternative method for assessing cell migration is the agarose technique, in which the cells under study are placed into a central well punched in a layer of agarose and are allowed to migrate under the agarose toward peripheral wells containing controls or chemoattractant (212,213). Although simple to perform, the agarose method is many times less sensitive than the micropore technique (211).

Ficoll-Hypaque density gradient separation and counterflow centrifugal elutriation have been found to provide the best cell suspensions of studies of neutrophil function (214). In addition, it is essential to maintain cells for studies of random locomotion and chemotaxis at room temperature because even short exposures to cold temperatures can impair these parameters.

Assays of Phagocytosis

Phagocytosis is determined by the use of a particle that is visible within the cell and can be counted after ingestion or an ingestible substance that can be extracted from the cell and measured. Live or heat-killed *Candida albicans* fulfills the first criterion and permits a convenient means of measuring phagocytosis. Neutrophils are incubated with the yeast particles in the fresh serum, Giemsa-stained cell preparations are made, and 200 neutrophils are examined and counted for the presence of ingested yeast (215,216). The results are expressed as the percentage of neutrophils showing phagocytosis and the number of particles per phagocytic neutrophil. Living staphylococci have also been used, with the number of bacteria remaining after phagocytosis determined by culture (217). Another means is to use oil droplets containing dissolved oil red O stain and coated with *E. coli*, liposaccharide, and C3 complement. After incubation, the cells are separated, the oil red O is extracted, and the results are measured spectrophotometrically (218).

Several flow cytometric assays of phagocytosis have been described that are quantitative and more accurate than subjective studies using visual counting of phagocytosed microorganisms. Most investigators (219,220,221,222,223,224,225,226 and 227) use fluorescent microspheres coated with opsonins. However, Wilson et al. (228) and Rothe and Valet (229) reported an increase in red fluorescence in cells stained with acridine orange during phagocytosis. In this, intracellular pH and cell esterase activity during phagocytosis was determined with 1,4-diacetoxy-2,3-dicyanobenzene. A commercial kit for the whole blood flow cytometric assessment of phagocytosis is available (Orpegen, Heidelberg, Germany) that can differentiate between surface-bound and internalized *E. coli* particles (210,230).

Assays of Enzymatic Activity and Intracellular Killing

The intracellular killing of microorganisms is mediated by complex enzymatic mechanisms. In this regard, oxygen-dependent mechanisms of cell killing are the most important, although oxygen-independent have been described. The increase in oxidative metabolism during phagocytosis has been referred to as the oxidative burst. During this process, the key event is the activation of a membrane-bound pyridine nucleotide-dependent oxidase (NADPH oxidase), which reduces oxygen to O_2^- (superoxide anion) and oxidizes NAD(P)H to NAD(P)$^+$. Subsequently, hydrogen peroxide (H_2O_2), hydroxyl radical (OH), and singlet oxygen, are generated. All these compounds are directly toxic to the cell, but the H_2O_2-halide-myeloperoxidase system is the most important mechanism of bacterial killing. In the presence of H_2O_2 and halide ions (I^-, Cl^-), myeloperoxidase catalyzes a direct halogenation of the bacterial cell wall and also produces damage to the cell wall by converting amino acids into aldehydes.

In the clinical laboratory, the Nitroblue tetrazolium dye (NBT) reduction assay has been the standard method for the determination of the oxidative capacity of the neutrophil (209,231,232).

NBT is a yellow dye in its oxidized state. However, when reduced, NBT forms formazan, a crystalline substance that has a deep blue color. NBT is attached to latex particles or microorganisms to permit phagocytosis, or it can be ingested in the soluble form by neutrophils that have been stimulated by endotoxin or other means. If the normal oxidizing environment in the phagolysosome is present, the resultant blue particles can be detected microscopically, or the formazan can be extracted and measured spectrophotometrically (Fig. 61.27). Approximately 10% of normal unstimulated neutrophils ingest and reduce NBT, but this increases to 80% with stimulated neutrophils.

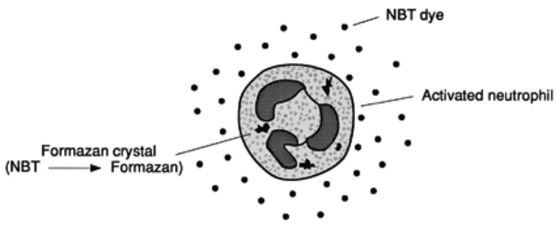


FIGURE 61.27. Principle of the nitroblue tetrazolium (NBT) reduction assay. Neutrophils are incubated with NBT. Enzymatic activity in the activated neutrophil converts NBT into an insoluble brownish-black precipitate (formazan) that is visible under the microscope. The proportion of neutrophils with intracytoplasmic precipitated material is determined.

Bactericidal activity by tritiated thymidine incorporation and autoradiography has been described by Cline (233). In this assay, leukocytes are incubated at 37°C with *Staphylococcus aureus*, ^{3}H-thymidine is added to the reaction mixture, and further incubation is carried out. Autoradiography is performed, and the total number of intracellular bacteria and the number of grains associated with the bacteria are determined in 50 phagocytic neutrophils. The incorporation of ^{3}H-thymidine is observed only in living bacteria.

Several different flow cytometric techniques have been reported for analysis of oxidative metabolism and the oxidative burst. Thorell quantitated NAD(P)H oxidation by measuring the loss of ultraviolet-excited fluorescence associated with NAD(P)H (234,235 and 236). By this technique, the respiratory burst induced by PMA was found to be an all-or-none phenomenon, probably reflecting a narrow threshold between $NAD(P)^{+}$ reduction and complete NAD(P)H oxidation. Bass et al. (237) described a flow cytometric technique for the quantitation of H_2O_2 production. In this technique, the fluorogenic substrate dichlorofluorescein-diacetate (DCFH-DA) is actively taken up by the cell and diacetylated by intracellular esterases into a nonfluorescent compound (2′,7′-dichlorofluorescein, DCFH), which is trapped within the cell. In the presence of H_2O_2 and peroxidase, DCFH is oxidized into dichlorofluorescein (DCF), which is highly fluorescent in the green region of the spectrum (Fig. 61.28).

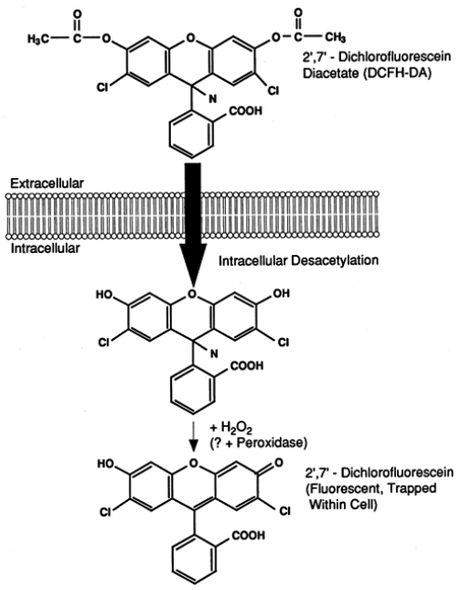


FIGURE 61.28. Flow cytometric analysis of the oxidative burst of polymorphonuclear leukocytes. In this assay, dichlorofluorescein-diacetate (DCFH-DA) is taken up by the cell and is trapped by deacetylation, with the formation of 2′, 7′-dichlorofluorescein. During the oxidative burst, the nonfluorescent 2′, 7′-dichlorofluorescein is converted into the fluorescent product dichlorofluorescein. The fluorescent intensity in the green region of the spectrum is proportional to the oxidative capacity of the cell. (From Bass DA, Parce JW, Dechatelet LR, et al. Flow cytometric studies of oxidative product formation by neutrophils: a graded response to membrane stimulation. *J Immunol* 1983;130:1910–1917, with permission.)

Bass et al. (237) and other investigators utilized this technique to study neutrophilic oxidative metabolism in normal individuals and in patients with infectious diseases and immunodeficiency diseases such as chronic granulomatous disease. Hydroethidine and DCFH were used by Rothe and Valet (238) to simultaneously quantitate both H_2O_2/peroxidase and KO_2. Hydroethidine is oxidized by KO_2 to ethidium bromide. Flow cytometric techniques for the simultaneous measurement of phagocytosis and intracellular killing have been reported (224,239,240).

Recent discoveries have revealed specific functional roles for the cellular receptors for complement components (CD11b, C3bi receptor), the Fc region of IgG (CD16, FcR III; CD32, FcR II; CD64, FcR I) and cell surface adhesion molecules (CD18) on phagocytes. In addition, an intrinsic deficiency of these receptors has been found in some patients with immune deficiency disease (241,242,243,244,245,246 and 247). Flow cytometric evaluation of peripheral blood mononuclear cells for the presence of these receptors and their regulation in the presence of various stimuli is routinely available (209).

Assays of Opsonization

The phagocytosis of microorganisms and other foreign bodies greatly enhanced by the presence of particular chemical substances (opsonins) on the surface of the foreign body. Opsonization is an important function of some antibodies and is demonstrated by some complement breakdown products (C3b, C4b). In patients with some immunoglobulin and complement deficiencies, an increased susceptibility to infection results in part because of decreased opsonization. Opsonins in the serum can be detected by incubating the serum with normal neutrophils and yeast particles and visually quantitating phagocytosis of the yeast particles by normal neutrophils. In addition, flow cytometric methods have been used for the measurement of serum opsonin function (248,249 and 250). In these studies, live FITC-labeled *Neisseria meningitidis* was used to study the opsonic response in patients with serogroup B meningococcal disease. In this assay, the labeled bacteria were preopsonized by incubation with serum for 7.5 minutes, and then incubated with human polymorphonuclear neutrophilic leukocytes (PMNs) for 5 minutes at a bacteria-to-PMNs ratio of 20:1. The results were expressed in the number of bacteria incubated per phagocyte. The opsonic response was markedly increased in acute compared with convalescent serum and remained elevated during the 3-year period examined. ELISA quantitation showed parallel increases in IgG1, IgG3, IgA, and IgM antibodies against type specific group B meningococcal outer membrane proteins (249,251).

Complement Assays

Complement is essential for the defense of the body against disease. In this regard, the complement system can eliminate target

cells or organisms by direct lysis. In addition, individual complement components and fragments produced during complement activation promote immune adherence, opsonization, and chemotaxis, release histamine, and increase vascular perme-ability.

Complement System

The complement system consists of a series of more than 20 proteins that interact in a specific, cascadelike manner (252,253,254,255 and 256) (Table 61.7). The system is regulated by the naturally short life of the activated components and by a number of specific inhibitory compounds. There are two pathways of complement activation (classic and alternative) that are initiated by different mechanisms and may function independently or in concert. Because quantitative and functional abnormalities of the complement system have been associated with autoimmune disease and increased susceptibility to infectious disease, laboratory studies of the complement system can be of diagnostic significance (255,257,258 and 259). Laboratory assays for the assessment of the complement system include immunochemical quantitation of individual complement components, determinations of the functional activity of complement, and tests of complement activation. The detection of complement in tissue can be performed by immunofluorescent or immunoenzymetric staining.

TABLE 61.7. LABORATORY EVALUATION OF THE COMPLEMENT SYSTEM

Functional complement assays
Hemolytic tube technique
Radial immunodiffusion technique
Quantitation of complement components
Immunonephelometry
Radial immunodiffusion
Other techniques
Complement activation assays
Immunoelectrophoresis
Immunofixation electrophoresis
Crossed immunoelectrophoresis
Electroimmunodiffusion
Radioimmunoassay
Enzyme-labeled immunoassay

Functional Complement Assays

The functional integrity of the complement system is determined by RBC lysis. The principle of this assay is that the complement components, acting in concert, can lyse RBCs coated with subagglutinating amounts of an anti-RBC antibody. Because lysis occurs only when all the complement components are functional and present in adequate amounts, a deficiency or

functional defect of a single component can prevent lysis. Therefore, the presence of complement dysfunction can be detected by this technique, but the specific nature of the defect must be determined by other methods. For this reason, these assays are usually referred to as total hemolytic complement assays.

Hemolytic techniques for functional complement can be performed in either a liquid suspension or an agarose gel matrix. The animal source of the RBCs used in the assay determines which pathway is measured. Because the classic pathway is activated by antigen-antibody complexes, SRBCs sensitized with a subhemolyzing dose of rabbit antibodies to SRBCs are utilized in its evaluation. In the alternative pathway hemolytic assay, rabbit RBCs are used because they selectively activate complement via the alternative pathway.

The conventional hemolytic tube technique for functional complement utilizes a standardized suspension of RBCs mixed with a series of dilutions of fresh serum. After an incubation period, the reaction tubes are centrifuged, and the supernatants are analyzed spectrophotometrically to determine the amount of free hemoglobin present and the percentage of RBCs lysed. The measurement of hemolysis is most sensitive in the region where 50% of the RBCs are lysed (Fig. 61.29). Therefore, a unit of total serum hemolytic complement is the dilution of serum required to lyse 50% of a standardized amount of RBCs coated with a standard amount of antibody. The tube titration assay is usually performed in the 50% lysis area, and results are expressed in CH50 (classic pathway) or AP50 units (alternative pathway). Standards can be determined by each manufacturer or purchased by commercial sources.

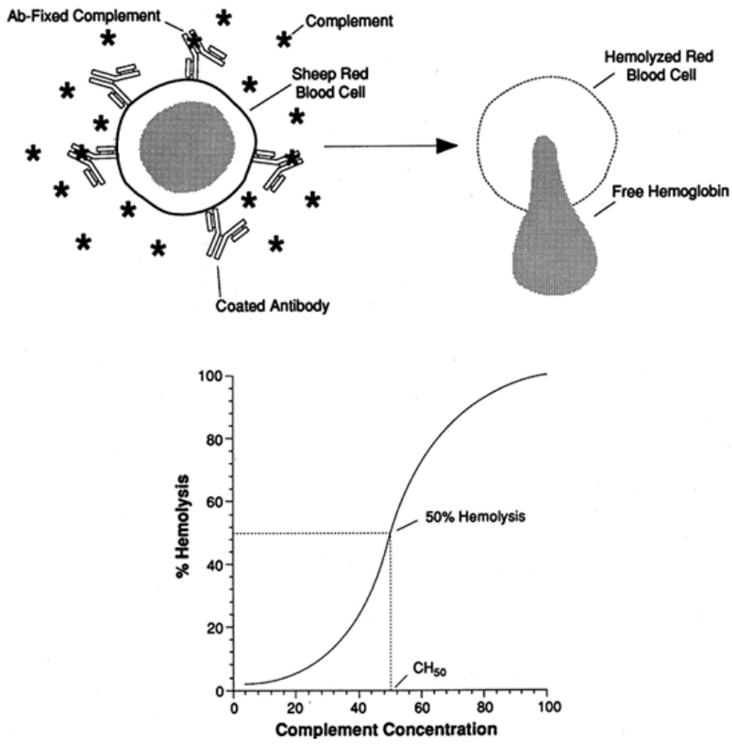


FIGURE 61.29. The hemolytic tube assay for functional complement activity. Sensitized sheep or rabbit red blood cells are the target cells in this assay. Lysis of sheep red blood cells occurs from complement activation on the surface of the red blood cell, resulting in hemolysis and the release of free hemoglobin. The assay can be performed in a liquid medium (hemolytic tube assay) or in agarose gel (gel matrix or radial diffusion method). Sheep red blood cells are used to detect activation of the classic complement pathway, whereas rabbit red blood cells are used for the alternative pathway.

In the gel matrix or radial diffusion method for the determination of functional complement, RBCs are suspended in a buffered agarose mixture that is poured into radial diffusion plates and allowed to solidify. Wells are punched into the agarose. Complement reference sera containing known units of functional complement activity are added to the wells, together with the unknown patient's sera. After an incubation period, a

cleared zone of lysed RBCs appears around each well where hemolysis has occurred. The diameter of the hemolytic zone is proportional to the functional activity of the complement in the well. Therefore, the diameter of each well is measured, and a reference graph is prepared from the known specimens. Unknowns are determined by reading from the graph.

The unit of measurement in the radial diffusion method is the dilution of serum necessary to cause 100% lysis of a standardized preparation of RBCs after activation via the classic or alternative pathways. The results are usually expressed in CH100 or AP100 units (259).

The results obtained by hemolytic tube titration are not directly comparable with those of hemolytic radial diffusion because the analytic techniques are different, and the end points of the assays are measured at different points on the complement response curve. Although the 50% hemolytic units generated by tube titration assays may be considered more sensitive and quantitative than 100% hemolytic units, tube titrations are also more complicated and labor intensive. Radial diffusion requires less elaborate laboratory equipment and is not as sensitive as tube titration to minor changes in reaction conditions. Both assays should be considered as qualitative screening tests for levels of functional complement activity and used in conjunction with quantitative tests. EIA techniques are under evaluation for complement activity (260,261).

Quantitation of Individual Complement Components

The quantitation of the complement components C3 and C4 is routinely performed in the laboratory by nephelometric and other immunoassay techniques. These components are present in the highest concentration in the serum and serve as monitors for the activation of complement. Depressed levels of C4 are usually indicative of classic pathway activation, whereas decreased serum C3 can result from activation of either the classic or alternative pathway. However, normal or elevated C3 and or C4 levels do not rule out complement activation because these substances are acute phase reactants, and their increased synthesis during acute inflammation or infection can more than compensate for consumption. Serum C1q, C2, C5, factor B, and other complement components can be quantitated by radial diffusion in gels or by other techniques if needed in the evaluation of selected patients (262,263).

Tests of Complement Activation

Quantitative assessment of a patient sample for *in vivo* complement activation may be necessary when complement activation is suspected but undetected by immunochemical or functional assays. One reason for this is the wide normal range of complement components, especially C3 and C4. If a patient's C3 or C4 level is normally near the upper range of normal, it may decrease into the lower range of normal during an illness but still be reported as normal.

In vivo activation could also go undetected because most complement components are acute-phase reactants that undergo increased synthesis during states of illness. This compensatory effect may cause normal *in vitro* complement levels or function, even in the presence of complement breakdown.

Testing for activation fractions can also be useful in studies of decreased complement levels because low values may be caused by increased consumption or decreased synthesis. If activation fragments are not found in the presence of low quantitative and functional levels, the cause is decreased synthesis.

The methods utilized for the laboratory detection of complement activation fragments include (a) IEP, (b) IFE, (c) crossed IEP, (d) electroimmunodiffusion, (e) RIA, and (f) ELISA (264). Using these techniques, it is possible to detect the smaller breakdown products of C3a, C3c, C3d, C4d, Ba, and Bb.

Immunohistochemical Staining

A new era of pathologic diagnosis evolved with the introduction of immunochemical techniques into the histology laboratory. These techniques utilize the natural specificity of the antibody to identify and localize a variety of antigens, including immunoglobulins, cell surface markers, hormones, enzymes, oncodevelopmental antigens, viral/bacterial substances, many of which were previously undetectable by cytochemical staining. Polyclonal or monoclonal antibodies labeled with a fluorescent dye were first used as immunohistochemical stains. Unfortunately, immunofluorescent staining is of limited sensitivity and requires frozen tissue sections. In addition, a fluorescent microscope is required for reading the slides. Immunoperoxidase techniques, utilizing monoclonal or polyclonal antibodies, are many times more sensitive than immunofluorescent methods and can usually be performed on fixed, paraffin-embedded tissue. Immunoperoxidase-stained sections can be examined with the ordinary light microscope and are able to withstand long storage.

Coons et al. first used antibodies conjugated with a fluorescent dye for the localization of antigens in tissue in 1941. FITC has been the most widely used fluorescent label because the human eye is very sensitive to its yellow-green fluorescence and because green autofluorescence is less common in nature than red autofluorescence (126). The direct and indirect methods of tissue immunofluorescence have been used, although inhibition and complement-staining techniques have been described.

Direct immunofluorescent staining is used to detect the presence of antigen in a tissue. In this procedure, the fluorochrome-labeled antibody is directly applied to the tissue section at its optimal dilution and incubated for a short period of time. After washing to remove unbound dye, the section is mounted in a water-soluble medium (usually buffered glycerol) and visualized by light microscopy. The sensitivity of direct immunofluorescent staining depends on the avidity of the antibody, the density of the antigen in the tissue, and the sensitivity of the eye.

Indirect immunofluorescent staining is primarily used to detect the presence of an antibody in body fluids, although the same procedure can also be applied for the identification of an unknown antigen in a tissue. In indirect staining, unlabeled antibody is first reacted with the antigen in the tissue. After incubation and washing to remove unlabeled antibody, a fluorochrome-labeled secondary antibody specific for the primary antibody is used. The tissue is visualized by fluorescent microscopy after a second incubation and washing. The indirect method is more sensitive than the direct technique because the second antibody amplifies the staining intensity. In addition, this technique is economical because the same-labeled antibody

can be used with multiple primary antibodies. However, nonspecific staining can be increased relative to direct staining.

There are numerous quality control considerations in immunofluorescent staining. First, the specificity of an antibody used in immunofluorescent staining must be extensively documented and the antibody must be used at its proper dilution. For the identification of antigens in tissue, controls with and without the antigen must be available. If possible, demonstration of the specificity of binding of the antibody should be documented by inhibition or blocking of fluorescent staining on known positive tissue by unlabeled antigen. In addition, loss of staining should be observed in the positive control after the antiserum is adsorbed with purified antigen. The dye:protein ratio is another consideration. The negative charge of a conjugate with a high dye:protein ratio can lead to increased nonspecific staining. Conversely, the sensitivity of detection of the staining reaction is reduced by the use of a conjugate with a dye:protein ratio that is excessively low. Electrophoresis with densitometric scanning of the gel can be used to detect the presence of unlabeled fluorescein and/or protein. The total protein concentration of a conjugate is measured by the biuret reaction of similar procedure. Unlabeled dye can be removed by dialysis, whereas highly negatively charged conjugates can be separated by DEAE-cellulose chromatography. The presence of nonspecific, extraneous antibodies in antisera is best detected by IEP or precipitin reactions in gels (126). Documentation of the sensitivity of a conjugate is performed by gel diffusion precipitation, using serially diluted preparations of the conjugate against a solution of the antigen at a concentration of 1 mg/mL. The highest dilution of the conjugate that gives a visible line of precipitation is the unitage. Conjugates for immunofluorescent staining should have at least 4 units per 1% protein (126). Although F:P ratios can be expressed as a weight ratio (micrograms of bound FITC per milligram of protein), they are usually expressed as molar ratios by multiplying the weight ratios by a factor of 0.411.

Commercial antisera and conjugates are supplied with information regarding the F:P ratio, protein concentration, and the results of assays for immunologic sensitivity and specificity. However, because the optimal working dilution will vary in each laboratory, it must be individually determined before using the reagent. For direct immunofluorescent staining, serial dilutions of the conjugate are used with a positive control to determine the optimal combination of bright fluorescence and low background staining. In indirect staining, where two antibodies are used, chessboard (or checkerboard) titrations are performed. In this technique, serial dilutions of a primary antibody are incubated with a well-characterized positive control, and each dilution is in turn developed with serial dilutions of the fluorochrome-labeled secondary antibody. The fluorescence intensity of each combination is recorded in tabular format. For each dilution of primary antibody, the fluorescent intensity will be constant for several dilutions of conjugate (plateau titer). In general, the highest conjugate dilution (plateau end point) can be diluted twofold to obtain a working dilution. For fluoresceinated antibodies with an F:P ratio of 3:5, a plateau end point between 1/6 and 1/64 is expected against a high-titered primary antibody, and a working dilution of approximately 1/4 unit/mL is generally used. For conjugates employed in direct immunofluorescent staining, dilutions of 1 unit per milliliter to 1/4 unit per milliliter provide satisfactory results in most circumstances (126). Because antibody conjugates are labile, storage conditions are critical. Generally, fresh conjugates should be ultracentrifuged, divided into small aliquots, and stored in a protein-rich medium (2 mg/mL or greater) at ultra-cold temperatures (-80°C or the vapor phase of liquid nitrogen) for long-term storage. Conjugates should be thawed only once before use and should be given a high-speed centrifugation before use (126). Stabilized, undiluted conjugates can usually be stored at 4°C for several weeks.

Proper preservation and sectioning of tissue is essential for the retention of immunologic reactivity and morphologic detail. Needle biopsies or small pieces of tissue (approximately 3 to 5 mm) are obtained in the fresh state and placed on pieces of moistened (not soaked) gauze for transport. The tissue is then placed on a piece of a wooden tongue blade, sponge, or cork, embedded in OCT compound (Ames Co., Elkhart, IN, U.S.A.), and snap or flash frozen by complete immersion in liquid nitrogen, isopentane cooled to the temperature of liquid nitrogen, or a dry ice-acetone mixture. The frozen tissue is then wrapped in aluminum foil, placed in a precooled, airtight storage container, and stored at a temperature of -70° or -80°C until sectioning and staining. Sections of 4 to 6 μ are cut under reduced temperature with a standard cryostat. Brief fixation of the tissue sections in acetone or absolute ethyl alcohol is sometimes employed to prevent the loss of soluble antigens or antibodies.

The fluorescent microscope provides a means of illuminating stained tissue with light of a proper wavelength under dark-field conditions, separating the excitation and emission wavelengths and visualizing the stained tissue. Fluorescent microscopes can be classified into two types based on optical configuration. The bright-field incident (epi-illumination) system described by Ploem in 1967 has been the most widely utilized. The heart of the epi-illumination fluorescent microscope is a dichroic (beam-splitting) mirror. In this system, excitation light from a high-intensity light source is passed through a primary filter, strikes a dichroic mirror placed at a 45 degree angle, and is deflected through the objective of the microscope unto the slide where a dark field is created. The light emitted from the specimen returns through the objective and passes through the dichroic filter and secondary filter to the oculars. The primary (excitation) filter is used to select light of the proper wavelength for excitation, whereas the secondary (barrier) filter transmits light at the emission wavelength of the fluorochrome under study, while removing excitation wavelengths. The dichroic filter is essential to reflect excitation light to the slide and transmit emitted light to the observer. One advantage of this system is that a condenser, with its associated light loss, is not required because the objective acts as a condenser. In addition, filter systems can be easily exchanged for viewing tissue sections labeled with multiple conjugated dyes, and fluorescence can be performed in conjunction with phase contrast and other types of microscopy. Fluorescent microscopes with transmitted illumination usually employ dark-field (oil-immersion) condensers. Excitation light passes through a primary filter and the dark-field condenser onto the glass slide. Emitted light passes through the microscope objective and secondary filter to the ocular lens (Fig. 61.30). This design is less efficient than epi-illumination systems because some light is diffused

in passing through the condenser and barrier filter. In addition, filter configurations are more difficult to change. In all microscopes, the objectives must be nonfluorescent and collect and transmit light with high efficiency (i.e., have high numerical aperture).

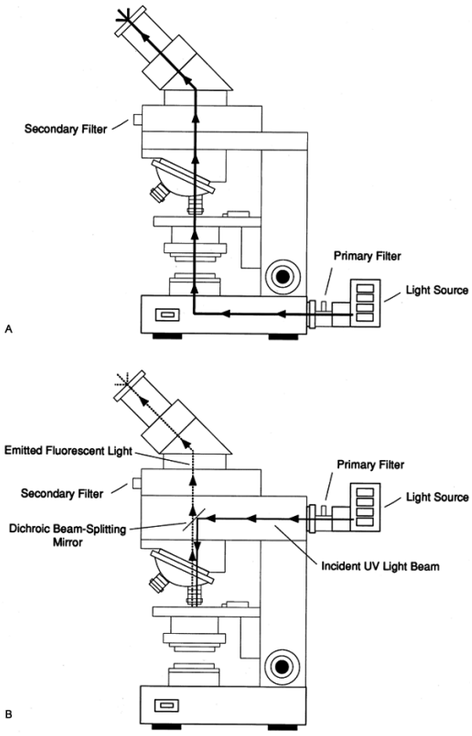


FIGURE 61.30. Schematic illustrations of epi-illumination **(a)** and transmitted illumination **(B)** fluorescent microscopes.

High-pressure arc (mercury, xenon) and halogen (quartz-iodine or tungsten-halogen) lamps have traditionally been the excitation sources for fluorescent microscopes. High-pressure arc lamps are the most widely used excitation sources because they produce light of very high intensity and have relatively long life spans (approximately 200 hours). However, light intensity decreases with age, and these lamps require a protective housing because they operate at high temperature and pressure. In the near future, the small, inexpensive lasers that have recently become

available in flow cytometers may be incorporated into the fluorescent microscope. These have very long life spans and produce monochromatic light of very high intensity.

If an assay for the detection of identification of antibodies is conducted, appropriate positive and negative sera must be used for controls. For antibody identification, serial dilutions of the specimen are tested, and the dilution (titer) that produces a minimal degree of fluorescence is reported as positive.

REFERENCES

1. Ronald A, Stimson WH. The evolution of immunoassay technology. *Parasitology* 1998;117(Suppl):S13–27.

2. Larkin DF. Antigen processing and presentation. *Dev Ophthalmol* 1997;28:127–134.

3. Jensen PE. Mechanisms of antigen presentation. *Clin Chem Lab Med* 1999;37:179–186.

4. Krensky AM. The HLA system, antigen processing and presentation. *Kidney Int* 1997;58(Suppl):S2–S7.

5. Whitton JL. An overview of antigen presentation and its central role in the immune response. *Curr Top Microbiol Immunol* 1998;232:1–13.

6. Van der Meide PH, Schellekens H. Cytokines and the immune response. *Biotherapy* 1996;8:243–249.

7. Callard R, George AJ, Stark J. Cytokines, chaos, and complexity. *Immunity* 1999;11:507–513.

8. Mire-Sluis AR. Cytokines: from technology to therapeutics. *Trends Biotechnol* 1999;17:319–325.

9. Inglis JR. Special feature on immunotechnology and industry: introduction. *Immunol Today* 1983;4:125–139.

10. Wild D, Davies C. Components. In: Wild D, ed. *The immunoassay handbook*. New York: Stockton Press, 1994:49–82.

11. Ritchie RF. Preparation of polyclonal antisera. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:4–8.

12. Hudson GA. Characterization of antisera. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:49–56.

13. Viadya HC, Beatty BG. Eliminating interference from heterophilic antibodies in a two-site immunoassay for creatine kinase MB by using F(ab′)2 conjugate and polyclonal mouse IgG. *Clin Chem* 1992;38:1727–1742.

14. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 1975;256:445–497.

15. Epstein N, Epstein M. The hybridoma technology: I. Production of monoclonal antibodies. *Adv Biotechnol Processes* 1986;6:179—218.

16. Bruggemann M, Taussig MJ. Production of human antibody repertoires in transgenic mice. *Curr Opin Biotechnol* 1997;8:455–458.

17. Green LL. Antibody engineering via genetic engineering of the mouse: XenoMouse strains are a vehicle for the facile generation of therapeutic human monoclonal antibodies. *J Immunol Methods* 1999;231:11–23.

18. Moore GP. Genetically engineered antibodies. *Clin Chem* 1989;35:1849–1853.

19. Peterson NC. Recombinant antibodies: alternative strategies for developing and manipulating murine-derived monoclonal antibodies. *Lab Anim Sci* 1996;46:8–14.

20. Wood RF, Pockley AG. Bioengineered monoclonals—easier on patients, but still hard on rejection (comment). *Transplantation* 1999;68:1625–1626.

21. Colcher D, Pavlinkova G, Beresford G, et al. Pharmacokinetics and biodistribution of genetically-engineered antibodies. *Q J Nucl Med* 1998;42:225–241.

22. Vaughan TJ, Osbourn JK, Tempest PR. Human antibodies by design. *Nat Biotechnol* 1998;16:535–539.

23. Kricka LJ. Human anti-animal antibody interferences in immunological assays. *Clin Chem* 1999;45:942–956.

24. Mitchell R, Hollis S, Crowley V, et al. Immunometric assays of luteinizing hormone (LH): differences in recognition of plasma LH by anti-intact and ß-subunit-specific antibodies in various physiological and pathological situations. *Clin Chem* 1995;41:1139–1145.

25. Breedveld FC. Therapeutic monoclonal antibodies. *Lancet* 2000;355:735–740.

26. Chang SS, Bander NH, Heston WD. Monoclonal antibodies: will they become an integral part of the evaluation and treatment of prostate cancer—focus on prostate-specific membrane antigen? *Curr Opin Urol* 1999;9:391–395.

27. Laurino JP, Shi Q, Ge J. Monoclonal antibodies, antigens and molecular diagnostics: a practical overview. *Ann Clin Lab Sci* 1999;29:158–166.

28. Weiner LM. An overview of monoclonal antibody therapy of cancer. *Semin Oncol* 1999; 26(Suppl 12):41–50.

29. Appelbaum FR. Antibody-targeted therapy for myeloid leukemia. *Semin Hematol* 1999;36(Suppl 6):2–8.

30. Dyer MJ. The role of CAMPATH-1 antibodies in the treatment of lymphoid malignancies. *Semin Oncol* 1999;26(Suppl 14):52–57.

31. Gopal AK, Press OW. Clinical applications of anti-CD20 antibodies. *J Lab Clin Med* 1999;134:445–450.

32. Multani PS, Grossbard ML. Monoclonal antibody-based therapies for hematologic malignancies. *J Clin Oncol* 1998;16:3691–3710.

33. Buske C, Feuring-Buske M, Unterhalt M, et al. Monoclonal antibody therapy for B cell non-Hodgkin's lymphomas: emerging concepts of a tumour-targeted strategy. *Eur J Cancer* 1999;35:549–557.

34. Cabanas C, Sanchez-Madrid F. Monoclonal antibodies specific for leukocyte adhesion molecules. Selective protocols of immunization and screening assays for generation of blocking, activating and activation reporter antibodies. *Methods Mol Biol* 1999;96:1–9.

35. Kosits C, Callaghan M. Rituximab: a new monoclonal antibody therapy for non-Hodgkin's lymphoma. *Oncol Nurs Forum* 2000;27:51–59.

36. Dunn CJ, Foster RH. Abciximab. A pharmacoeconomic review of its use in percutaneous coronary revascularisation. *Pharmacoeconomics* 1999;16:711–741.

37. Harlan JM. Introduction: anti-adhesion therapy in sickle cell disease (comment). *Blood* 2000;95:365–367.

38. Topol EJ. Evolution of improved antithrombotic and antiplatelet agents. *Am J Cardiol* 1998;82:63P–68P.

39. Wall GC, Heyneman C, Pfanner TP. Medical options for treating Crohn's disease in adults: focus on antitumor necrosis factor-alpha chimeric monoclonal antibody. *Pharmacotherapy* 1999;19:1138–1152.

40. Brenner TL, Adams VR. First MAb approved for treatment of metastatic breast cancer. *J Am Pharm Assoc* 1999;39:236–238.

41. Ditzel HJ. Human monoclonal antibodies: a tool for cancer detection *in vivo*. *APMIS Suppl* 1999;94:5–42.

42. Fleckenstein G, Osmers R, Puchta J. Monoclonal antibodies in solid tumours: approaches to therapy with emphasis on gynaecological cancer. *Med Oncol* 1998;15:212–221.

43. Welt S, Ritter G. Antibodies in the therapy of colon cancer. *Semin Oncol* 1999;26:683–690.

44. Davies C. Principles. In: Wild D, ed. *The immunoassay handbook*. New York: Stockton Press, 1994:3–47.

45. Hogg N. The structure and function of Fc receptors. *Immunol Today* 1988;9:185–186.

46. Kasahara Y. Agglutination immunoassays. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:7–12.

47. Coombs RR, Scott ML, Cranage MP. Assays using red cell-labelled antibodies. *J Immunol Methods* 1987;101:1–14.

48. van Hell H, Leuvering JHW, Gribnau TCJ. Particle immunoassays. In: Collins WP, ed. *Alternative immunoassays*. New York: John Wiley & Sons, 1985:39–58.

49. Boyden SV. The adsorption of protein on erythrocytes treated with tannic acid and subsequent hemagglutination by antiprotein sera. *J Exp Med* 1951;93:107–120.

50. Singer JM, Plotz CM. The latex fixation test. I. Application to the serological diagnosis of rheumatoid arthritis. *Am J Med* 1956;21:888–891.

51. Collet-Cassart D, Limet JN, Van Krieken L, et al. Turbidimetric latex immunoassay of placental lactogen on microtiter plates. *Clin Chem* 1989;35:141–143.

52. Sladoljev S, Trescec A, Cvetkovic M, et al. The development of a latex agglutination inhibition assay for the determination of foetal calf serum in viral vaccines. *Biologicals* 1990;18:45–48.

53. Mongkolsirichaikul D, Tarnchompoo B, Ratanabanangkoon K. Development of a latex agglutination inhibition reaction test for amphetamines in urine. *J Immunol Methods* 1993;157:189–195.

54. Kaszanyitzky E, Czifra G, Stipkovits L. Detection of *Mycoplasma gallisepticum* antibodies in turkey blood samples by ELISA and by the slide agglutination and haemagglutination inhibition tests. *Acta Vet Hung* 1994;42:69–78.

55. Ryo R, Sugano W, Goto M, et al. Platelet release reaction during EDTA-induced platelet agglutinations and inhibition of EDTA-induced platelet agglutination by anti-glycoprotein II b/III a complex monoclonal antibody. *Thromb Res* 1994;74:265–272.

56. Aoki K, Shikama Y, Yoshida T, et al. Enzyme-linked immunosorbent assay and latex agglutination inhibition reaction test for cocaine and benzoylecgonine in urine. *Forensic Sci Int* 1996;77:151–157.

57. Aoki K, Shikama YS, Kokado A, et al. Enzyme-linked immunosorbent assay and latex agglutination inhibition reaction test for morphine in urine. *Forensic Sci Int* 1996;81:125–132.

58. Kubota N, Momose K, Maekawa H, et al. Turbidimetric latex agglutination inhibition immunoassay for primidone. *Biol Pharm Bull* 1994;17:463–466.

59. Tsunekawa Y, Hasegawa T, Nadai M, et al. Clinical application of the Biotrack 516 system for determination of theophylline by a turbidimetric latex agglutination inhibition reaction. *Ther Drug Monit* 1996;18:86–91.

60. Martins TB, Jaskowski TD, Mouritsen CL, et al. Comparison of commercially available enzyme immunoassay with traditional serological tests for detection of antibodies to Coccidioides immitis. *J Clin Microbiol* 1995;33:940–943.

61. Johnson AM. Immunoprecipitation in gels. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*. Washington, DC: American Society for Microbiology, 1986:14–24.

62. Bridges AJ, Lorden TE, Havighurst TC. Autoantibody testing for connective tissue diseases. Comparison of immunodiffusion, immunoblot, and enzyme immunoassay. *Am J Clin Pathol* 1997;108:406–410.

63. Mancini G, Carbonara AO, Heremans JF. Immunochemical quantitation of antigens by single radial immunodiffusion. *Immunochemistry* 1965;2:235–254.

64. Sharma M, Datta U, Roy P, et al. Low sensitivity of counter-current immuno-electrophoresis for serodiagnosis of typhoid fever. *J Med Microbiol* 1997;46:1039–1042.

65. Svendsen PJ. Fused rocket immunoelectrophoresis. In: Axelsen NH, Kroll J, Weeke B, eds. *A manual of quantitative immunoelectrophoresis*. Oslo: Universitetsforlaget, 1973:69–70.

66. Kyle RA, Katzmann JA. Immunochemical characterization of immunoglobulins. In. Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:156–176.

67. Alper CA, Johnson AM. Immunofixation electrophoresis. a technique for the study of protein polymorphism. *Vox Sang* 1969;17:445–452.

68. Arnaud P, Wilson GB, Koistinen J, et al. Immunofixation after electrofocusing. Improved method for specific detection of serum proteins with determination of isoelectric points. I. Immunofixation print technique for detection of alpha-1-protease inhibitor. *J Immunol Methods* 1977;16:221–231.

69. Richie RF, Smith R. Immunofixation. I. General principles and application to agarose gel electrophoresis. *Clin Chem* 1976;22:1735–1737.

70. Gerson B, LaBrie J, Copeland BE. Immunofixation on thin-layer agarose. *Clin Chem* 1979;25:197.

71. Ledue TB, Garfin DE. Immunofixation and immunoblotting. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:54–64.

72. Keren DF. Clinical indications for electrophoresis and immunofixation. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:65–74.

73. van Lente F. Light-scattering immunoassays. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:13–19.

74. Thater JT, Price CP, Spencer K. Immunonephelometric and immunoturbidimetric assays for proteins. *CRC Crit Rev Clin Lab Sci* 19XX;18:213–257.

75. Sternberg JC. Rate nephelometry. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: ASM Press, 1986:304–307.

76. Hellsing K, Laurent TC. The influence of dextran on the precipitin reaction. *Acta Chem Scand* 1964;18:1303.

77. Foster RC, Ledue TB. Turbidimetry. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:25–32.

78. Normansell DE. Quantitation of serum immunoglobulins. *CRC Crit Rev Clin Lab Sci* 1982;17:103–170.

79. Opheim KE, Glick MR, Ou C-N, et al. Particle-enhanced turbidimetric inhibition immunoassay for theophylline evaluated with the Du Pont ACA. *Clin Chem* 1984;30:870–874.

80. Ekins RP. Ligand assays. from electrophoresis to miniaturized microarrays. *Clin Chem* 1998;44:2015–2030.

81. Gosling JP. A decade of development in immunoassay methodology. *Clin Chem* 1990;36:1408–1427.

82. Ekins RP. The estimation of thyroxine in human plasma by an electrophoretic technique. *Clin Chem Acta* 1960;5:453–459.

83. Diamandis EP. Immunoassays with time-resolved fluorescence spectroscopy. Principles and applications. *Clin Biochem* 1988;21:139–150.

84. Price CW. The evolution of immunoassay as seen through the journal *Clinical Chemistry*. *Clin Chem* 1998;44:2071–2074.

85. Schall RF, Tenoso HJ. Alternatives to radioimmunoassay. Labels and methods. *Clin Chem* 1981;27:1157–1164.

86. Howanitz JH. Immunoassay. Innovations in label technology. *Arch Pathol Lab Med* 1988;112:775–779.

87. Yalow RS, Berson SA. Immunological specificity of human insulin. Application to immunoassay of insulin. *J Clin Invest* 1961;40:2190–2198.

88. Larsen J, Odell WD. General principles of radioimmunoassay. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory Immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:110–115.

89. Binoux MA, Odell WD. Use of dextran-coated charcoal to separate antibody-bound from free hormone. a critique. *J Clin Endocrinol Metab* 1973;36:303–310.

90. Desbuquois B, Aurbach GD. Use of polyethylene glycol to separate free and antibody-bound peptide hormones in radioimmunoassays. *J Clin Endocrinol* 1971;33:732–738.

91. Carpenter AB. Enzyme-linked Immunoassays. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*, 5th ed. Washington, DC: ASM Press, 1997:20–29.

92. Johannsson A, Ellis DH, Bates DL, et al. Enzyme amplification for immunoassays. Detection limit of one hundredth of an attomole. *J Immunol Methods* 1986;87:7–11.

93. Ishikawa E, Hashida S, Tanaka K, Kohno T. Development and applications of ultrasensitive enzyme immunoassays for antigens and antibodies. *Clin Chim Acta* 1989;185:223–230.

94. Oellerich M. *Enzyme-immunoassay*. A review. *J Clin Chem Clin Biochem* 1984;22:895–904.

95. Kabakoff DS. Chemical aspects of enzyme-immunoassay. In: Maggio ET, ed. *Enzyme-immunoassay*. Boca Raton, FL: CRC Press, 1980:71–104.

96. Rowley GL, Rubenstein JE, Huisjen J, et al. Mechanism by that antibodies inhibit hapten-malate dehydrogenase conjugates. *J Biol Chem* 1975;250:3759–3766.

97. Ullman EF, Blakemore J, Leute RK, et al. Homogeneous enzyme immunoassay for thyroxine. *Clin Chem* 1975;21:1011.

98. Boguslaski RC, Li TM. Homogeneous immunoassays. *Appl Biochem Biotechnol* 1982;7:404.

99. Carrio RJ, Christner JE, Boguslaski RC, et al. A method for monitoring specific binding reactions with cofactor labeled ligands. *Anal Biochem* 1976;72:271–282.

100. Schroeder HR, Carrico RJ, Boguslaski RC, et al. Specific binding reactions monitored with ligand-cofactor conjugates and bacterial luciferase. *Anal Biochem* 1976;72:283–292.

101. Ngo TT, Lenhoff HM. Enzyme modulators as tools for the development of the homogeneous enzyme immunoassays. *FEBS Lett* 1980;116:285–288.

102. Gibbons I, Hanlon TM, Skold CN, et al. Enzyme-enhancement immunoassay. A homogeneous assay for polyvalent ligands and antibodies. *Clin Chem* 1981;27:1602–1608.

103. Wilkins TA, Brouwers G, Mareschal J-C, et al. High sensitivity, homogeneous particle-based immunoassay for thyrotropin (Multipact™). *Clin Chem* 1988;34:1749–1752.

104. Auditore K, Houghton RL, Monji N, et al. Phase-separation immunoassays. *Clin Chem* 1987;33:1509–1516.

105. Litchfield WJ, Freytag JW, Adamich M. Highly sensitive immunoassays based on use of liposomes without complement. *Clin Chem* 1984;30:1441–1445.

106. Henderson DR, Friedman SB, Harris JD, et al. CEDIA™, a new homogeneous immunoassay system. *Clin Chem* 1986;32:1637–1641.

107. Khanna PL, Dworschack RT, Manning WB, et al. A new homogeneous enzyme immunoassay using recombinant enzyme fragments. *Clin Chim Acta* 1989;185:231–240.

108. Blum LJ, Gautier SM, Coulet PR. Design of luminescence photobiosensors. *J Bioluminescence Chemiluminescence* 1989;4:543–550.

109. Davis KA, Leary TR. Continuous liquid-phase piezoelectric biosensor for kinetic immunoassays. *Anal Chem* 1989;61:1227–1230.

110. Green MJ. Electrochemical immunoassays. *Philos Trans R Soc Lond B Biol Sci* 1987;316:135–142.

111. Heineman WR, Halsall HB, Wehmeyer KR, et al. Immunoassay with electrochemical detection. *Methods Biochem Anal* 1987;32:345–393.

112. Ngeh NJ, Suleiman AA, Guilbault GG. Piezoelectric crystal biosensors. *Biosens Bioelectron* 1990;5:13–26.

113. Rishpon J, Rosen I. The development of an immunosensor for the electrochemical determination of the isoenzyme LDH5. *Biosensors* 1989;4:61–74.

114. Stanley CJ, Cox RB, Cardosi MF, et al. Amperometric enzyme-amplified immunoassays. *J Immunol Methods* 1988;112:153–161.

115. Tsuji I, Eguchi H, Yasukouchi K, et al. Enzyme immunosensors based on electropolymerized polytyramine modified electrodes. *Biosens Bioelectron* 1990;5:87–101.

116. Mayo CS, Hallock RB. Immunoassay based on surface plasmon oscillations. *J Immunol Methods* 1989;120:105–114.

117. Giegel JL, Brotherton MM, Cronin P, et al. Radial partition immunoassay. *Clin Chem* 1982;28:1894–1898.

118. Anderson RR, Lee TT, Saewart DC, et al. Internally reference ImmunoCentration™ assays. *Clin Chem* 1986;32:1692–1695.

119. Waggoner AS. Fluorescent probes for analysis of cell structure, function, and health by flow and imaging cytometry. In: *Applications of fluorescence in the biomedical sciences*. New York: Alan R. Liss, 1986:3–28.

120. Mathies RA, Stryer L. Single-molecule fluorescence detection. A feasibility study using phycoerythrin. In: *Applications of fluorescence in the biomedical sciences*. New York: Alan R. Liss, 1986:3–28.

121. Oi VT, Glazer AN, Stryer L. Fluorescent phycobiliprotein conjugates for analyses of cells and molecules. *J Cell Biol* 1982;93:981–986.

122. Glazer AN, Stryer L. Phycofluor probes. *Trends Biol Sci* 1984;9:423–427.

123. Kronick MN. The use of phycobiliproteins as fluorescent labels in immunoassay. *J Immunol Methods* 1986;92:1–13.

124. Smith DS, Al-Hakiem MHH, Landon J. A review of fluoroimmunoassay and immunofluorometric assay. *Ann Clin Biochem* 1981;18:253–274.

125. Hemmila I. Fluoroimmunoassays and immunofluorometric assays. *Clin Chem* 1985;31:359–370.

126. Nakamura RM, Tucker ES III. Antibody as reagent. In: Henry JB, ed. *Clinical diagnosis and management by laboratory methods*, 17th ed. Philadelphia: Saunders, 1984:893–923.

127. Smith DS. Enhancement fluoroimmunoassay of thyroxine. *FEBS Lett* 1977;77:25–27.

128. Calvin J, Burling K, Blow C, et al. Evaluation of fluorescence excitation transfer immunoassay for measurement of specific proteins. *J Immunol Methods* 1985;86:249–256.

129. Peterson JD, Kim JY, Melvold RW, et al. A rapid method for quantitation of antiviral antibodies. *J Immunol Methods* 1989;119:83–94.

130. Dandliker WB, De Saussure VA. Fluorescence polarization in immunochemistry. *Immunochemistry* 1970;7:799–828.

131. Dandliker WB, Kelley RJ, Dandliker J, et al. Fluorescence polarization immunoassay. Theory and experimental method. *Immunochemistry* 1973;10:219–227.

132. Jovin TM. Fluorescence polarization and energy transfer. Theory and application. In: Melamed MR, Mullaney PF, Mendelsohn ML, eds. *Flow cytometry and sorting*. New York: John Wiley & Sons, 1979.

133. Szollosi J, Trons L, Damjanoviek S, et al. Fluorescence energy transfer measurements on cell surfaces. A critical comparison of steady state fluorimetric and flow cytometric methods. *Cytometry* 1984;5:210–216.

134. Watson RAA, Landon J, Shaw EJ, et al. Polarization immunoassay of gentamicin. *Clin Chim Acta* 1976;73:51–55.

135. McGregor AR, Crookall-Greening JO, Landon J, et al. Polarization fluoroimmunoassay of phenytoin. *Clin Chim Acta* 1978;83:161–166.

136. Colbert DL, Gallacher G, Mainwaring-Burton RW. Single-reagent polarization fluoroimmunoassay for amphetamine in urine. *Clin Chem* 1985;31:1193–1195.

137. Halfman CJ, Jay DW. Homogeneous, micelle quenching fluoroimmunoassay for detecting amphetamines in urine. *Clin Chem* 1986;32:1677–1681.

138. Soini E, Kojola H. Time-resolved fluorometer for lanthanide chelates—a new generation of nonisotopic immunoassays. *Clin Chem* 1983;29:65–68.

139. Saunders GC, Jett JH, Martin JC. Amplified flow-cytometric separation-free fluorescence immunoassays. *Clin Chem* 1985;31:2020–2023.

140. Lindmo T, Bormer O, Ugelstad J, et al. Immunometric assay by flow cytometry using mixtures of two particle types of different affinity. *J Immunol Methods* 1990;126:183–189.

141. Wehry EL. Molecular fluorescence, phosphorescence, and chemiluminescence spectrometry. *Anal Chem* 1986;58.

142. Campbell AK, Simpson JSA. Chemi- and bioluminescence as an analytical tool in biology. *Tech Life Sci (Metab Res)* 1979;B213:1–56.

143. Patel A, Morton MS, Woodhead JS, et al. A new chemiluminescent label for use in immunoassay. *Biochem Soc Trans* 1982;10:224–225.

144. Barnard GJ, Collins WP. The development of luminescence immunoassays. *Med Lab Sci* 1987;44:249–266.

145. Patel A, Campbell AK. Homogeneous immunoassay based on chemiluminescence energy transfer. *Clin Chem* 1983;29:1604–1608.

146. Strasburger CJ, Kohen F. Two-site and competitive chemiluminescent immunoassays. *Methods Enzymol* 1990;184:481–496.

147. Weeks I, Sturgess ML, Woodhead JS. Chemiluminescence immunoassay. An overview. *Clin Sci* 1986;70:403–408.

148. Selby C. Interferences in immunoassay. *Ann Clin Biochem* 1999;36:704–721.

149. Levinson S. Test interferences from endogenous antibodies. *J Clin Lig Assay Soc* 1997;20:180–189.

150. Charrie A, Charriere G, Guerrier A. Hook effect in immunometric assays for prostate-specific antigen [Letter]. *Clin Chem* 1995;41:480–481.

151. Morrissey NE, Quadri SF, Kinders R, et al. Modified method for determining carcinoembryonic antigen in the presence of munan anti-murine antibodies. *Clin Chem* 1993;39:522–529.

152. Bray RA, Gebel HM. Basic concepts of cell isolation. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:17–22.

153. Nisperos B. Density gradient isolation of peripheral blood lymphocytes. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:23–27.

154. Wottawa A, Klein G, Altman H. A method for the isolation of human and animal lymphocytes with Ficoll-Urografin. *Wien Klin Wochenschr* 1974;86:161–163.

155. Perer RJ, Tina WZ, Mickelson MM. Purification of lymphocytes and platelets by gradient centrifugation. *J Clin Lab Med* 1968;72:842–848.

156. Boyum A. Isolation of mononuclear cells and granulocytes from human blood. *Scand J Clin Invest Suppl* 1968;97:77–89.

157. Boyum A. Separation of lymphocytes, lymphocyte subgroups and monocytes. A review. *Lymphology* 1968;10:71–76.

158. Boyum A. Isolation of leukocytes from human blood. Further observations—methyl cellulose, dextran, and Ficoll as erythrocyte aggregating agents. *Scand J Clin Invest Suppl* 1968;97:31–49.

159. Kaplan ME, Clark C. An improved rosetting assay for detection of human T lymphocytes. *J Immunol Methods* 1974;5:131–135.

160. Stewart D. Rosetting as a method for separating human B cells and T cells. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:49–64.

161. Johnson AH. Anti-Fab separation of T and B lymphocytes. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:71–75.

162. Grier JO, Abelson LA, Mann DL, et al. Enrichment of B lymphocytes using goat anti-human F(ab′)2. *Tissue Antigens* 1977;10:236.

163. Lou CD, Cunniffe KJ, Garovoy MR. Histocompatibility testing by immunologic methods. Humoral assays. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:1087–1097.

164. Sleasman JW, Leon BH, Aleixo LF, et al. Immunomagnetic selection of purified monocyte and lymphocyte populations from peripheral blood mononuclear cells following cryopreservation. *Clin Diagn Lab Immunol* 1997;4:653–658.

165. Leffell MS. Assessment of purity and viability. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:38–48.

166. Patarca R, Maher K, Goodkin K, et al. Cryopreservation of peripheral blood mononuclear cells. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:281–286.

167. Bell RS, Bourret LA, Bell DF, et al. Evaluation of fluorescein diacetate for flow cytometric determination of cell viability in orthopaedic research. *J Orthop Res* 1988;6:467–474.

168. Ross DD, Joneckis CC, Ordonez JV, et al. Estimation of cell survival by flow cytometric quantification of fluorescein diacetate/propidium iodide viable cell number. *Cancer Res* 1989;49:3776–3782.

169. Crowley JP, Rene A, Valeri CR. The recovery, structure and function of human blood leukocytes after freeze-preservation. *Cryobiology* 1974;13:442–447.

170. Birkland SA. The influence of different freezing procedures and different cryoprotective agents on the immunological capacity of frozen-stored lymphocytes. *Cryobiology* 1976;13:442–447.

171. Strong DM, Sell KW. Functional properties of cryopreserved lymphocytes. *Cryoimmunology* 1976;62:81–92.

172. Gramatzki M, Strong DM, Grove SB, et al. Cryopreserved human cultured T cells as responder cells for the qualitative measurement of interleukin-2. Improvement of the assay. *J Immunol Methods* 1982;53:209–220.

173. Lorentzen DF. Cell preservation. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:154–157.

174. Kleeberger CA, Lyles RH, Margolick JB, et al. Viability and recovery of peripheral blood mononuclear cells cryopreserved for up to 12 years in a multicenter study. *Clin Diagn Lab Immunol* 1999;6:14–19.

175. Maluish AE, Strong DM. Lymphocyte proliferation. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*. Washington, DC: American Society for Microbiology, 1986:274–281.

176. Lavergne JA, del Llano AM. Assessment of cellular activation by flow cytometric methods. *Pathobiology* 1990;58:107–117.

177. Palutke M, KuKuruga D, Tabaczka P. A flow cytometric method for measuring lymphocyte proliferation directly from tissue culture plates using Ki-67 and propidium iodide. *J Immunol Methods* 1987;105: 97–105.

178. Bontadini A, Conte R, Dinota A, et al. Mixed lymphocyte reactions evaluated by means of bromodeoxyuridine incorporation. *Haematologica* 1990;75:7–11.

179. Allsopp CE, Nicholls SJ, Langhorne J. A flow cytometric method to assess antigen-specific proliferative responses of different subpopulations of fresh and cryopreserved human peripheral blood mononuclear cells. *J Immunol Methods* 1998;214:175–186.

180. Dean JH, Connor R, Herberman RB, et al. The relative proliferation index as a more sensitive parameter for evaluating lymphoproliferative responses of cancer patients to mitogens and alloantigens. *Int J Cancer* 1977;20:359–370.

181. Cram LS, Gomez ER, Thoen CO, et al. Flow microfluorometric quantitation of the blastogenic response of lymphocytes. *J Histochem Cytochem* 1976;24:383–387.

182. Bach F, Segall M. Pooled stimulating cells as a “standard stimulator” in mixed lymphocyte culture. *Transplantation* 1974;22:117–135.

183. Mickelson EM, Hansen JA. The mixed lymphocyte culture (MLC) test. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:339–356.

184. Dupont B, Hansen JA, Yunis EJ. Human mixed lymphocyte culture reactions. Genetics, specificity, and biological implications. *Adv Immunol* 1976;23:107–202.

185. Reinsmoen NL. Histocompatibility testing by immunologic methods. Cellular. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:1080–1086.

186. Kanda H, Matsuura T, Akiyama T, et al. A flow-cytometric DNA analysis in mixed lymphocyte reaction. *Transplant Proc* 1985;17:2157–2158.

187. Slezak SE, Horan PK. Cell-mediated cytotoxicity. A highly sensitive and informative flow cytometric assay. *J Immunol Methods* 1989;117:205–214.

188. Talbot D, Shenton BK, Givan AL, et al. A rapid, objective method for the detection of lymphocytotoxic antibodies using flow cytometry. *Clin Genet* 1987;31:25–34.

189. Zarcone D, Tilden AB, Cloud G, et al. Flow cytometry evaluation of cell-mediated cytotoxicity. *J Immunol Methods* 1986;94:247–255.

190. Van Leeuwen A. Two color fluorescence. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:202–205.

191. Van Rood JJ, Van Leeuwen A, Ploem JS. Simultaneous detection of two cell populations by two-color fluorescence and application to the recognition of B-cell determinants. *Nature* 1976;204:795–797.

192. Tardif GN, LeFor WM. Single color fluorescence for HLA-DR typing. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:206–211.

193. Carpenter CB. Lymphocyte-mediated cytotoxicity. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:304–307.

194. Thomas FT, Lee HM, Lower RR, et al. Immunological monitoring as a guide to the management of transplant recipients. *Surg Clin North Am* 1979;59:253–281.

195. Beatty PG. The induction and assay of human cytotoxic T lymphocytes in vitro. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:399–414.

196. Dunkley M, Miller RG, Shortman K. A modified 51Cr release assay for cytotoxic lymphocytes. *J Immunol Methods* 1974;6:39–51.

197. Morales A, Ottenhof PC. Clinical application of a whole blood assay for human natural killer (NK) cell activity. *Cancer* 1983;52:667–670.

198. Moy PM, Holmes EC, Golub SH. Depression of natural killer cytotoxic activity in lymphocytes infiltrating human pulmonary tumors. *Cancer Res* 1985;45:57–60.

199. Bonavida B, Bradley TP, Grimm EA. The single cell assay in cell-mediated cytotoxicity. *Immunol Today* 1983;4:196–200.

200. Bradley TP, Vonavida B. Mechanism of cell-mediated cytotoxicity at the single cell level. III. Evidence that cytotoxic T lymphocytes lyse both antigen-specific and -nonspecific targets pretreated with lectins or periodate. *J Immunol* 1981;126:200–213.

201. Huttunen K, Lionen J. The determination of natural killer activity of human peripheral blood lymphocytes by measuring the DNA synthesis of proliferating target cells (K 567 cell line). *Acta Pathol Microbiol Scand (C)* 1983; 91:197–201.

202. McGinnes K, Chapman G, Marks R, et al. A fluorescence NK assay using flow cytometry. *J Immunol Methods* 1986;86:7–15.

203. Shi T-X, Tong MJ, Bohman R. The application of flow cytometry in the study of natural killer cell cytotoxicity. *Clin Immunol Immunopathol* 1987;45:356–365.

204. Vitale M, Neri LM, Comani S, et al. Natural killer function in flow cytometry. II. Evaluation of NK lytic activity by means of target cell morphological changes detected by right angle light scatter. *J Immunol Methods* 1989;121:115–120.

205. Herberman RB. Natural killer cell activity and antibody-dependent cell-mediated cytotoxicity. In: Zachary AA, Teresi GA, eds. *ASHI laboratory manual*, 2nd ed. New York: American Society for Histocompatibility and Immunogenetics, 1990:308–314.

206. Stelzer GT, Robinson JP. Flow cytometric evaluation of leukocyte function. *Diagn Clin Immunol* 1988;5:223–231.

207. Bjerknes R, Bassoe C-F, Sjursen H, et al. Flow cytometry for the study of phagocyte functions. *Rev Infect Dis* 1989;11:16–33.

208. Duque RE, Ward PA. Quantitative assessment of neutrophil function by flow cytometry. *Anal Quant Cytol Histol* 1987;9:42–48.

209. Campbell DE, Douglas SD. Phagocytic cell functions. I. Oxidation and chemotaxis. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of Clinical laboratory immunology*. Washington, DC: ASM Press, 1997:320–328.

210. O'Gorman MRG. Phagocytic cell functions. II. Adhesion and phagocytosis. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:320–328.

211. Maderazo EG, Ward PA. Leukocyte chemotaxis. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:290–294.

212. Nelson RD, Quie PG, Simmions RL. Chemotaxis under agarose. A new simple method for measuring chemotaxis and spontaneous migration of human polymorphonuclear leukocytes and monocytes. *J Immunol* 1975;115:1650–1656.

213. Glasser L, Fiederlein RL. Neutrophil migration under agarose. Quantitation and variables. *Am J Clin Pathol* 1979;72:956–962.

214. Glasser L, Fiederlein RL. The effect of various cell separation procedures on assays of neutrophil function. *Am J Clin Pathol* 1990;93:662–669.

215. Lehrer RI, Cline MJ. Interaction of *Candida albicans* with human leukocytes and serum. *J Bacteriol* 1969;98:996–1004.

216. Lehrer RI. Measurement of candidacidal activity of specific leukocyte types in mixed cell populations. I. Normal, myeloperoxidase-deficient, and chronic granulomatous disease neutrophils. *Infect Immunol* 1970;1970:42–47.

217. Quie PG, White JG, Holmes B, et al. *In vitro* bactericidal capacity of human polymorphonuclear leukocytes. Diminished activity in chronic granulomatous disease of childhood. *J Clin Invest* 1967;46:668–679.

218. Southwick FS, Stossel TP. Phagocytosis. In: Rose NR, Friedman H, Fahey JL, eds. *Manual of clinical laboratory immunology*, 3rd ed. Washington, DC: American Society for Microbiology, 1986:326–331.

219. Bassoe CF, Solsvik J, Laerum OD. Quantitation of single cell phagocytic capacity by flow cytometry. In: Laerum OD, Lindmo T, Thorud E, eds. *Flow cytometry IV*. Bergen: Universitetsforlaget, 1980:170–174.

220. Bassoe C-F, Laerum OD, Solberg CO, et al. Phagocytosis of bacteria by leukocytes measured by flow cytometry. *Proc Soc Exp Biol Med* 1983;174:182–186.

221. Bassoe C-F, Laerum OD, Glette J, et al. Simultaneous measurement of phagocytosis and phagosomal pH by flow cytometry. Role of polymorphonuclear neutrophilic leukocyte granules in phagosome acidification. *Cytometry* 1983;4:254–262.

222. Bassoe C-F, Bjerknes R. Phagocytosis by human leukocytes, phagosomal pH, and degradation of seven species of bacteria measured by flow cytometry. *J Med Microbiol* 1985;19:115–125.

223. Bjerknes R, Bassoe C-F. Human leukocyte phagocytosis of zymosan particles measured by flow cytometry. *Acta Pathol Microbiol Immunol Scand (C)* 1983;91:341–348.

224. Bjerknes R. Flow cytometric assay for combined measurement of phagocytosis and intracellular killing of Candida albicans. *J Immunol Methods* 1984;72:229–241.

225. Steinkamp JA, Wilson JS, Saunders GC, et al. Phagocytosis. Flow cytometric quantitation with fluorescent microspheres. *Science* 1982;215:64–66.

226. Ogle JD, Noel JG, Sramkoski RM, et al. Phagocytosis of opsonized fluorescent microspheres by human neutrophils. A two-color flow cytometric method for the determination of attachment and ingestion. *J Immunol Methods* 1988;115:17–29.

227. Dunn PA, Tyrer HW. Quantitation of neutrophil phagocytosis, using fluorescent latex beads. Correlation of microscopy and flow cytometry. *J Lab Clin Med* 1981;98:374–381.

228. Wilson RM, Galvin AM, Robins RA, et al. A flow cytometric method for the measurement of phagocytosis by polymorphonuclear leucocytes. *J Immunol Methods* 1985;76:247–253.

229. Rothe G, Valet G. Phagocytosis, intracellular pH, and cell volume in the multifunctional analysis of granulocytes by flow cytometry. *Cytometry* 1988;9:316–324.

230. Nebe CT. Phagocytic activity of monocytes and granulocytes. In: Robinson JP, ed. *Handbook of flow cytometry methods*. New York: Alan R. Liss, 1993:181–182.

231. Jarstrand C, Lahnborg G, Berghem L. Human granulocyte function during dextran infusion. *Acta Chir Scand Suppl* 1979;489:279–285.

232. Ochs HD, Igo RP. The NBT slide test. A simple screening method for detecting chronic granulomatous disease and female carriers. *J Pediatr* 1973;65:77–82.

233. Cline MJ. A new white cell test that measures individual phagocyte function in a mixed leukocyte preparation. 1. A neutrophil defect in acute myelocytic leukemia. *J Lab Clin Med* 1973;81:311–316.

234. Hafeman DG, McConnell HM, Gray JW, et al. Neutrophil activation monitored by flow cytometry. Stimulation by phorbol diester is an all-or-none event. *Science* 1982;215:673–675.

235. Thorell B. Intracellular red-ox steady states as basis for cell characterization by flow cytofluorometry. *Blood Cells* 1980;6:745–751.

236. Thorell B. Flow cytometric analysis of cellular endogenous fluorescence simultaneously with emission from exogenous fluorochromes, light scatter, and absorption. *Cytometry* 1981;2:39–43.

237. Bass DA, Parce JW, Dechatelet LR, et al. Flow cytometric studies of oxidative product formation by neutrophils: a graded response to membrane stimulation. *J Immunol* 1983;130:1910–1917.

238. Rothe G, Valet G. Flow cytometric analysis of respiratory burst activity in phagocytes with hydroethidine and 2′,7′-dichlorofluorescin. *J Leukoc Biol* 1990;47:440–448.

239. Fattorossi A, Nisini R, Le Moli S, et al. Flow cytometric evaluation of nitro blue tetrazolium (NBT) reduction in human polymorphonuclear leukocytes. *Cytometry* 1990;11:907–912.

240. Trinkle LS, Wellhausen SR, McLeish KR. A simultaneous flow cytometric measurement of neutrophil phagocytosis and oxidative burst in whole blood. *Basic Appl Histochem* 1987;2:62–68.

241. Etzioni A. Adhesion molecule deficiencies and their clinical significance. *Cell Adhes Commun* 1994;2:257–260.

242. Bogomolski-Yahalom V, Matzner Y. Disorders of neutrophil function. *Blood Rev* 1995;9:183–190.

243. Mazzone A, Ricevuti G. Leukocyte CD11/CD18 integrins. Biological and clinical relevance. *Haematologica* 1995;80:161–175.

244. Malech HL, Nauseef WM. Primary inherited defects in neutrophil function. Etiology and treatment. *Semin Hematol* 1997;34:279–290.

245. Etzioni A. Leukocyte adhesion defects. *Arch Immunol Ther Exp (Warsz)* 1997;45:31–36.

246. Becker DJ, Lowe JB. Leukocyte adhesion deficiency type II. *Biochim Biophys Acta* 1999;1455:193–204.

247. Etzioni A, Doerschuk CM, Harlan JM. Of man and mouse. Leukocyte and endothelial adhesion molecule deficiencies. *Blood* 1999;94: 3281–3288.

248. Bassoe C-F, Solberg CO. Phagocytosis of *Staphylococcus aureus* by human leukocytes. Quantitation by a flow cytometric and a microbiological method. *Acta Pathol Microbiol Immunol Scand (C)* 1984;92:43–50.

249. Sjursen H, Bjerknes R, Halstensen A, et al. Serum opsonins to group B meningococci in patients and vaccinated volunteers. *Cytometry* 1987;8(Suppl 1):38.

250. Sjursen H, Bjerknes R, Halstensen A, et al. Flow cytometric assay for the measurement of serum opsonins to *Neisseria meningitidis* serogroup B, serotype 15. *J Immunol Methods* 1989;116:235–243.

251. Sjursen H, Bjerknes R, Halstensen A, et al. Serum opsonins to group B meningococci. *Acta Pathol Microbiol Immunol Scand (C)* 1987;95:283–289.

252. Ratnoff WD. Inherited deficiencies of complement in rheumatic diseases. *Rheum Dis Clin North Am* 1996;22:75–94.

253. Cicardi M, Bergamaschini L, Cugno M, et al. Pathogenetic and clinical aspects of C1 inhibitor deficiency. *Immunobiology* 1998;199:366–376.

254. Jelezarova E, Lutz HU. Assembly and regulation of the complement amplification loop in blood. the role of C3b-C3b-IgG complexes. *Mol Immunol* 1999;36:837–842.

255. Speth C, Wurzner R, Stoiber H, et al. The complement system. pathophysiology and clinical relevance. *Wien Klin Wochenschr* 1999;111:378–391.

256. Lindahl G, Sjobring U, Johnsson E. Human complement regulators. A major target for pathogenic microorganisms. *Curr Opin Immunol* 2000;12: 44–51.

257. Ahmed AE, Peter JB. Clinical utility of complement assessment. *Clin Diagn Lab Immunol* 1995;2:509–517.

258. Callegari PE, Williams WV. Laboratory tests for rheumatic diseases. When are they useful? *Postgrad Med* 1995;97:65–68, 71–74.

259. Giclas PC. Complement tests. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:181–186.

260. Jaskowski TD, Martins TB, Litwin CM, et al. Comparison of three different methods for measuring classical pathway complement activity. *Clin Diagn Lab Immunol* 1999;6:137–139.

261. Zwirner J, Wittig A, Kremmer E, et al. A novel ELISA for the evaluation of the classical pathway of complement. *J Immunol Methods* 1998;211:183–190.

262. Coppola R, Tombesi S, Cristilli P, et al. Comparison of two immunoassays for the complement protein C4b-binding protein in health and disease. *Int J Clin Lab Res* 1995;25:88–92.

263. Gemmell CH. A flow cytometric immunoassay to quantify adsorption of complement activation products (iC3b, C3d, SC5b-9) on artificial surfaces. *J Biomed Mater Res* 1997;37:474–480.

264. Rabson AR. Complement activation. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of clinical laboratory immunology*. Washington, DC: ASM Press, 1997:187–191.

62

Flow Cytometry

J. Philip McCoy Jr.

Flow cytometry has evolved from a research tool to an integral component of contemporary laboratory medicine. Flow cytometry, which offers the ability to examine multiple features of individual cells in rapid succession, has become useful, if not required, in the evaluation and diagnosis of many diseases. Key to the evolution of flow cytometry has been the development of instrumentation which simplified maintenance and quality control procedures. Additional factors that have spurred the transition into the clinical laboratory include the development of new and better fluorochromes and a broad array of monoclonal antibodies. These have permitted more multiparameter analyses of specimens and thus a more precise dissection of disease processes.

This chapter reviews the current status of flow cytometry in laboratory medicine. Only cursory mention is made of instrument design and theory of operation; several excellent texts thoroughly cover these topics (1, 2). While numerous assays are discussed in terms of methodology and interpretation, a goal of this chapter also is to convey, in a broader sense, an appreciation for the abilities, potential abilities, and limitations of flow cytometry. It is essential to bear in mind that flow cytometric analyses frequently are adjunctive laboratory assays: clinical diagnoses must integrate the flow cytometric findings with other laboratory findings and the clinical presentation. Interpretation of data derived from flow cytometric analysis is nearly impossible without appropriate correlation of morphological and clinical parameters. Finally, flow cytometry is a single technology for which many applications have been found. The clinical applications described in this text are by no means complete; rather, they reflect those that are most commonly performed. For a more in-depth discussion of these applications, a variety of excellent texts are available (3,4,5 and 6).

DISEASE STATES

Part of "62 - Flow Cytometry"

Flow cytometric assays are used in the evaluation of a multitude of disease states to examine a wide variety of cellular features. The features examined include cell surface proteins and carbohydrates, intracellular antigens, cellular enzymes, and nucleic acids. Of these, the most common flow cytometric assays involve the detection of cell surface antigens. This is most often accomplished through the use of fluorochrome-conjugated monoclonal antibodies directed against lineage-specific (or -associated) antigens and/or antigens associated with maturation or clonality.

The plethora of these reagents has led to a series of workshops designed to better classify the distribution of the antigens on hematopoietic cells. The result of these workshops is a series of cluster designations (CD), which describe the distribution of a number of antigens defined by monoclonal antibodies from a variety of sources. Representative cluster designations, their distribution, and commonly available antibodies for each CD are given in Table 62.1 (7). Armed with these tools for the understanding of hematopoiesis, one can utilize flow cytometric immunophenotyping to characterize disease processes.

As mentioned above, flow cytometric assays used in the clinical laboratory are by no means limited to cell surface immunophenotyping. Assays have been described for the quantitation of cellular DNA (for ploidy and proliferation studies), RNA (for reticulocyte maturation), cell surface receptors (e.g., estrogen), carbohydrates (by lectin binding), intracellular antigens (e.g., cytoplasmic mu and Tdt), ion fluxes (such as calcium fluxes in triggering target cells), and phagocytosis, to name a few. In many instances, these parameters may be examined concomitantly with cell surface markers. This section will illustrate the application of some of these flow cytometric assays in the study of selected disease in the clinical laboratory.

Leukemia

Acute

The diagnosis of acute leukemia is dependent upon clinical presentation of the patient as well as the examination of a number of laboratory parameters such as morphology and cytochemistry. A detailed description of the complete laboratory evaluation of acute leukemia is found in Chapter 43, Acute Leukemias and Myelodysplastic Syndromes. Immunophenotypic analyses of abnormal cells in peripheral blood and bone marrow by flow cytometry complements other laboratory studies and has substantially enhanced our ability to classify leukemias (8,9,10,11,12 and 13). These phenotypic analyses are primarily focused upon delineating cell lineage and stage of maturation of the cells in question, thus helping to delineate lymphoid from myeloid acute leukemias, and to subclassify leukemias within each lineage. A schema for surface marker expression as a function of lineage and maturation can be constructed as illustrated in Fig. 62.1. Immunophenotyping data is particularly useful to establish a diagnosis when standard morphology or cytochemistry provides ambiguous, conflicting, or incomplete results.

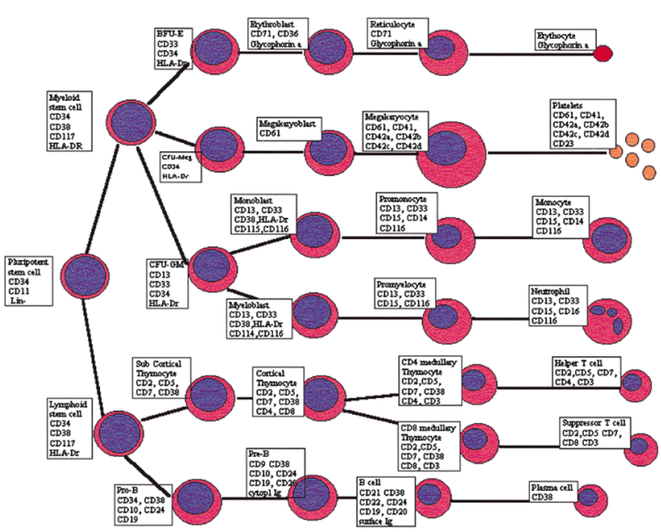


FIGURE 62.1. A maturation schema for hematopoiesis in relationship to marker expression.

TABLE 62.1. CD ANTIGENS AND CELLULAR DISTRIBUTION

Antigen	Cellular Distribution	Name/Membrane Component
CD1a	thymocytes, Langerhans dendritic dermal cells, B lymphocyte subset	gp49
CD1b	thymocytes, dendritic dermal cells, B lymphocyte subset	gp45
CD1c	thymocytes, dendritic dermal cells, B lymphocyte subset	gp43
CD2	all T lymphocytes, 95% thymocytes, natural killer cells	CD58 (LFA-3) receptor, gp50
CD2R	activated T lymphocytes	CD2 epitopes restricted to activated T cells
CD3	mature T lymphocytes	CD3 receptor complex (5 chains), gp/p26,20,16
CD4	T lymphocyte subset, helper/inducer T lymphocytes, monocytes	MHC Class II/HIV receptor, gp59
CD5	all T lymphocytes, some B lymphocyte subsets	gp67
CD6	mature T lymphocytes, B lymphocyte subset	gp 100
CD7	all T lymphocytes	gp40
CD8	T lymphocyte subset, cytotoxic/suppressor T lymphocytes, Natural killer cells	MHC Class 1 receptor gp72 heterodimer (gp30,32)
CD9	monocytes, B lymphocytes, platelets, B progenitors	p24
CD10	B lymphocytes, common acute lymphoblastic leukemia, progenitors, granulocytes, germinal center B cells	neutral endopeptidaseB gp100, CALLA
CD11a	leukocytes	LFA-1, gp 180/95
CD11b	monocytes, granulocytes, natural killer cells, macrophages	C3bi receptor (CR3), gp 155/95
CD11c	monocytes, granulocytes, macrophages, B lymphocyte subset natural killer cells,	gp 150/95 heterodimer
CDw12	monocytes, granulocytes, platelets	p90–120
CD13	granulocytes, monocytes	aminopeptidase N, gp 150
CD14	monocytes, granulocytes, follicular dendritic reticulum cells, epidermal Langerhans cells	gp55
CD15	granulocytes, monocytes	Hapten X on 3-FAL
CD15s	granulocytes, monocytes	Sialyl-Lewis X
CD16	granulocyte, natural killer cells, macrophages	lgGFcRIII (low affinity), gp50–65
CDw17	granulocytes, platelets, monocytes	lactosylceramide
CD18	leukocytes, broad	ß-chain to CD11a,b,c
CD19	all B lymphocytes, some myeloid progenitors	gp95
CD20	B lymphocytes, follicular dendritic reticulum cells, all peripheral B lymphocytes, some late B progenitors, some Reed-Sternberg cells	gp35, p37/32, ion channel?
CD21	B lymphocytes, B lymphocyte subset, follicular dendritic reticulum cells	C3d Epstein-Barr virus receptor, gp 140, CR2 for C3d fragment
CD22	B lymphocytes, B progenitors, and most peripheral B lymphocytes to myelin	gp 135, homology weak on germinal center B, assoc.gp (MAG) cytoplasmic B/surface B lymphocyte subset
CD23	B lymphocyte subset, activated monocytes, follicular dendritic reticulum cells, eosinophil	IgE FcRil receptor (low affinity), gp45–50
CD24	B lymphocytes, granulocytes, all B lymphocytes including progenitors, polymorphs	gp45,55,65 heterotrimer
CD25	activated T lymphocytes, B lymphocytes, monocytes	II-2 receptor ß-chain, gp55
CD26	activated T lymphocytes	dipeptidylpeptidas e IV, gp110
CD27	T lymphocytes	p55(dimer)
CD28	T lymphocyte subset	gp44
CD29	T lymphocyte subset (others)	VLA ß, Integrin ß-chain, platelet GPIIa, p135
CD30	activated T lymphocytes, activated B lymphocytes, activated Hodgkin's and Reed-Sternberg cells	gp120, K1-1 prototype
CD31	monocytes, granulocytes, platelets, T lymphocytes, bone marrow cells, epithelium, B lymphocytes	gp 140, platelet GPIIa
CD32	monocytes, granulocytes, platelets, B lymphocytes	FcRII receptor, gp40
CD33	myeloid leukemia, myeloid progenitors, monocytes	gp67
CD34	myeloid and lymphoblastic leukemia, some bone marrow cells, endothelium, lymphocyte progenitor cells	gp 115
CD35	granulocytes, monocytes, B lymphocytes, red blood cells, follicular dendritic reticulum cells	CRI for C3b fragment of C3,p220
CD36	monocytes, macrophages, platelets, early erythroid	GP90, platelet GPIV
CD37	B lymphocytes, all peripheral B lymphocytes, some late B progenitors, some leukocytes, some T lymphocytes, monocytes	GP40–52
CD38	early B lymphocytes, T lymphocyte, T blasts, plasma cells, leukocyte progenitors	p45/12 heterodimer
CD39	B lymphocytes, endothelial cells, activated T cells and NK cells	gp78
CD40	B lymphocytes, interdigitating reticulum cells, dendritic cells, carcinomas	gp50, homology to NGF receptor
CD41	platelets, megakaryocytes	CPIIb/IIIa complex (receptor for fibrinogen)
CD42a	platelets, megakaryocytes	GPIX, gp22
CD42b	platelets, megakaryocytes	GPIb receptor for factor VIII antigen, gp135/25
CD42c	platelets, megakaryocytes	gp22, GPIb receptor for factor VIII antigen
CD42d	platelets, megakaryocytes	gp85, GPV
CD43	T lymphocytes, most leukocytes except B cells	leukosialin, sialoopherin, gp95
CD44	most leukocytes, red blood cells	Pgp-1, pg80–95
CD44R	erythrocytes, monocytes, epithelial cells	Gp130,160,190
CD45	leukocytes	leukocyte common antigen, T200

CD45RA	T lymphocyte subset, B lymphocytes, granulocytes, monocytes, myeloid cells	restricted T200, gp220
CD45RB	T lymphocyte subset, B lymphocytes, granulocytes, monocytes, macrophage	restricted T200, gp220
CD45RO	T lymphocyte subset, B lymphocytes, granulocytes, monocytes	restricted T200, gp 180
CD46	broad	membrane cofactor protein (MCP), gp66/56
CD47	broad	Integrin associated protein (IAP)
CD48	leukocytes	gp43, GPI-linked, BLAST-1
CD49a	activated T and B cells, monocytes	Integrinα1, VLAα1, gp210
CD49b	activated T and B cells, platelets, monocytes	Integrinα2, VLAα2, GP1a, gp160
CD49c	broad, non-hematopoietic cells	Integrinα3, VLAα3, gp150
CD49d	broad	Integrinα4, VLAα4, gp150
CD49e	platelets, monocytes, neutrophils	Integrinα5, VLAα5, gp150
CD49f	platelets, T lymphocyte subset, monocytes, megakaryocytes	Integrinα6, VLAα6, gp140, platelet GPIc
CD50	leukocytes, broad	ICAM-3, gp120, PI-linked
CD51	platelets, some B lymphocytes, monocytes, endothelial cells	VNR-α-chain
CD52	leukocytes	Campath-1, gp21-28
CD53	leukocytes	gp32-40, PI-linked
CD54	broad, activated, endothelial cells	ICAM-1
CD55	broad	decay accelerating factor (DAF)
CD56	natural killer cells, some T lymphocytes	Gp180, N-CAM
CD57	subset of natural killer cells, T lymphocyte subset	gp110, HNK1
CD58	leukocytes, epithelial cells, endothelial cells	LFA-3, gp55-65
CD59	broad	gp18-20
CDw60	T lymphocyte subset, platelets	9-0-acetyl GD3
CD61	platelets, megakaryocytes, B lymphocytes	integrin ß3-, VNR ß-chain, platelet GPIIIa
CD62E	endothelial cells	ELAM-1, E- selectin
CD62L	NK cells, some T and B cells, monocytes	L-selectin
CD62P	activated platelets	GMP-140 (PADGEM), gp140 P-selectin
CD63	activated platelets, monocytes, granulocytes,	gp53 T lymphocytes, B lymphocytes
CD64	monocytes	FcγRI, gp75
CD65	granulocytes, monocytes	poly-n- acetylactosamine
CD65s	granulocytes, monocytes	poly-sialyl-n- acetylactosamine
CD66a	granulocytes	Biliary glycoprotein (BGP) gp 180-200
CD66b	granulocytes	CGM-6, gp95-100
CD66c	granulocytes	NCA, gp90-95
CD66d	granulocytes	CGM-1, gp30
CD66e	granulocytes	CEA, gp 180-200
CD66f	granulocytes	Pregnancy-specific antigen (PSG)
CD67	now CD66b	
CD68	macrophages, activated platelets, monocytes	gp 110
CD69	activated B and T lymphocytes	gp32/38, AIM
CD70	activated B lymphocytes, T lymphocytes, Reed-Sternberg cells	Ki-24, CD27 ligand
CD71	proliferating cells, erythroid cells	transferrin receptor
CD72	B lymphocytes	gp42
CD73	B lymphocyte subset, T lymphocyte subset	ecto-5'-nucleotidase, gp69
CD74	B lymphocytes, monocytes	Class II assoc. invariant chain, gp41/35/33
Cdw75	mature B lymphocytes, T lymphocyte subset	p53
Cdw76	mature B lymphocytes, T lymphocyte subset	gp85/67
CD77	restricted B lymphocytes, Burkitt's lymphoma	globotriasylceramide (Gb3)
Cdw78	B lymphocytes, macrophages	unknown
CD79a	B lymphocytes	Igα, mb-1 gene product, gp33
CD79b	B lymphocytes	Igß, gp40
CD80	B cell subset	B7, BBI, gp60
CD81	broad on hematopoietic cells	TAPA-1, p26
CD82	broad on hematopoietic cells	R2, gp50-53
CD83	dendritic cells, Langerhans cells	gp43
Cdw84	monocytes, macrophages, platelets, some B and T cells	p73
CD85	plasma cells, monocytes, B cell subset	P120
CD86	activated B cells, monocytes	gp80
CD87	activated T cell, monocytes, NK cells	UPAR, gp50-65
CD88	granulocytes, monocytes, macrophages	C5a-R, gp40
CD89	granulocytes, monocytes, macrophages, lymphoid subsets	Fcα-R, gp55-70
CD90	hematopoietic stem cell subset	Thy-1, gp18
CD91	monocytes, macrophages, erythroblasts	α2MR, p600
Cdw92	granulocytes, monocytes	p70
Cdw93	granulocytes, monocytes	p120
CD94	NK cells, T cell subsets	Kp43, gp43
CD95	activated lymphocytes, broad	FAS, APO-1, gp42
CD96	activated T and NK cells	TACTILE, gp160
CD97	activated lymphocytes, monocytes	p74,80,89
CD98	broad, activated lymphocytes	4F2, gp80/40
CD99	broad on all hematologic cells	E2/MIC2, gp32
CD99R	broad on all hematologic cells	gp32
CD100	broad on hematopoietic cells	gp150

CD101	activated T cells, monocytes, granulocytes, macrophages	p140
CD102	lymphocytes, monocytes, platelets, endothelial cells	ICAM-2, gp55-65
CD103	mucosal associated lymphocytes	integrin αE, HML-1, gp150/25
CD104	thymocytes, epithelial cells, endothelial cells, monocytes	integrin ß4, gp205
CD105	activated monocytes, endothelial cells	endoglin, gp95
CD106	endothelial cells	VCAM-1, INCAMI 10, gp100
CD107a	platelets	LAMP-1, gp110
CD107b	platelets	LAMP-2, gp110
Cdw108	activated T cells	JMH, gp80
CD109	activated T cells, endothelial cells, platelets	gp170/150
CD110	reserved	
CD111	reserved	
CD112	reserved	
CD113	reserved	
CD114	granulocytes, monocytes, platelets	G-CSFR, gp130
CD115	monocytes, macrophages	M-CSFR, gp150
CD116	granulocytes, monocytes	GM-CSFR, gp80
CD117	hematopoietic stem cells	C-kit, SCFR, gp145
CD118	reserved	
CDw119	broad	IFNγ1-R, gp90
CD120a	broad	TNFR-1, gp55
CD120b	broad	TNFR-2, gp75
CD121a	broad	ILI-Rl
CD121b	broad	ILI-R2
CD122	lymphocytes, monocytes	IL2R, gp75
CDw123	myeloid cells and progenitors	IL3Rα, gp70
CD124	mature T and B cells, some hematopoietic precursors	IL4R, gp140
CDw125	myeloid progenitors	IL5Rα, gp60
CD126	activated B cells, plasma cells, monocytes, epithelial cells	IL6Rα, gp80
CD127	precursor B cells, T cells, monocytes	IL7Rα, gp75
CDw128	neutrophils, monocytes, fibroblasts, T cell subset	IL8R, gp60
CD129	reserved	
CD130	broad	gp130
CDw131	monocytes, granulocytes, early B cells	ß chain of IL3, IL5, GM-CSF receptors, gp95-120
CD132	lymphocytes, monocytes, granulocytes	γ chain of IL2, IL4, IL7, IL,9, IL15 receptors, gp64
CD133	reserved	
CD134	activated T cells	OX40, gp50
CD135	hematopoietic progenitor cells	FLT3, flk2, stk1, gp130-150
CDw136	broad	MSPR
CDw137	T lymphocytes	4-1BB, gp30
CD138	B lymphocytes, plasma cells	syndecan-1, GP 85-92
CD139	B cells, follicular dendritic cells	gp209-228
CD140a	broad	PDGFRα, gp180
CD140b	endothelial cells	PDGFRß, gp180
CD141	endothelial cells, smooth muscle cells, granulocytes, monocytes thrombomodulin	gp100
CD142	endothelial cells, epithelial cells, monocytes, keratinocytes	gp 45
CD143	endothelial cells, epithelial cells, macrophages	ACE, gp 170
CD144	endothelial cells,	gp135
CDw145	endothelial cells, basement membrane	gp25-90-110
CD146	endothelial cells, activated T cells, dendritic cells	MUC18, SENDO, gp113
CD147	endothelial cells, leukocytes	neurothelin, basigin, gp50-60
CD148	broad	DEP1, HPTPη, gp260
CDw149	lymphocytes, monocytes	MEM-133
CDw150	T and B lymphocytes, thymocytes, dendritic cells	SLAM, IPO3, gp75-95
CD151	endothelial cells, epithelial cells, platelets	PETA-3, gp27
CD152	activated T cells	CTLA-4, gp44
CD153	activated T cells, granulocytes, activated macrophages	CD30 ligand, gp40
CD154	activated T cells	CD40 ligand, gp32-39
CD155	monocytes, macrophages, thymocytes	PVR, gp90
CD156	monocytes, macrophages, granulocytes	ADAM8, gp60
CD157	bone marrow stromal cells, monocytes, granulocytes	BST-1, gp42-45
CD158a	NK cell, T cell subset	P58.1, p50.1, gp58/50
CD158b	NK cell, T cell subset	P58.2, p50.2, gp58/50
CD158c	reserved	
CD159	reserved	
CD160	reserved	
CD161	NK cell, T cell subset	NKRPIA, gp60
CD162	monocytes, granulocytes, T cells, B cell subset	PSGL-1, gp110
CD163	macrophages, monocytes	M130, gp130
CD164	monocytes, granulocytes	MGC24
CD165	platelets, some T cells, NK cells	AD2, gp37
CD166	endothelial cells, activated T cells, monocytes	ALCAM, gp100

Table 62.2 and Table 62.3 illustrate common immunophenotypes associated with lymphoid and myeloid leukemias, respectively. Acute leukemias are not limited to these immunophenotypes, rather these point to typical findings within each subtype. Some variation occurs even within subtype. It also is fairly common in acute leukemias to have aberrant expression of a marker associated with another lineage (14,15,16 and 17). Caution must be taken in attempting to interpret immunophenotyping data in the absence of clinical information and other laboratory findings. An example of this is the similarity in immunophenotypes of precursor B-ALL and chronic myelogenous leukemia in acute lymphoid blast crisis.

TABLE 62.2. COMMON IMMUNOPHENOTYPES ASSOCIATED WITH VARIOUS TYPES OF ACUTE LYMPHOBLASTIC LEUKEMIA (ALL)

Precursor B ALL	Pre-B ALL	B-ALL	T-ALL
CD10+	CD10+	CD10±	CD1+
CD19+	CD19+	CD19+	CD2+
CD20-	CD20±	CD20+	cytoplasmic CD3+
CD24+	CD24+	CD24+	CD5+
CD34+	CD34-	CD34-	CD4/CD8 dual +
HLA-Dr+	HLA-DR+	HLA-DR+	CD10±
Tdt+	Tdt±	Tdt-	Tdt+
	cytoplasmic IgM+	surface Ig+	

TABLE 62.3. COMMON IMMUNOPHENOTYPES ASSOCIATED WITH VARIOUS TYPES OF ACUTE MYELOGENOUS LEUKEMIA (AML)

AML-M0	AML-M1	AML-M2	AML-M3	AML-M4	AML-M5	AML-M6	AML-M7
CD13+	CD13+	CD13+	CD13+	CD13+	CD13+	CD34±	CD34+
CD33+	CD33+	CD33+	CD33+	CD33+	CD33+	HLA-Dr+	CD41+
CD34+	CD34+	CD34+	CD34±	CD34±	CD34±	Glycophorin A±	CD61+
HLA-Dr+	HLA-Dr+	HLA-Dr+	HLA-Dr-	HLA-Dr+	HLA-Dr+	CD71±	CD13±
Tdt±	CD15-	CD15-	CD15+	CD15+	CD15+	CD13±	CD33±
CD15-				CD14±	CD14±	CD33±	
MPO-*				CD4±	CD4±		

* MPO negative by cytochemistry, but frequently positive by flow cytometry

Flow cytometric immunophenotyping is useful to identify both biphenotypic and biclonal leukemias (18,19,20,21,22 and 23). The former form of leukemia displays cell surface markers, or other characteristics, of two lineages or types of leukemias on the same cell whereas the latter form of leukemia involves the identification of two separate leukemic clones within the same patient. There is not a sharp delineation between what is considered lineage-aberrant marker expression and biphenotypic leukemia (8). While there is no universally agreed upon standard for distinguishing between these two entities, one widely accepted criterion is the biphenotypic leukemia must have at least two lineage-specific markers from each lineage present. Aberrant marker expression on the other hand, would have only one lineage-specific marker. Biclonal leukemias, by contrast, display two distinct populations of cells each bearing a separate and clearly defined array of surface markers. Biclonal leukemias may have different clones from the same or differing lineages (21,22 and 23).

In addition to immunophenotyping of cell surface markers, flow cytometry may also be used to examine intracellular components of leukemic cells. Flow cytometric assays have been described for intracellular enzymes such as myeloperoxidase (MPO) and terminal deoxynucleotidyl transferase (TdT), and cytoplasmic immunoglobulin and cytoplasmic CD3 (24). Terminal deoxynucleotidyl transferase (TdT), a template-independent DNA polymerase, is useful as a marker for ALLs. While present in the majority of ALLs, TdT is not found in the majority of AMLs and mature non-Hodgkin's lymphomas. As part of a panel including immunophenotyping, morphological examination, and cytochemistries, TdT is a useful tool in diagnosis and in the detection of recurrent disease. MPO, an enzyme marker found in myeloid, but not lymphoid, leukemias is another useful intracellular marker in leukemia to distinguish lineage. Detection of MPO has historically been performed by cytochemical methods, however antibody-based cytometric methods for this are now available and appear to offer greater sensitivity as well as ease of testing (25).

Measurement of cellular DNA content is often of use in the study of acute leukemias either alone or in conjunction with surface marker analysis. Measurement of DNA content serves to identify aneuploid populations of tumor (although minor DNA alterations and chromosome translocations cannot be detected) and to estimate proliferation rates. Aneuploidy is a strong indi-

cator of malignancy and, as such, can aid in diagnosis when immunophenotypic and morphological examinations are not conclusive (26, 27). Aneuploidy also can be used as a means to detect recurrent disease (26), and prognosis has been associated with ploidy in selected diseases. For example, Look and colleagues (28) demonstrated a better prognosis in childhood acute leukemia associated with a hyperdiploid (DI >1.16) DNA content. The observations by Look and colleagues in childhood ALL do not apply in adult ALL. Thus, it should be noted that ploidy-prognosis relationships are by no means universal; each type of lesion and group of patients must be independently assessed to determine the validity of ploidy evaluation in prognosis.

Minimal Residual Disease

Flow cytometry is one of several technologies that is being applied to detect residual disease in patients with acute leukemia prior to overt relapse (29,30,31 and 32). Flow cytometric assays to detect minimal residual disease (MRD) have relied on aberrant marker expression to distinguish leukemic blasts from normal marrow precursors. Here aberrant marker expression is considered any one of the following: expression of lineage aberrant markers, unusual intensity of marker expression, failure of cells to express expected lineage markers, and asynchronous marker combinations. In order to detect surface marker combinations which are aberrant or which are somewhat unique to the leukemic cells, a rather large panel of antibodies must be analyzed at diagnosis. This serves to “fingerprint” the leukemic cells in sufficient detail that they may be distinguished in small numbers from normal hematologic precursor cells. CD45/log side scatter gating (33), described below, is particularly useful in this endeavor. Arguably, in some patients, a unique “fingerprint” may not be identifiable due to the leukemic blasts close immunophenotypic appearance to normal cells. In those patients where a “fingerprint” is identified, sequential analyses of remission marrows will provide the best identification of residual disease.

Chronic

Chronic lymphoproliferative disorders (CLPD) are a group of diseases characterized by atypical proliferations of small, mature lymphoid cells. A number of diseases fall within this category and there is a varying degree to which immunophenotyping is used in the diagnosis of each disorder. Chronic lymphocytic leukemia (CLL) is the most common CLPD in Europe and North America and thus represents the most frequent of these diseases encountered in the clinical flow cytometry laboratory.

CLL are monoclonal proliferations of B cells with cell surface immunoglobulins (34,35 and 36). CLLs are mature B-cell lesions, typically identified by the expression of CD19, CD20, HLA-DR, cytoplasmic immunoglobulin heavy chain, and the weak, but monoclonal, expression of heavy and light immunoglobulin chains on the cell surface (seldom seen in B-ALLs). In contrast to B-ALLs, CLLs do not express the CALLA antigen; however, most are positive for CD5. Thus, CD19+, CD20+, CD5+ cells with dim, clonal, light chain expression is highly characteristic of CLL. Additionally, the presence of CD23 helps to distinguish CLL from mantle cell lymphoma, which expresses a similar immunophenotype (37). By contrast, prolymphocytic leukemia is generally CD5 negative with stronger expression of clonal surface light chain (38, 39) (Table 62.4).

TABLE 62.4. IMMUNOPHENOTYPES ILLUSTRATIVE OF CHRONIC LYMPHOPROLIFERATIVE DISORDERS

B-CLL	B-PLL	HCL	PCD	T-LGLD	NK-LGL
CD19+	CD19+	CD19+	CD19-	CD2+	CD2+
CD20+	CD20+	CD20+	CD20-	CD3+	CD3-
CD5+	CD5 usually-	CD5–	CD22–	CD5+	CD5-
CD22+	CD22+	CD22+	CD38+	CD7+	CD7+
CD23+	CD23-	CD23–	cytoplasmic light chain	CD8+	CD11c+
CD45 bright	CD45 bright	CD11c+	+	CD11c±	CD16+
light chain dim+	light chain strong+	CD25+	CD45 dim or-	CD16±	CD56+
HLA-Dr+	HLA-Dr+	CD103+		CD56±	CD57+
		light chain+		CD57±	

CLL, chronic lymphocytic leukemia; PLL, prolymphocytic leukemia; HCL, hairy cell leukemia; PCD, plasma cell dyscrasia; T-LGLD, T lineage large granular lymphocyte disorder; NK-LGLD, natural killer large granular lymphocyte disorder

Myelodysplasia and Chronic Myelogenous Leukemia

The role of flow cytometry and immunophenotyping in the evaluation of myelodysplasia is somewhat uncertain (40, 41). CMLs

are classically described by the presence of a consistent chromosomal abnormality, the Philadelphia chromosome (Ph1). Surface marker analysis is of limited utility during the chronic phase of this disease, as many immature and mature myeloid elements are present, but it is of substantial value during the acute phase (blast crisis) (42, 43). The acute phase is similar to an acute leukemia, and here the cells may bear either lymphoid or myeloid characteristics. Delineation of these two forms of blast crisis is critical because each responds differently to chemotherapeutic regimens. Lymphoid variants generally bear the CD10, CD19, and HLA-DR surface markers, and the myeloid variants express CD13, CD33, and HLA-DR. Additionally, blasts can display erythroid or megakaryocytic features and thus can be identified through the use of glycophorin A or Plt-1, respectively.

Lymphoma

The lack of a specific antigen, or combination of antigens, which can identify Hodgkin's lymphoma has severely limited the use of flow cytometry in the diagnosis of these lymphomas, however, flow cytometric immunophenotyping is quite useful in the diagnosis and staging of non-Hodgkin's lymphomas (NHL) (8, 44,45,46,47 and 48). The majority of NHL are of the B cell lineage, flow cytometric identification of a light chain restriction is important in establishing a diagnosis of lymphoma in many cases. In addition to identifying clonality, immunophenotyping also is useful in subclassifying B-NHL, as illustrated in Table 62.5. In addition to studying the lymph nodes themselves, examination of the bone marrow by immunophenotyping permits a more precise and sensitive method for detection of marrow involvement, and thus staging, than possible by morphologic studies alone.

TABLE 62.5. IMMUNOPHENOTYPES REPRESENTATIVE OF LYMPHOMAS

Follicular	Mantle Cell	MALT	Burkitt's
CD19+	CD19+	CD19+	CD19+
CD20+	CD20+	CD20+	CD20+
CD5-	CD5+	CD5-	CD5-
CD10+	CD10-	CD10-	CD10+
light chain+	CD23-	light chain+	light chain+
	CD43+		
	cyclin D1+		
	light chain+		

T lineage lymphomas may be immunophenotyped by flow cytometry, but clonality is much more difficult to establish (44). As with all lymphoma immunophenotyping, flow cytometric analysis of putative T cell lesions must be conducted only in the context of appropriate morphologic studies. Within a morphologic diagnosis of lymphoma, surface markers may be used to delineate B lineage from T lineage. T lineage lymphomas generally display aberrant immunophenotypes, particularly characterized by the absence of an expected pan T-cell antigen, and this may be used as additional. In mature forms with aßT cell receptor complexes (TCR), antibodies to variable ß the TCR regions of may be used as indicators of clonality (49). Because of the large number of antibodies required for these studies, this task is generally not feasible in the clinical laboratory.

Analysis of the cellular DNA content of lymphomas may be of use when one is presented with a difficult differential diagnosis such as a reactive node, although, in general, molecular studies to assess clonality are preferable. As with leukemias, aneuploidy in lymphomas is very strongly correlated with malignancy. High proliferation rates in lymphomas are closely associated with high-grade disease and a high frequency of an initial complete remission of short duration (50,51 and 52). Recent surveys reveal that, despite the increased utilization of molecular techniques in the diagnosis of malignant lymphoma, flow cytometric assessment of DNA content remains a widely performed test (53, 54).

Immunodeficiencies

Human Immundeficiency Virus (HIV) and Acquired Immunodeficiency Syndrome

Perhaps the most common application of immunophenotyping is in the diagnosis and therapy of HIV disease and acquired immunodeficiency syndrome (AIDS) (54). The HIV virus has been shown to infect and destroy helper T (CD4+) lymphocytes (55, 56). Thus laboratory findings in support of a diagnosis of AIDS include a decrease in the absolute number of the CD4+ cells and thus a low ratio of CD4+ to CD8+ T lymphocytes (57,58 and 59). Normal individuals generally have a CD4/CD8 ratio of approximately 2.0; AIDS patients quite often have ratios less than 0.5. Enumeration of CD4+ cells provides a direct indicator of disease progression in HIV-infected individuals and of survival in patients with AIDS (58, 59). CD4 enumeration and HIV viral load

testing usually is performed in parallel to assess both the quantity of virus detectable in the patient and the effect of the virus on the immune system. In addition to diagnostic testing, both assays are extremely important in monitoring of therapy, particularly therapy including protease inhibitors where declining viral loads and increasing CD4 counts may be observed (60, 61).

Immunophenotyping also may be used to examine subsets of both CD4 T cells and CD8 T cells in HIV infection which may be important in disease progression. CD45RA, CD45RO, CD28, and CXCR4 have, among other antigens, been examined on helper or suppressor subsets during HIV infection (62,63 and 64). While these additional markers to subset helper and suppressor T cells have provided interesting information regarding the nature of HIV disease and are being used in various clinical studies, none have become a universal standard of practice thus far.

Historically, a CD4 percentage from a flow cytometer and an absolute lymphocyte count obtained from a hematology analyzer, or by manual methods, had to be used to calculate an absolute CD4 count (a “dual-platform” method). Recently, though the use of bead standards placed in the blood to be analyzed, absolute CD4 counts as well as percentages can be obtained directly from a flow cytometer (a single-platform method) (65). Data has been published demonstrating less variability in CD4 counts obtained using single platform methods in comparison to a dual platform technique (66, 67).

Congenital Immunodeficiencies

Surface marker analysis is used in conjunction with other laboratory tests, such as those for serum immunoglobulins, cell function, and enzyme deficiencies, to examine a variety of primary immunodeficiency syndromes such as severe combined immunodeficiency (SCID), DiGeorge syndrome, and common variable immunodeficiency (CVID) (68,69,70,71 and 72). Immunophenotyping in patients with suspected immunodeficiencies may be performed to: (i) assess the lineages and subsets of leukocytes present; (ii) detect the absence of normal cell surface antigens; and (iii) to evaluate the presence of inducible antigens.

Determination of the lineages and subsets of cells present on the peripheral blood is of value in immunodeficiencies characterized by an absence, or diminution in numbers, such as DiGeorge's syndrome (thymic hypoplasia). This is a disease of T cell deficiency and immunophenotyping can be used to pinpoint the paucity of T lymphocytes in the peripheral blood (72). It is important to note, however, that a diagnosis of DiGeorge's syndrome cannot be made solely on the basis of immunophenotyping data.

The absence of normal cell surface antigens is observed in immunodeficiencies such as leukocyte adhesion deficiency (LAD) (68, 73, 74). LAD-type 1 patients, who suffer from recurrent bacterial infections, have a defect in leukocyte integrins including CD11a, CD11b, and CD11c in conjunction with CD18. In most cases, leukocytes from LAD-1 patients display decreased expression of these antigens as well as decreased function. Cell surface immunophenotyping is quite useful in the overall evaluation of the patients.

The detection of inducible antigens may be used in the overall evaluation of some immunodeficiency diseases such as DiGeorge's syndrome and SCID. An example of this is the expression of CD69 in response to mitogen stimulation (75, 76). Increases in CD69 expression, an early activation marker, occur on lymphocytes as early as 4 to 6 hours of culture with mitogens and thus may be used as a measure of cellular activation. In addition to providing quicker data than classical methods, such as thymidine incorporation, CD69 staining for lymphocyte activation also readily permits multiparameter studies in which cellular activation can be examined on various lymphoid subsets. Caution must be exercised in using this approach, as CD69 expression may be induced under conditions other than mitogenesis and might be less sensitive than thymidine incorporation (77).

Various immunodeficiencies are associated with specific enzymatic defects in various cell types. Chronic granulomatous disease (CGD), for example, is classically associated with a defect in the oxidative metabolism of neutrophils, resulting in a lack of hydrogen peroxide production. This defect is classically detected using nitroblue tetrazolium tests (NBT), but also may be assayed by flow cytometry using dichlorofluorescein diacetate (DCF-DA)-a nonfluorogenic substrate capable of entering live cells (78). Production of hydrogen peroxide within the cell converts DCF-DA into an insoluble, fluorogenic compound, dichlorofluorescein. Alternative, dihydrorhodamine can be used for the same purpose (79). Phagocytosis, which should be comparable in CGD and normal individuals, also may be examined through the use of beads or opsonized fluorochrome labeled bacteria. It is possible to assay both oxidative burst and phagocytosis simultaneously in a single flow cytometric assay, screening for both possible defects in individual cells (80).

Transplantation

Solid Organ

Flow cytometry can be used either pretransplantation or posttransplantation in support of solid organ transplantation. Flow cytometric assays for the detection of alloantibodies are widely used in the pretransplantation setting. These assays involve the incubation of donor cells with recipient serum followed by a fluorochrome-conjugated antihuman reagent. Compared to complement dependent-cytotoxicity crossmatching, flow cytometric assays are more sensitive, objective, rapid and permit concomitant examination of multiple types of cells (81, 82).

Flow cytometric assays for post-transplantation are directed at monitoring therapy and predicting or detecting rejection. Flow cytometric analysis of cell surface markers has been used to assist in delineating graft rejection *vis á vis* viral [cytomegalovirus (CMV)] infection, to monitor the efficacy of cytoreductive/immunosuppressive therapy, and to detect antibodies to murine immunoglobulin during OKT3 therapy. In the study of graft rejection versus CMV infection, several studies have demonstrated an association between an increase in interleukin 2 (IL-2) receptor-positive, CD3+ T lymphocytes and graft rejection (83, 84). There is controversy concerning the value of total T-cell enumeration, CD4/CD8 ratios, and the expression of HLA-DR (Ia) on T cells in discriminating rejection from CMV infection (85, 86). The monoclonal antibody OKT3 is often used as antirejection therapy to block the activity of cytotoxic T lymphocytes (87).

This therapy relies on the establishment of sufficient plasma levels of OKT3 to bind the T cells and remove them from the circulation (at least temporarily) without the development of antimurine antibodies. Flow cytometric assays are available to monitor CD3+ cells and plasma (or serum) levels of OKT3 and, therefore, to provide an overall assessment of the likelihood of successful OKT3 therapy (88, 89). Similarly, through the use of blocking assays (90,91 and 92), it is possible to use flow cytometry to detect the presence of anti-OKT3 antibodies that may develop in response to therapy with this murine antibody.

Flow cytometry also can be used to detect intracellular cytokines, such as IL-2 and interferon-gamma (93, 94). These cytokines are produced in activated T lymphocytes, and their production is impaired by many antirejection agents. Multiparametric flow cytometry assays are capable of detecting alterations in the levels of intracellular cytokines in lymphoid subsets, permitting such parameters to be used in monitoring the efficacy of immunosuppressive agents.

Hematopoietic Stem Cell

Transplantation of pluripotent hematopoietic stem cells has become an accepted form of therapy for a myriad of diseases (95,96 and 97). Common to this therapy is the eradication of the patient's existing immune system with subsequent introduction of new immune cells from either an autologous or allogeneic source. Keys to successful engraftment of the new immune system include transplantation of adequate numbers of stem cells to the host and transplantation of stem cells of appropriate purity. Both of these tasks are well-suited to monitoring by flow cytometric immunophenotyping.

Enumeration of hematopoietic stem cells is now generally accomplished through the use of CD34 monoclonal antibodies (98,99,100,101,102,103,104 and 105). Although the precise immunophenotype of the true pluripotent stem cells remains controversial, CD34 is the most universally accepted marker for identifying these cells, or at least cells which are capable of engraftment and reconstituting the immune system. Simplistically, enumeration of CD34+ cells is cell surface immunophenotyping similar to that performed for CD4 enumeration or a number of other assays. Unfortunately, accurate CD34 determination is more complex due a variety of factors such as the low percentage of stem cells in most specimens, the intensity of CD34 staining, and the recognition of different epitopes by various CD34 antibodies. CD34 monoclonal antibodies have been assigned three classes (I, II, and III) based on the sensitivity of the epitopes recognized to enzyme digestion (102). To promote inter-institutional consistency of therapy and design of cooperative studies, attempts have been made to design “consensus” protocols for the enumeration of CD34+ stem cells, and commercial products have become available for stem cell enumeration (Table 62.6). While there is no agreement concerning which of these protocols is optimal, there is increasing acceptance that stem cell enumeration is best performed with class III antibodies with phycoerythrin as the fluorochrome. Fig. 62.2 illustrates the data analysis strategy for CD34+ stem cells enumeration proposed by the International Society for Hematotherapy and Graft Engineering (ISHAGE).

TABLE 62.6. STEM CELL ENUMERATION PROTOCOL COMPARISON

Protocol (ref)	Mabs and Fluorochromes Used	Number of Events Collected
ISHAGE (100)	CD34 (class II or III for PE, class III for FITC), CD45	75,000 CD45+
SIHON (103)	CD34, CD14, CD66e, LDS-751	50,000 nucleated
Milan (98)	CD34	50,000 gated
Procount* (104)	CD34 (class III), CD45, nucleic acid dye	Variable
Stem-kit** (105)	CD34, CD45 (modified ISHAGE)	75,000 CD45+

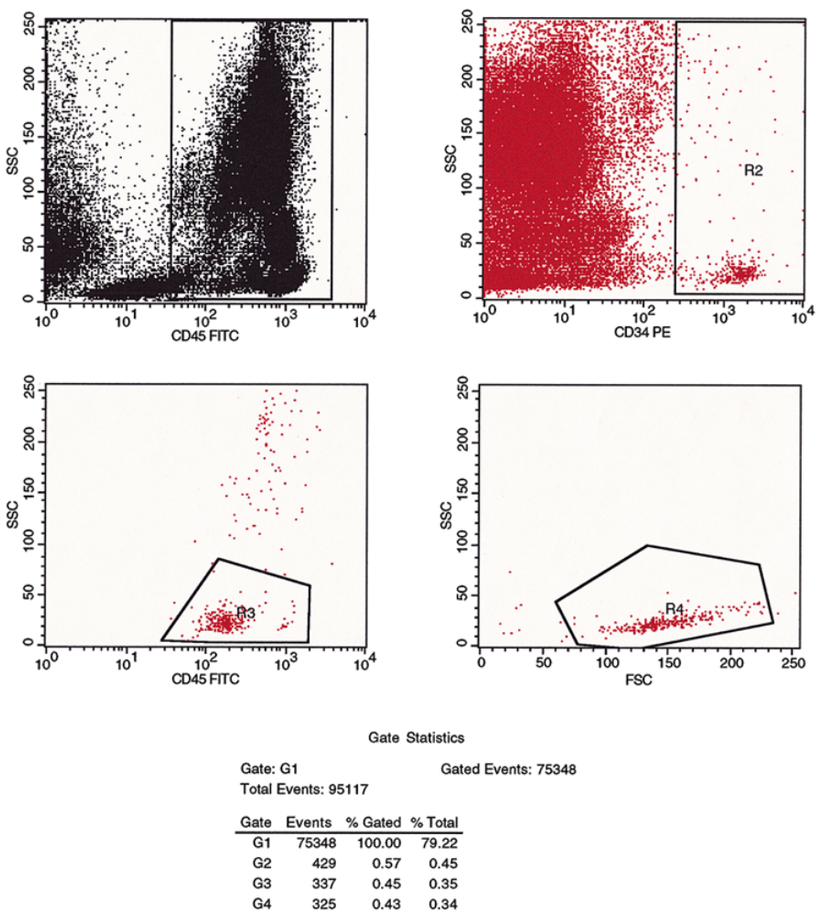
ISHAGE, International Society for Hematotherapy and Graft Engineering; SIHON, Foundation for Immunophenotyping in Hemato-Oncology of the Netherlands

Key clinical issues associated with poststem cell transplantation management include assessing the engraftment of the marrow, preventing, detecting, and managing graft versus host disease, and in certain instances, detecting minimal residual disease. Engraftment is often assessed by bone marrow biopsies and differential counts on smears or peripheral blood. Immunophenotyping may be used as an ancillary assay to provide a more comprehensive picture of the reconstituted immune system in these patients. Flow cytometric quantitation of reticulocyte RNA also has been proposed as a method of predicting successful engraftment of bone marrow (43). Here thiazole orange is used to stain the RNA in reticulocytes, providing an index of maturity in these cells, which is thought to be an early indicator of generally functioning marrow.

Graft versus host (GVH) disease in allogeneic marrow transplants is a major concern with grave consequences. Here the transplanted immune system recognizes the host as foreign. Reduction of mature T cells in the donor stem cells reduces the incidence of GVH, and immunophenotyping is often of use in monitoring T-cell depletion of these preparations before transplantation (107). A variety of studies have presented immunophenotypic correlates or predictors of GVH, but none of these has yielded an assay that has become a widespread and standardized practice in the monitoring of these patients (108).

Antibody-Mediated Diseases

There are autoimmune diseases characterized by the production of antibodies that react with, and destroy, specific cells or tissue. Included among these diseases are several forms of thrombocytopenia and neutropenia. Examples of the former are neonatal alloimmune thrombocytopenia, transfusion-induced thrombocytopenia, and autoimmune thrombocytopenia purpura; examples of the latter include transfusion-induced neutropenia and neonatal isoimmune neutropenia. In the evaluation of these diseases, it is important to establish whether the paucity of cells is because of immune-mediated destruction or lack of production or maturation.



Gate Statistics

Gate: G1 Gated Events: 75348
Total Events: 95117

Gate	Events	% Gated	% Total
G1	75348	100.00	79.22
G2	429	0.57	0.45
G3	337	0.45	0.35
G4	325	0.43	0.34

FIGURE 62.2. An example of one method for CD34+ stem cell enumeration, the ISHAGE protocol. Specimens are stained with CD45 to identify leukocytes and CD34 to identify stem cells. Stem cells are identified and quantitated using a series of gates. Gating parameters include forward light scatter, side light scatter, CD45 staining, and ultimately CD34 staining.

Flow cytometry is well suited for assaying peripheral blood specimens for the presence of circulating antibody in the serum or plasma or antibody bound *in vivo* to the cell surfaces, using indirect and direct immunofluorescence techniques, respectively (109,110,111,112,113 and 114).

Paroxysmal Nocturnal Hemoglobinuria

Paroxysmal nocturnal hemoglobinuria (PNH), a rare, chronic hemolytic disease, is from a deficiency of cell surface glycoproteins that are glycosylphosphatidylinositol-anchored. Classic laboratory tests for PNH include the sucrose hemolysis test and the acid hemolysis test (Ham's test). CD55 (decay accelerating factor) and CD59 (membrane inhibitor of reactive lysis, protectin) are two glycosylphosphatidylinositol-anchored membrane glycoprotein whose expressions are diminished in PNH. In addition to being expressed on erythrocytes in normal individuals, these two markers also are present on granulocytes. The defect in PNH patients that diminishes the expression of CD55 and CD59 on erythrocytes, and contributes to their lysis, also diminishes the expression of these glycoproteins on granulocytes, but without resulting lysis. A flow cytometric assay has been developed to quantitate CD55 and CD59 expression on granulocytes for the diagnosis of PNH (115, 116). The specificity of

these antibodies together with the quantitative aspects of flow cytometry permit a rapid, accurate diagnosis and identification of partially (PNH II) or completely (PNH III) deficient cells.

Abnormal Hematopoiesis

The laboratory evaluation of anemia frequently includes reticulocyte enumeration as a measure of hematopoiesis. The normal response of bone marrow to anemia is the increased production of erythrocytes, although this is not observed when the anemia is the result of abnormal marrow function, which is often because of vitamin or iron deficiencies (hypoproliferative states). Thus, reticulocyte enumeration is often of use in determining the cause of anemia and monitoring the effect of therapy. Reticulocyte enumeration also has been proposed as a method of assessing the function of engrafted bone marrow. Reticulocytes are immature red blood cells that, although lacking nuclei, retain RNA in higher amounts than is found in more mature red blood cells. This residual RNA can be measured rapidly and quantitatively on a per-cell basis by flow cytometry using dyes such as thiazole orange (117,118,119 and 120). In addition to their use for the enumeration of reticulocytes, the quantitative aspects of the flow cytometric assay have been used to devise a reticulocyte maturity index (RMI) based upon the quantity of RNA in these cells (118).

Platelet Disorders

Platelet activation is associated with a variety of diseases, including adult respiratory distress syndrome (121), renal failure requiring dialysis (122), heparin-induced thrombocytopenia-thrombosis (123) and a number of cardiovascular ailments (124,125 and 126). Activation of platelets is generally assessed by flow cytometry using monoclonal antibodies directed against activation-related antigens, such as CD62P directed against P-selectin, a granule membrane protein GMP-140, and CD61 directed against the activated form of the platelet membrane glycoprotein IIb-IIIa complex (126,127 and 128). Flow cytometric assays of this type have several advantages over other approaches: subpopulations of platelets can be analyzed; multiple parameters may be examined simultaneously (129,130 and 131).

Non-Neoplastic Diseases of Erythroid Cells — Sickle Cell, ß-Thalessemia, Feto-Maternal Hemorrhage

While most flow cytometric assays are performed on nonerythroid components of hematologic specimens, recent advances have made certain studies of erythroid cells by flow cytometry practical and useful tests in the clinical laboratory. The development of monoclonal antibodies direct against fetal hemoglobin (HbF) has been the most prominent advance in this area and have permitted the development of flow cytometric assays for the detection of fetal hemoglobin (132,133,134 and 135). With the use of appropriate agents to permeabilize the cell, specific staining can be achieved with fluorochrome conjugated anti-HbF. This can be used as a replacement for the (acid elution) Kleihauer-Betke test, which has been the classical method of detecting HbF. The detection of cells containing fetal hemoglobin is of importance in monitoring therapy of sickle cells disease with butyrate and hydroxyurea, in detection of feto-maternal hemorrhage, and in the evaluation of various ß-thalessemias.

Ankylosing Spondylitis and Other Arthropathies

A strong genetic association exists between the presence of the HLA-B27 antigen and certain arthritic diseases including ankylosing spondylitis, a chronic, progressive form of arthritis affecting the spine. HLA-B27 typing has become a valuable component in the diagnosis of these diseases. The classical method of typing for HLA-B27 is through complement-mediated lymphocytotoxicity. This assay, while reliable, is time consuming and somewhat laborious. Several groups have described flow cytometric methods for HLA-B27 determination using fluorochrome-labeled anti-HLA-B27 monoclonal antibodies (136,137,138 and 139). The use of CD3 gating in multiparameter flow cytometry studies for HLA-B27 may improve the reliability and sensitivity of this method (137), as will the use of multiple monoclonal antibodies (139, 139).

Solid Tumors

DNA Analysis

The analysis of cellular DNA by flow cytometry is limited by the fact that only gross alterations in the DNA can be measured: translocations and small alterations in total DNA content escape detection. Furthermore, the need to dissociate solid tissue into monodispersed suspensions also places restrictions on this endeavor. The total G0/G1 nuclear DNA content (ploidy) and the rate of proliferation as determined by calculating the percentage of cells in each phase of the cell cycle are the two most common applications of flow cytometry in the clinical laboratory. Abnormal DNA content (i.e., aneuploidy) is perhaps the single cellular feature demonstrating the strongest correlation with neoplasia (140). While many neoplastic lesions are not aneuploid, the presence of aneuploidy is closely associated with neoplastic or preneoplastic lesions. Thus, in specimens where morphologic or phenotypic studies are ambiguous or not available, the detection of aneuploidy can assist in arriving at a diagnosis. This use may be extended to exfoliative cytology or body cavity effusions from suspected neoplasms in which the assessment of tissue architecture is not possible and sample preparation for flow cytometry is enhanced by the more dispersed nature of the cells. DNA studies also may be conducted on paraffin-embedded tissue as well as fresh tissue, thus allowing for retrospective studies or the ploidy assessment of a specimen at a later time without the necessity of procuring extra tissue. Studied of paraffin-embedded tissue, however, is fraught with technical difficulties and great care must be directed at obtaining reliable data (141,142,143 and 144).

Great controversy exists concerning the use of DNA content/cell cycle analysis as a prognostic marker for various types of solid tumors. Numerous studies have been published examining these parameters in a wide assortment of solid tumors, largely using paraffin-embedded tissue (reviewed in 145,146,147,148 and 149). Quite likely the technical artifacts associated with the study of paraffin-embedded material resulted in much of the conflicting data and conclusions. The American Society of Clinical Oncology concluded that insufficient data exists to warrant use of DNA content/cell cycle analysis as a prognostic marker in breast and colorectal cancer (150).

Despite this report and the controversy over the merits of DNA analysis, a recent survey indicated that DNA/cell cycle analysis remains to be a widely performed test (54).

Methodologies

Flow cytometry relies upon the intersection of a stream of cells or particles with a stable light beam of known wavelength and the subsequent detection of fluorescence or light scattering from cells. Before any specimens are analyzed on a flow cytometer, there must be assurance that the stream of cells and the light beam are intersecting at a point where the cells are optimally illuminated and the resulting signals are capable of being collected by the appropriate detectors. The process of bringing the sample stream and light beam into proper orientation is termed alignment. Improper alignment of an instrument can lead to weak, unusual, or even missing signals from cells. In most recently manufactured clinical instruments alignment is optimized by service engineers and does not need to be adjusted on a daily basis. Alignment should be checked daily by analyzing specimens (beads or cells) with known light scatter and fluorescence characteristics to ensure that these are detected in a consistent manner. In addition to alignment, calibration of flow cytometers must be performed prior to the analysis of any clinical specimens to determine how sensitive the detection of standardized signals is from day to day. Again, this is normally accomplished by analyzing standard beads or chicken erythrocytes. Fluctuations in the signal obtained with predetermined instrument settings and laser power generally are indicative of misalignment, a dirty optical path, or a failing light source, detector, or circuitry, among other causes. Reproducible calibration thus ensures standard sensitivity of a cytometer, which is particularly important in DNA studies or when studying markers that result in dim fluorescence. A third general procedure in preparing to analyze samples by flow cytometry is setting proper compensation for specimens stained with multiple fluorochromes. Because fluorochromes generally have wide emission spectra, they will overlap to various extent into the band of emission being collected for other fluorochromes. By using a set of standard beads with commonly used fluorochromes, this overlap can be minimized using an electronic compensation to subtract contaminating signals. These procedures are more thoroughly examined elsewhere (1,2,3,4 and 5, 151).

Gating

Flow cytometers are incapable of distinguishing cells of interest from other particles that may be present in specimens submitted for analysis unless procedures are implemented to uniquely distinguish these cells. Cytometers are capable of measuring (in addition to fluorescence signals) the forward and orthogonal light scatter signals emanating from cells, which are roughly indicative of particle size and granularity, respectively. Empirically, and through the use of staining with monoclonal antibodies, the light scatter patterns of normal peripheral blood have been determined and correlated with cell type (152). Armed with this knowledge, it is possible to examine only cells of interest for immunophenotyping, excluding the extraneous cells, debris, or particles through a process referred to as light scatter gating. Here the cytometer is directed to ignore events (particles) whose light scatter characteristics do not fall within a predefined set of parameters. In routine lymphocyte subset analysis of peripheral blood, erythrocytes can be lysed from the whole blood to reduce the number of particles flowing through the cytometer. The lymphocytes can then be analyzed separately from the monocytes and granulocytes, which also remain through the use of gating (Fig. 62.3). Failure to exclude cells of extraneous lineages from analysis can, and usually does, lead to misinterpretation of marker expression.

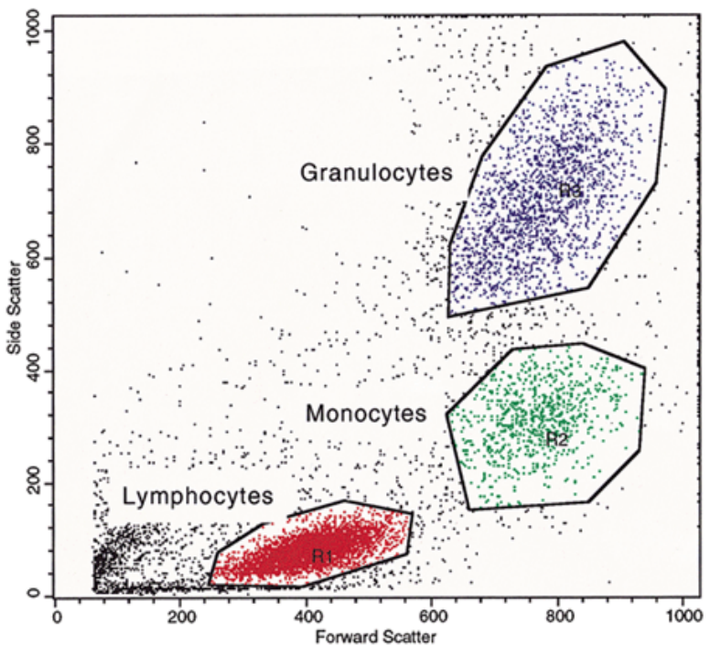


FIGURE 62.3. A dot density, two-parameter histogram illustrating the forward light scatter (FSC) on the x-axis and side (90°) light scatter (SSC) on the y-axis of human peripheral blood after lysis of the red blood cells. The red area surrounded by the box represents the gated lymphoid cells; monocytes are indicated by the green cells; the granulocytes are indicated in blue. Debris and small particles are to the left of the lymphoid cells.

While gating using forward versus side scatter signals has been the standard methodology for a decade, gating also may be performed using features other than, or in addition to, light scatter signals. For example, the identification of leukocytes in heterogeneous samples can be accomplished by gating on cells that stain with CD45. Leukocyte subsets and their precursors can be distinguished with high sensitivity in bone marrow using a combination of CD45 staining and side light scatter collected in a logarithmic mode rather than the more common linear mode (Fig. 62.4) (33).

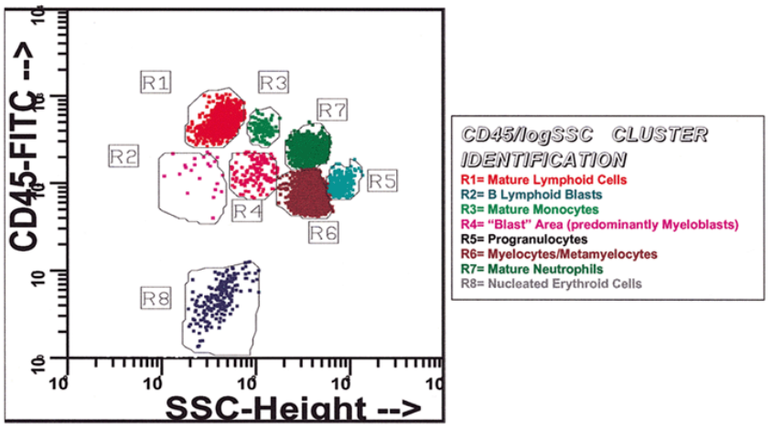


FIGURE 62.4. CD45 versus log SSC gating of human bone marrow. By using the pan-leukocyte reagent (CD45) and log side scatter, a better resolution of leukocyte subsets is obtained than with the FSC versus SSC gating demonstrated in Fig. 62.3.

Surface Marker Staining of Lymphocyte Subsets

The most common clinical applications of flow cytometry involve the analysis of cell surface markers on leukocytes. The key advantage to this technology is the examination of multiple cellular features in a rapid, quantitative manner on individual cells. In contrast to manual immunofluorescence techniques, flow cytometric immunophenotyping often examines multiple antigenic markers simultaneously to more precisely determine lineage, maturity, and/or state of activation of each cell type. This is possible through the use of multiple fluorochromes that have distinct emission wavelengths. Quantitation of the emissions at prespecified wavelengths will yield a value proportional to the amount of fluorochrome-conjugated antibody bound to the cells and, ultimately, to the amount of antigen present. Examples of commonly used fluorochromes are given in Table 62.7. Staining of cells with multiple antibodies is best accomplished using a direct immunofluorescence technique, i.e., having fluorochrome-conjugated antibodies rather than using a sandwich (indirect) technique to avoid cross-reactivity. Currently, three- and four-"colors" (fluorochromes) normally are examined together with light scatter signals, yielding five or six parameter studies (153,154 and 155).

TABLE 62.7. COMMONLY USED FLUOROCHROMES FOR IMMUNOPHENOTYPIC STUDIES

Fluorochrome	Excitation* (nm)	Emission* (nm)	Comments
Fluorescein isothiocyanate (FITC)	488	525	Low molecular weight, stable
Phycoerythrin (PE)	488	580	High quantum yield
Cy3	488	575	
Peridin chlorophyll protein (PerCP)	488	672	Bleached by high laser power
Tandem conjugate (ECD, Red613)	488	613	Tandem conjugate of PE and Tx Red
Tandem conjugate (Red670)	488	670	Tandem conjugate of PE and Cy5
Cy7	743	785	
Allophycocyanin	633	660	Requires He, Ne, krypton, or dye laser
Texas red	600	620	Requires dye laser
Cy5	633	667	Binds monocytes, requires He, Ne, krypton, or dye laser

* The values listed for excitation and emission wavelengths are not necessarily the maximum, but rather those that are commonly used in the laboratory.

Care must be taken in sample procurement and handling to avoid the introduction of artifacts into the analysis of surface markers. A wide variety of factors such as the anticoagulant used for the collection of the peripheral blood specimen and the collection time of day have been demonstrated to influence the enumeration of various lymphocyte subsets (151, 156,157 and 158). Further more, maintenance of the specimens at room temperature has been demonstrated to be a strict requirement; storage at 4°C for even short periods of time will induce capping of the antigens and alter the detection of lymphocyte subsets (151, 158,159 and 160).

Peripheral blood specimens are prepared for staining either by density-gradient separation of lymphocytes or by the lysis of red blood cells with agents such as ammonium chloride. Removal of erythrocytes from peripheral blood specimens is generally thought to be necessary because it allows the flow cytometer to analyze other cellular populations more quickly. The latter technique, commonly referred to as the "whole blood lysis" technique, generally is used because it is quicker, it may be performed before or after staining, and it permits the examination of monocytic and granulocytic components concomitant to the analysis of lymphocytes. The major questions regarding this approach are whether the specific antigen is altered by the lysing agent and whether there is a need for the cells of interest to be further concentrated. Phenotyping data from both methods of sample preparation are generally similar, although density-gradient procedures may produce loss of certain populations of cells (151, 161, 162). Recently, a consensus conference on the immunophenotyping of leukemias and lymphomas recommended the density gradient

separation **not** be used on these specimens (163). Nonetheless, approximately one fourth of laboratories performing lymphocyte immunophenotyping continue to use density gradient separation as a preparative technique (54). One approach to the flow cytometric analysis of lymphocyte subsets, not commonly in use, would preclude the need for removal of erythrocytes. This approach involves the use of a monoclonal antibody to the CD45 (common leukocyte) antigen as well as nucleic acid dyes to distinguish leukocytes from erythrocytes and adjusting the cytometer to "recognize" only the cells staining for this antigen (164). Subsetting is then accomplished through the use of additional monoclonal antibodies conjugated to other fluorochromes.

Staining of peripheral blood and bone marrow with monoclonal antibodies may be performed either prior to or after lysis of red blood cells. Most laboratories use the stain-then-lyse method (54). The amount of antibody used is variable, but in most instances should be of sufficient quantity to saturate all of the antigen sites on the cells. The time of incubation with these antibodies may also vary, but for most commonly examined cell surface markers, incubation time between 15 and 30 minutes are sufficient. After this incubation, the cells are washed with an isotonic solution and, in most clinical laboratories, fixed with paraformaldehyde. Besides preserving the cells better for later analysis, this fixation also has been demonstrated to inactivate the human immunodeficiency virus (151, 155, 165). Cells prepared in this manner, if stored at 4°C and protected from light, may be stored approximately 10 to 12 days without substantial loss of staining or light scatter properties.

Surface Marker Staining of Cells Other Than Lymphocytes

A myriad of cells other than lymphocytes can be analyzed for cell surface marker expression in the clinical flow cytometry laboratory. Each cell type will have unique specimen handling or staining requirements. Certain specimens, such as platelets, require that special care be taken in procurement, transport, and staining to reduce artifacts such as aggregation. Analysis of platelet activation is highly influenced by the type of anticoagulant used, the length of time prior to analysis, and the use of aspirin or similar products by the patient. In these assays (in contrast to lymphocyte immunophenotyping), it is often desirable to fix the specimen with paraformaldehyde prior to staining to minimize the artifacts introduced after procurement of the specimen (54, 60). Other nonlymphoid hematologic cells also present unique requirements. Monocytes, for example, bind the fluorochrome Cy5 independent of monoclonal antibodies against monocytoid antigens (166), therefore the use of Cy5-labeled monoclonal antibodies to study monocytes is not recommended. Granulocytes survive for much shorter periods of times *ex vivo* than do most lymphoid cells, thus studies of these cells must be performed in a more rapid manner.

The staining of platelets, neutrophils, and lymphocytes for autoantibodies against each of these cell types may be performed by using either direct or indirect immunofluorescence techniques. Direct staining, using a xenogenic antihuman immunoglobulin antibody, is used to detect autoantibody bound *in vivo* to the cells. To control for the specificity of the staining, an antibody directed against an irrelevant antigen that is derived from the same species as the anti-Ig antibody and matched for isotype and F/P ratio is used. An indirect staining method is used to detect autoantibodies in the circulation. In this approach, relevant donor cells are incubated with the serum (usually heat inactivated) from the patient with the suspected autoimmunity. After an appropriate length of incubation, the cells are washed and incubated with a fluorochrome-conjugated antihuman immunoglobulin reagent. Controls in this instance normally include identical staining with serum obtained from a normal donor. The isotype of the autoantibodies can be identified through the use of isotype-specific antihuman reagents.

Staining of Intracellular Antigens

A wealth of clinically relevant information is present within cells as well as on their cell surfaces. Flow cytometry may be used to detect intracellular antigens in addition to, or instead of, cell surface antigens. Examples of these include cytoplasmic immunoglobulin, cytoplasmic CD3, terminal deoxynucleotidyl transferase (Tdt), and myeloperoxidase (167,168 and 169). To stain intracellular components it is necessary to permeabilize the cell membrane for appropriate penetration of the monoclonal antibodies into the cell. If distinct cell surface antigen expression is to be examined at the same time as intracellular antigens, cell surface staining must precede permeabilization of the cells. Cells

may be permeabilized using a number of different reagents, including a number of proprietary formulations. Perhaps the most common nonproprietary permeabilization reagents are a combination of paraformaldehyde and Tween 20 (169, 170). Key concerns in selection of permeabilization reagents is not only assuring adequate accessibility of the monoclonal antibodies to the target antigens, but also ensuring that the target antigen is not destroyed or altered by the permeabilization reagent.

DNA Staining

The staining of hematological material for DNA content is easily accomplished because the cells to be examined are predominantly in single-particle suspension. This is not the case, however, for the DNA analysis of tissue derived from solid tumors. Here the specimen must be disaggregated to obtain a monodisperse suspension. This is accomplished by mechanical dissociation and/or enzyme (protease or collagenase) digestion, generally followed by filtration through a 50- to 70-mm mesh (171). The yields from dissociation procedures vary because of the procedure itself and the nature of tumor being dissociated. Therefore each laboratory needs to assess the merits of each procedure on each type of tumor to be routinely analyzed. Cellular losses should be, at worst, random, or at best, exclusive of the cells of interest. Dissociation procedures must be performed appropriately for the type of analysis to be conducted, e.g., enucleation procedures may be used if only DNA is to be studied, however if both surface markers and DNA are to be examined, then care must be taken to ensure an intact cell membrane.

DNA analyses are not limited to specimens freshly obtained from the patient or *in vitro* cultivation. Cryopreserved specimens generally yield coefficients of variation (CVs) approximating those obtained from fresh tissue. As mentioned previously, DNA assays also may be performed on formalin-fixed, paraffin-embedded tissue (172, 173). Here sample preparation and data interpretation are more rigorous than required in DNA analysis of fresh tissue. Sample preparation concerns begin at the time of original tissue fixation and embedding. The type of fixative has a profound effect on the resulting DNA histograms, and neutral-buffered formalin has been demonstrated to be the fixative of choice (174). However, even this fixative introduces some DNA artifacts, by reducing the accessibility of the DNA to the fluorescent dye (175). Sections usually are processed from paraffin blocks rather than the entire tissue. This generally is because of a desire to retain some tissue for any future medical or legal issues. Section thickness has a substantial effect on the resulting DNA histograms, and it generally is agreed that 50-mm sections provide the optimal results (176). The sections are deparaffinized using xylene or the equivalent and subsequently rehydrated using a series of decreasing ethanol solutions. The rehydrated tissue is now disassociated with enzymes (pepsin or trypsin) and filtered to obtain single cell suspensions. To reverse the effects of the formalin fixation, the cells should be subjected to heat treatment prior to staining (141). These cell suspensions are then stained with propidium iodide as described for fresh tissue.

A number of fluorescent dyes that may be used to stain cellular DNA are available; however, the dye most commonly used for this purpose in the clinical laboratory is propidium iodide (PI). PI intercalates into cellular DNA and yields fluorescence signals in the orange/red region proportional to the amount of DNA. It is readily excitable at 488 nm and is stable in solution, and these factors contribute to its widespread use. The staining achieved with propidium iodide is uniform and reproducible, yielding very small CVs. CVs reflect the “tightness” of a histogram peak; therefore, the tighter the CV of a G0/G1 DNA peak, the easier one can detect slight aneuploidy in a specimen (e.g., the sensitivity of an aneuploidy detection increases). A number of methods have been described for staining cellular DNA with propidium iodide (92, 93). The classic procedures (177, 178) rely upon lysis of the surface membrane of the cell, digestion of the RNA, and staining of the residual intact nuclei. Using this approach, a one-step method has been developed for DNA staining. Using a filtered solution of 10 mM Trizma base, 10 mM NaCl, 0.7 U/mL RNAse, 7′ 10 5 M PI, and 0.1% NP-40, cells can be lysed, enzyme-digested, and stained with PI concomitantly. This solution is added to the cells while vortexing, and the cells are then incubated at 4′C for 30 minutes. The stained nuclei can be analyzed immediately by flow cytometry or fixed with paraformaldehyde and stored for as long as a week until analysis. As DNA standards (discussed later), chicken or trout erythrocytes may be added to the specimen prior to staining.

Specimens stained with propidium iodide may be analyzed to determine the ploidy of the specimen and, in most instances, to assess proliferation. This is possible because cellular DNA is increased as cells progress through S-phase and enter G2/M. In G2/M, cells have 4N DNA and yield PI fluorescence twice that observed in G0/G1 nuclei (Fig. 62.5). Proliferation may be assessed in freshly-obtained specimens to determine the *in vivo* proliferative rate (particularly for anaplastic specimens), or specimens may be stimulated *in vitro*, then stained with PI to assess their proliferative response. The latter is performed in the evaluation of immunodeficiencies in which the response of peripheral blood lymphocytes to mitogen stimulation is important in pinpointing the immune defect. Patients' lymphocytes are cocultured with mitogen (e.g., phytohemagglutinin) for 72 hours, harvested, and stained with PI as described above. The percentage of cells in S-phase and G2/M is determined and compared to patient cells cultured without mitogen or to normal, resting lymphocytes stained in the same manner. This assay also may have a positive control consisting of normal lymphocytes cocultured with mitogen and responding in the predicted manner.

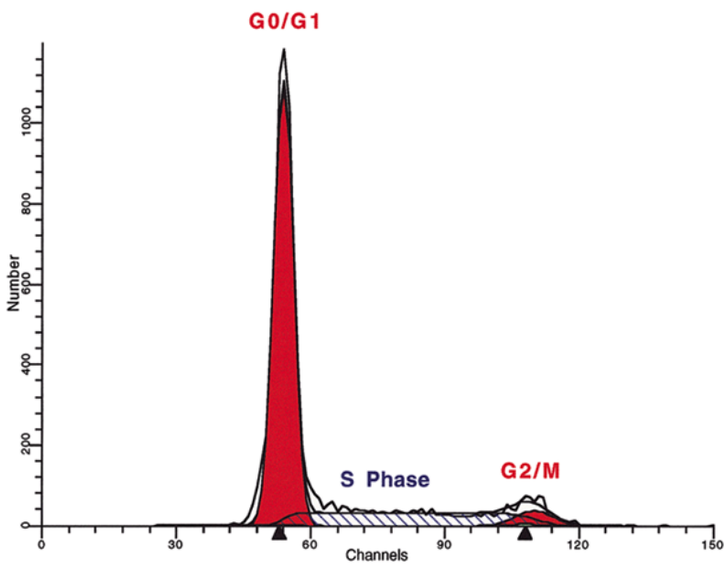


FIGURE 62.5. A representative one-parameter histogram of cell nuclei stained with propidium iodide. Phases of the cell cycle are as indicated. (Note the broad, irregular S phase.) Ploidy determinations among specimens are made by comparing the position of the G0/G1 peaks.

Combined Surface Marker/DNA Staining

There are obvious advantages to being able to concomitantly examine the surface markers and DNA content of the same cells. The proliferation of neoplastic cells can be assessed much more accurately if one can gate on the relevant cells through the use of a surface marker. The detection of aneuploidy in minimal residual disease becomes more sensitive if only cells of the pertinent lineage are examined. The assessment of proliferation can be examined by cell lineage and subset through the use of concomitant cell surface marker/DNA staining. A technique for the dual staining of cells for surface markers and DNA was introduced by Braylan et al. (179) and remains in use with varying degrees of modification. Cells from fresh tissue are incubated with the

fluorochrome (usually FITC)-conjugated antibody against the surface marker in sodium azide at 4°C for an appropriate length of time, then are washed thoroughly. The cells, suspended in PBS or medium, are vortexed while an equal volume of 100% ethanol is added. The cells are incubated in the ethanol for 1 hour at 4°C, washed, treated with RNAse, and finally stained with propidium iodide. A major disadvantage of this procedure is the loss of resolution in the light scatter patterns. Cell surface/DNA staining has also been achieved through the use of sequential paraformaldehyde/detergent treatment of cells or detergent treatment alone (180, 181).

Methods also have been introduced that permit the concomitant analysis of surface markers, cytoplasmic antigens, and DNA content while retaining most light scatter characteristics of the fresh cells (182, 183). In these methods, cells are stained for a specific surface marker, washed, then briefly fixed with a low concentration of paraformaldehyde and washed again. The cells are then permeabilized with 70% methanol and again washed. The cells may now be stained for the expression of intracellular antigens. If FITC was used to mark the surface antigen, a distinct fluorochrome such as allophycocyanin or Duochrome is used to stain the intracellular antigen. The cells are again washed and finally incubated with an RNAse/propidium iodide solution to stain the DNA. This method has been demonstrated to reliably stain all parameters, with a fair preservation of light scatter signals, and only minor diminution of the propidium iodide fluorescence because of the paraformaldehyde. A single-step proprietary reagent has also been reported which maintains light scatter properties while permitting staining of both surface markers and intracellular components (184).

Reticulocyte Staining

Reticulocyte staining is dependent on the staining of residual RNA in immature erythrocytes. A number of RNA-binding dyes have been used for this purpose, including pyronin Y, acridine orange, thioflavin T, and thiazole orange. Thiazole orange (117) is a membrane-permeable dye, excited at 488 nm, that intercalates into nucleic acids and fluoresces at 533 nm. Staining of reticulocytes with this dye is easily accomplished: 5 mL of whole blood is incubated for 1 hour at room temperature with 1 mL of 0.1 mg/mL thiazole orange in PBS containing EDTA and sodium azide.

Interpretation

Surface Marker Staining

Flow cytometric data are customarily displayed as histograms, or plots, of the digitized data. With digitized data, the light scatter and/or fluorescence signals emanating from individual cells are assigned to a channel proportional to the strength of the signal. The number of channels varies from instrument to instrument

but is normally 64, 256, or 1024. The stronger the signal from the cell (e.g., the more fluorescent the cell), the higher the channel number to which that cell is assigned. Various numbers of parameters may be displayed on a histogram, the most basic being a one-parameter display. Here the histogram consists of the intensity of signal (channel number) on the x-axis (on either a log or linear scale) versus the number of particles (cells) in each channel on the y-axis. Most contemporary flow cytometry examines more than one parameter of a cell, and therefore one parameter histograms are rarely used. Two-parameter histograms display the intensity of two separate signals on the x- and y-axes, with the number of events being displayed either by dot density (Fig. 62.4) or contours or in isometric form. Three-parameter histograms are occasionally used to display multiparameter data and rely on three-dimensional (cubic) displays. The display of the number of events often requires color density graphics; therefore, the use of three-parameter histograms has been somewhat limited. Probably the most common display used in the clinical laboratory is the two-parameter histogram dot-plot. Here correlated data from individual cells may be visualized in a readily understandable manner. For three- and four-color (or more) studies, data are displayed in multiple two-parameter histograms.

Virtually all cells will display some degree of autofluorescence if stimulated with sufficient light. This would result in nonstained cells being placed at various positions in the histogram displays. This usually is not desirable because fluorescence intensity is typically displayed on a log scale and the ability to delineate small shifts in fluorescence decreases on the upper portion of the scale. Therefore, using an unstained control, the laser power and/or voltages to the photomultiplier tubes are adjusted to bring the unstained samples near the origin of the histogram display. This defines where “negative” cells will appear on future histograms.

Historically, the interpretation of “positive” staining relied on the “background” staining observed with isotypic controls (monoclonal antibodies of the same isotype as the test antibody but directed against irrelevant antigens) (151, 185). The isotype control defined negative staining. Recently, the use of isotype controls has been proposed to be superfluous, and potentially misleading at times (163, 186,187 and 188). Isotype controls ultimately have use in certain in clinical situations, however total, blind reliance on isotype controls is not appropriate (189).

The interpretation of immunophenotyping in the clinical laboratory is highly reliant on the immunophenotyping of normal control specimens. This serves two purposes: first, immunophenotyping normal controls on a daily basis, or with each assay, ensures that the assay is being properly conducted and the reagents are performing as expected; second, maintaining a historical database of normal control values allows better determination if a specific patient's phenotype falls within the range considered to be normal. Specimens used as normal controls should be as similar to the patient specimen as possible. For example, both should be collected in the same anticoagulant, if possible, and both should be prepared in the same manner (density gradient or whole blood lysis). Age, race, sex, and time of collection have all been shown to influence immunophenotyping, although the variations are generally regarded as insufficient to warrant separate controls.

The results of immunophenotyping assays are often reported in terms of percentages of cells (i.e., lymphocytes) positive for a specific marker. While this is sufficient in most instances, there are disease states that require the determination of absolute numbers of a given cell population, particularly if the patient is lymphopenic. In AIDS, for example, the absolute number of CD4+ cells is usually desired (155). The absolute number is readily obtained by multiplying the percentage of positive cells (for CD4 in this instance) by the absolute lymphocyte concentration. Alternatively, absolute numbers of CD4+ cells can be determined using bead standards (65). In the immunophenotyping of leukemia and lymphomas, reporting the percentage of cells positive for a given marker is considered of less value than describing the markers present on an abnormal population of cells (188, 190).

Quantitative Flow Cytometric Immunophenotyping

Flow cytometry is often used to determine if cells express a particular antigen, however, it also may be used to quantitate the amount of fluorescence resulting from antibody binding to a cell (191,192,193 and 194). As with manual immunofluorescence studies, staining evaluated by flow cytometry often has been reported as “bright” or “dim.” Through the use of a variety of standard beads, flow cytometric immunophenotyping may be reported in quantitative units such as molecules of equivalent soluble fluorochrome MESF) or antibody-binding capacity (ABC) (191,192 and 193). This quantitative approach to reporting flow cytometric immunophenotyping data offers a much more rigorous interlaboratory standardization for interpreting intensity of marker staining, as intensity of surface marker staining is often of potential value in the study of hematologic neoplasia and in HIV disease (163, 195, 196). A recent survey indicated that, in the analysis of leukemias and lymphomas, less than 4% of respondents used MESF units, although more than 70% routinely commented on intensity of marker staining (54).

DNA Analysis

One of the major reasons for undertaking DNA studies is to determine the ploidy of a specimen. Flow cytometry is a convenient method for this purpose, given the limitations of sensitivity discussed previously, and the interpretation of these data generally is straightforward (for more detail see 197, 198). The inclusion of a known diploid control, such as normal peripheral blood, is crucial to this interpretation. This control may be run in a separate tube or admixed with the patient specimen as an internal control. Preferably, the analysis is performed with the following array of samples: (a) diploid control, (b) patient specimen, and (c) diploid control and patient specimen. The third sample reduces the possibility that any shift in the G0/G1 peak from control to patient is because of fluctuations in the instrument. It remains desirable, however, to retain a separate diploid control. This serves to confirm that a single peak in a patient specimen reflects a diploid population rather than a technical error (e.g., failure to mix the control and specimen at the proper ratio). In simplest terms, the ploidy, or DNA index (DI), is determined by dividing the peak channel number of the G0/G1 peak of the specimen by the peak channel number of the G0/G1 peak of the diploid control. A

diploid specimen will have a DI equal to 1.00. A hyperdiploid specimen will have a DI greater than 1.00, and a hypodiploid specimen will have a DI less than 1.00.

The interpretation of DNA histograms is occasionally more complex than the proceeding would suggest. This primarily relates to the CV of the G0/G1 peaks. As mentioned previously, the lower the CV, the better and more precise the detection of aneuploidy. A specimen that is only slightly aneuploid may not yield a distinct G0/G1 peak, even with low CVs. Rather, one would observe a “shoulder” on the diploid G0/G1 peak because of the slightly aneuploid population. Discriminating a true an-euploid shoulder from an uneven diploid G0/G1 peak may be quite difficult. Experience, the “three tube” approach to analysis, and repeated analysis of the specimen in question are the means of addressing this question. Caution also must be taken in the interpretation of the case in which no distinct aneuploid peak appears in a patient specimen. This does not necessarily imply a true diploid genome; the deletion of some chromosomes, accompanied by the overreplication of others, could yield a net DNA content resembling normal diploidy. Therefore, the failure to detect an aneuploid G0/G1 peak should not be reported as diploidy, but rather as “no aneuploid peak detected.”

The interpretation of tetraploidy also must be approached cautiously because this peak will fall in the same channels as the normal G2/M peak and as doublets of diploid G0/G1 cells. The latter artifacts (doublets) may be detected using pulse processing, a technique by which the shape of the pulse signal may be used to delineate tetraploid cells from doublets. In lieu of pulse processing, or in addition to it, the manual microscopic examination of specimens concomitant to their analysis by flow cytometry will reveal whether or not doublets are present in the specimen. Tetraploidy should never be reported without at least one of these quality assurances. In specimens in which doublets are not present, a question may remain whether a tetraploid peak reflects a true aneuploid population or the G2/M population of the diploid cells. In some cases, this may never be ultimately resolved, although the following guidelines are helpful: (i) in most instances, an increase will be observed in S-phase if the tetraploid cells are G2/M cells; and (ii) if the tetraploid cells represent a true aneuploid population, the presence of an octoploid (8N) population might be expected, which would represent the G2/M population of the tetraploid cells.

A second reason for undertaking studies of cellular DNA is to assess proliferation. Here the DNA histograms must be analyzed in a manner that most accurately determines the percentage of cells in each phase (G0/G1, S, and G2/M) of the cell cycle. In the commonly used one-parameter DNA analysis, it is not possible to distinguish between G0 and G1 or G2 and M; therefore, these are considered collectively as G0/G1 and G2/M, respectively. For the comparison of proliferation rates among cell populations, the S-phase or combined S plus G2/M phases are computed. Because these phases represent periods of DNA synthesis, the higher the percentage of cells in these phases, the more rapidly a population is proliferating.

S-phase is best determined in cultured cells by the incorporation of bromodeoxyuridine, followed by two-color flow cytometric analysis (199). Most clinical specimens are not cultured *in vitro*, thus, clinical specimens usually are stained only with nucleic acid dyes such as propidium iodide. A number of computer programs have been written to determine the percentage of cells in G0/G1, S, and G2/M in one-color DNA analysis, and have been elegantly described elsewhere (200,201,202 and 203). In general, the more normal appearing the DNA histograms and the tighter the CVs, the more accurate are the data derived from these programs. The determination of cell cycle compartments may be complicated by the presence of debris and multiploid specimens. In these instances, cell cycle compartments, particularly the S-phase, may be contaminated by extraneous events. The debris or cells from other clones in a multiploid specimen may erroneously become a component of the cell cycle determination. New software that considers debris and aneuploid G0/G1 peaks in the estimation of cellular proliferation are available and are particularly helpful in the analysis of solid tumors (201, 202).

Data Collection

There are several methods to collect and store data from flow cytometers. Data may be collected and stored as histogram representations of the specimen analysis, or they may be collected and stored in “list mode” fashion (204). The essence of list mode data collection and storage is the fact that each piece of information regarding each cell is stored and may be recalled and reanalyzed in a large number of methods. By contrast, data stored in a histogram form may only be reanalyzed in the same manner in which it was stored. Storage of list mode data versus histogram storage is somewhat comparable to a histopathologist's storage of tissue embedded in paraffin blocks rather than a single stained tissue section. The list mode data can be used to generate a histogram, as a paraffin block may be used to generate a stained tissue section. Yet, list mode data, like paraffin blocks, contain information not necessarily evident in the processed histogram or slide. For example, list mode storage allows regating of the data, whereas histogram storage would not. With the increasing complexity of flow cytometric analyses and the availability of large, low-cost data storage devices on computers, list mode data collection has become the standard in the clinical laboratory.

Cytometer Reliability and Performance

Flow cytometers are complex arrays of optic, pneumatic, hydraulic, and computer systems integrated in such a way as to allow interrogation of multiple cellular features. Failures in any of these systems can lead to an overall malfunction of the instrument. Total failure is less of a concern than a partial failure in which the instrument appears to remain functional, yet the resulting data are flawed. For example, a partially obscured orifice will alter light scatter patterns, and fluctuations in either the laser output or the sample pressure will alter apparent fluorescence intensity of a specimen. Experienced cytometer operators will generally detect abrupt shifts in these patterns and cease performing clinical analyses until the problem is identified and corrected. Subtle changes, which occur over a long period of time, are more difficult to detect; however, they may equally affect the quality of the flow cytometric data. Requirements to assure proper instrument performance and maintenance are outlined by Clinical Laboratory Improvement Amendments (CLIA) and are enforced

by routine inspection of flow cytometry laboratories by an agency endorsed by CLIA (205). The daily alignment (if required) and calibration of the cytometer, together with the analysis of normal controls on a daily or per-run basis, serve as methods to detect gradual deterioration in instrument performance (151, 155, 185, 206). An additional and most important quality assurance measure is the participation in interlaboratory proficiency testing programs (207,208 and 209). In these programs (e.g., College of American Pathologists), aliquots of specimens are routinely distributed among participating laboratories for analysis. The data from the different institutions are compared, and institutions reporting data significantly different from the mean are recommended to check their instrumentation and procedures. Clearly, these programs are necessary to ensure a certain level of consistency in data analysis and interpretation among clinical flow cytometry laboratories.

Standardization and Consensus

With the firm entrenchment of flow cytometry on the clinical laboratory, much effort has been directed in recent years toward standardizing the practice of clinical flow cytometry. One component of standardization comes from the regulatory requirements discussed above. As in all other areas of clinical pathology, interlaboratory standardization of flow cytometric techniques and interpretation will benefit the clinician and patient alike. To achieve standardization of practice, a number of consensus conferences have been held on a range of pertinent topics. These topics include DNA analysis (52, 145,146,147,148 and 149, 210), immunophenotyping of hematologic neoplasms (48, 163, 188, 190, 211), and stem cell enumeration (100, 101). Additionally, the National Committee for Clinical Laboratory Standards (NCCLS) has put forward guidelines for immunophenotyping of peripheral blood lymphocytes (151), immunophenotyping of leukemia and lymphoma (212), flow cytometric reticulocyte enumeration (213), as well as numerous general laboratory procedures. Guidelines for performing CD4+ T-cell enumeration in HIV disease have been published by the Centers for Disease Control (155). These published guidelines and consensus recommendations are crucial for standardization of flow cytometry practices and should be strongly considered when implementing and performing pertinent procedures.

SUMMARY

Part of "62 - Flow Cytometry"

Flow cytometric assays are now commonplace in the clinical laboratory. The techniques used for these assays are highly varied, and it is increasingly difficult for any one laboratory to maintain expertise in all applications. The field of clinical flow cytometry continues to be a dynamic realm. Improved, yet easier-to-use instrumentation is becoming increasingly available; new reagents are making it possible to examine more cellular features simultaneously on a given cell; and the benefits of research to date provide an impetus for further study and the implementation of new clinical assays. Perhaps the most exciting prospect lies in the merger of flow cytometric analysis with molecular genetics to provide a study of disease processes of a cellular basis, from the level of the gene through the surface proteins. Clearly, flow cytometry does not, and will not, stand as a separate diagnostic procedure. Rather, it is one of many approaches used to ascertain the status of a patient's health.

REFERENCES

1. Shapiro HM. Practical Flow Cytometry, 3rd ed., New York, Wiley Liss, 1994.

2. Melamed M, Lindmo T, and Mendelsohn ML. *Flow Cytometry and Sorting*, 2nd ed., New York: Wiley-Liss, 1994.

3. Keren DF, McCoy JP, Carey JL. *Flow Cytometry in Clinical Diagnosis*, 3rd ed., Chicago: ASCP Press, 2000.

4. Owens MA and Loken MR. *Flow Cytometry: Principles for Clinical Laboratory Practice* New York: Wiley-Liss, 1995.

5. Riley RS, Mahin EJ, and Ross W. *Clinical Applications of Flow Cytometry*. New York: Igaku-Shoin Press, 1993.

6. Macey MG. *Flow Cytometry: Clinical Applications*. Oxford: Blackwell Scientific Publications, 1994.

7. Kishimoto T, Kikutani, H, Borne A, et al. *Leucocyte Typing VI*. Levittown, PA: Garland Publishing, 1997.

8. Jennings CD and Foon KA. Recent advances in flow cytometry: application to the diagnosis of hematologic malignancy. *Blood* 1997;90:2863–2892.

9. Freedman AS. Cell surface antigens in leukemias and lymphomas. *Cancer Invest* 1996;14:252–276.

10. Orfao A, Schmitz G, Brando B, et al. Clinically useful information provided by the flow cytometric immunophenotyping of hematological malignancies: current status and future directions. *Clinical Chemistry* 1999;45:1708.

11. Krause JR, Penchansky L, Contis L, et al. Flow cytometry in the diagnosis of acute leukemia. *Am J Clin Pathol* 1989;89:314–346.

12. Bradstock KF. The diagnostic and prognostic value of immunophenotyping in acute leukemia. *Pathology* 1993;25:367–374.

13. Sun T. *Color Atlas and Text of Flow Cytometric Analysis of Hematologic Neoplasms*. New York: Igaku-Shoin, 1993.

14. Khalidi HS, Medeiros LJ, Chang KL, et al. The immunophenotype of adult acute myeloid leukemia: high frequency of lymphoid antigen expression and comparison of immunophenotype, French-American-British classification, and karyotypic abnormalities. *Am J Clin Path* 1998;109:211–220.

15. Uckun FM, Sather HN, Gaynon PS, et al. Clinical features and treatment outcome of children with myeloid antigen positive acute lymphoblastic leukemia: a report from the Children's Cancer Group. *Blood* 1997;90:28–35.

16. Drexler HG, Thiel E, Ludwig WD. Acute myeloid leukemias expressing lymphoid-associated antigens: diagnostic incidence and prognostic significance. *Leukemia* 1993;7:489–498.

17. Launder TM, Bray RA, Stempora L, et al. Lymphoid-associated antigen expression by acute myeloid leukemia. *Am J Clin Path* 1996;106:185–191.

18. Hanson CA, Abaza M, Sheldon S, et al. Acute biphenotypic leukemia: immunophenotypic and cytogenetic analysis. *Br J Hematol* 1993;84:49.

19. Matutes E, Morilla R, Farahat N, et al. Definition of acute biphenotypic leukemia. *Haematologica* 1997;82:64–66.

20. McCoy JP, Johnson E, Catalano E, et al. Concomitant myeloproliferative disorder and chronic lymphocytic leukemia (CLL): detection and monitoring by flow cytometry. *Arch Pathol Lab Med* 1995;119:1038–1043.

21. Mitterbauer G, Schwarzmeier J, Mitterbauer M, et al. Myelodysplastic syndrome/acute myeloid leukemia supervening previously untreated chronic B-lymphocytic leukemia: demonstration of the concomitant presence of two different malignant clones by immunologic and molecular analysis. *Annals of Hematology* 1997;74:193–197.

22. Novak PM, Mattson JC, Crisan D, et al. Separate clones in concomitant multiple myeloma and a second B-cell neoplasm demonstrated by molecular and immunophenotypic analysis. *Eur J Haematol* 1995;54:254–261.

23. Esteve J, Cervantes F, Rives S, et al. Simultaneous occurrence of B-cell chronic lymphocytic leukemia and chronic myeloid leukemia with further evolution to lymphoid blast crisis. *Haematologica* 1997;82:596–599.

24. Groeneveld K, te Marvelde JG, van den Beemd MW, et al. Flow cytometric detection of intracellular antigens for immunophenotyping of normal and malignant leukocytes. *Leukemia* 1996;10:1383–1389.

25. Nakase K, Sartor M, Bradstock. Detection of myeloperoxidase by flow cytometry in acute leukemia. *Cytometry* 1998;34:198–202.

26. Nowak R, Oelschlagel U, Hofmann R, et al. Detection of aneuploid cells in acute lymphoblastic leukemia with flow cytometry before and after therapy. *Leukemia Res* 1994;18:897–901.

27. Barlogie B, Stass S, Dixon D, et al. DNA aneuploidy in adult acute leukemia. *Cancer Genet Cytogenet* 1987;28:213–228.

28. Look AT, Roberson PK, Williams DK, et al. Prognostic importance of blast cell DNA content in childhood acute lymphoblastic leukemia. *Blood* 1985;65:1079–1086.

29. Sievers EL, Loken MR. Detection of minimal residual disease in acute myelogenous leukemia. *J Pediatr Hematol Oncol* 1995;17:123–133.

30. Pui CH, Raimondi SC, Head DR, et al. Characterization of childhood acute leukemia with multiple myeloid and lymphoid markers at diagnosis and at replase. *Blood* 1991;78:1327–1337.

31. Macedo A, Orfao A, Gonzalez M, et al. Immunological detection of blast cell subpopulations in acute myeloblastic leukemia at diagnosis: Implications for minimal residual disease studies. *Leukemia* 1995;9:993–998.

32. Stelzer GT, Shults KE, Wormsley SB, et al. Detection of occult lymphoma cells in bone marrow aspirates by multidimensional flow cytometry. *Prog Clin Biol Res* 1992;377:629–635.

33. Stelzer GT, Shults KE, and Loken MR. CD45 gating for routine flow cytometric analysis of human bone marrow specimens. *Ann NY Acad Sci* 1993;677:265–280.

34. Foon KA. Chronic lymphoid leukemias: recent advances in biology and therapy. *Stem Cells* 1995;13:1–21.

35. DiGiuseppe JA, Borowitz MJ. Clinical utility of flow cytometry in the chronic lymphoid leukemias. *Semin Oncology* 1998;25:6–10.

36. Kilo MN, Dorfman DM. The utility of flow cytometric immunophenotypic analysis in the distinction of small lymphocytic lymphoma/chronic lymphocytic leukemia from mantle cell lymphoma. *Am J Clin Pathol* 1996;105:451–457.

37. Dorfman DM, Pinkus GS. Distinction between small lymphocytic and mantle cell lymphoma by immunoreactivity for CD23. *Modern Pathology* 1994;7:326–331.

38. Huang JC, Finn WG, Goolsby CL, et al. CD5 — small B-cell leukemias are rarely classifiable as chronic lymphocytic leukemia. *Am J Clin Path* 1999;111:123–130.

39. Cabezudo E, Carrara P, Morilla R, et al. Quantitative analysis of CD79b, CD5 and CD19 in mature B-cell lymphoproliferative disorders. *Haematologica* 1999;84:413–418.

40. Oertel J, Huhn D. CD34 immunophenotyping of blasts in myelodysplasia. *Leukemia Lymphoma* 1994;15:65–69.

41. Jensen IM. Myelopoiesis in myelodysplasia evaluated by multiparameter flow cytometry. *Leukemia Lymphoma* 1995;20:17–25.

42. Ch'ang HJ, Tien HF, Wang CH, et al. Comparison of clinical and biologic features between myeloid and lymphoid transformation of Philadelphia chromosome positive chronic myeloid leukemia. *Cancer Genetics & Cytogenetics* 1993;71:87–93.

43. Cervantes F, Villamor N, Esteve J, et al. 'Lymphoid' blast crisis of chronic myeloid leukaemia is associated with distinct clinicohaematological features. *British Journal of Haematology* 1998;100:123–128.

44. Tbakhi A, Edinger M, Myles J, et al. Flow cytometric immunophenotyping of non-Hodgkin's lymphomas and related disorders. *Cytometry* 1996;25:113–124.

45. Zuckerberg LR, Medeiros JL, Ferry JA, et al. Diffuse low-grade lymphomas: four clinically distinct subtypes defined by a combination of morphologic and immunophenotypic features. *Am J Clin Pathol* 1993;100:373–385.

46. Jaffe ES. The role of immunophenotypic markers in the classification of non-Hodgkin's lymphomas. *Semin Oncol* 1990;17:11–19.

47. Weisenburger DD and Chan WC. Lymphomas of follicles: mantle cell and follicle center cell lymphomas. *Am J Clin Pathol* 1993;99:409–420.

48. Davis BH, Foucar K, Szczarkowski W, et al. US-Canadian Consensus Conference on the immunophenotypic analysis of hematologic neoplasia by flow cytometry: medical indications. *Cytometry* 1997;30:249–263.

49. Charley M, McCoy JP, Deng JS, et al. Anti-V region antibodies as "almost clonotypic" reagents for the study of cutaneous T cell lymphomas and leukemias. *J Invest Dermatol* 1990;95:614–617.

50. Merkel DE, Dressler LG, McGuire WL. Flow cytometry, cellular DNA content, and prognosis in human malignancy. *J Clin Oncol* 1987;5:1690–1703.

51. Bauer KD, Merkel DE, Winter JN, et al. Prognostic implications of ploidy and proliferative activity in diffuse large cell lymphomas. *Cancer Res* 1986;46:3173–3178.

52. Duque RE, Andreef M, Braylan RC, et al. Consensus review of the clinical utility of DNA flow cytometry in neoplastic hematopathology. *Cytometry* 1993;14:492–496.

53. McCoy JP, Overton WR. A survey of current practices in clinical flow cytometry. *Am J Clin Pathol* 1996;106:82–86.

54. McCoy JP and Keren DF. Current practices in clinical flow cytometry: a practice survey by the American Society of Clinical Pathologists. *Am J Clin Pathol* 1999;111:161–168.

55. Dalgleish AG, Beverly PCL, Clapham PR, et al. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 1984;312:763–767.

56. Simonsen JN, Fowke KR, MacDonald KS, et al. HIV pathogenesis: mechanisms of susceptibility and disease progression. *Current Opinion in Microbiology* 1998;1:423–429.

57. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann Intern Med* 1997;126:946–954.

58. Taylor JMG, Fahey JL, Detels R, et al. CD4 percentage, CD4 number, and CD4:CD8 ratio in HIV infection: which to choose and how to use. *J AIDS* 1989;2:114–124.

59. Giorgi JV, Detels R. T-cell subset alterations in HIV-infected homosexual men: NIAID multicenter AIDS cohort study. *Clin Immunol Immunopathol* 1989;52:10–18.

60. Horowitz HW, Telzak EE, Sepkowitz KA, et al. Human immunodeficiency virus infection, part I. *Disease-A-Month* 1998;44:545–606.

61. British HIV Association guidelines for antiretroviral treatment of HIV seropositive individuals. BHIVA Guidelines Co-ordinating Committee. *Lancet* 1997;349:1086–1092.

62. Riley JL, Levine BL, Craighead N, et al. Naive and memory CD4 T cells differ in their susceptibilities to human immunodeficiency virus type 1 infection following CD28 costimulation: implications for transmission and pathogenesis. *J Virol* 1998;72:8273–8280.

63. Giorgi JV. Hausner MA. Hultin LE: Detailed immunophenotype of CD8+ memory cytotoxic T-lymphocytes (CTL) against HIV-1 with respect to expression of CD45RA/RO, CD62L and CD28 antigens. Immunology Letters. 66(1–3):105–10, 1999.

64. Hengel RL, Jones BM, Kennedy MS, Hubbard MR, McDougal JS, Markers of lymphocyte homing distinguish CD4 T cell subsets that turn over in response to HIV-1 infection in humans. *Journal of Immunology* 163(6):3539-48, 1999.

65. Nicholson JKA, Stein D, Mui T, Mack R, et al. Evaluation of a method for counting absolute numbers of cells with a flow cytometer. *Clin Diagn Lab Immunol* 1997;4:309–313.

66. O'Gorman MR, Gelman R. Inter- and intrainstitutional evaluation of automated volumetric capillary cytometry for the quantitation of CD4- and CD8-positive T lymphocytes in the peripheral blood of persons infected with human immunodeficiency virus. Site Investigators and the NIAID New CD4 Technologies Focus Group. *Clin Diagn Lab Immunol* 1997;4:173–179.

67. Johnson D, Hirschkorn D, Busch MP, et al. Evaluation of four alternative methodologies for determination of absolute CD4+ lymphocyte counts. *J Acq Immun Def Syndr Human Retrovirol* 1995;10:522–530.

68. Fleisher TA, Tomar RH. Introduction to diagnostic laboratory immunology. *JAMA* 1997;278:1823–1834.

69. Nicholson JKA. Use of flow cytometry in the evaluation and diagnosis of primary and secondary immunodeficiency disease. *Arch Pathol Lab Med* 1989;113:598–605.

70. Gupta S. T cell subpopulations defined with monoclonal antibodies in patients with primary immunodeficiencies. *Immunol Lett* 1982;4:129–133.

71. White WB, Ballow M. Modulation of suppressor cell activity by cimetidine in patients with common variable hypogammaglobulinemia. *N Engl J Med* 1985;312:198–202.

72. Sirianni MC, Buscinco RL, Fiore L, et al. T cell subsets and natural killer cells in DiGeorge and SCID patients. *Birth Defects* 1982;19:107–101.

73. Harlan JM. Leukocyte adhesion deficiency syndrome: insights into the molecular basis of leukocyte emigration. *Clinical Immunology & Immunopathology* 1993;67(3 Pt 2):S16–24.

74. Phillips ML, Schwartz BR, Etzioni A, et al. Neutrophil adhesion in leukocyte adhesion deficiency syndrome type 2. *J of Clin Invest* 1995;96:2898–906.

75. Maino VC, Suni MA, Ruitenberg JJ. Rapid flow cytometric method for measuring lymphocyte subset activation. *Cytometry* 1995;20:127–133.

76. Mardiney M 3rd, Brown MR, Fleisher TA. Measurement of T-cell CD69 expression: a rapid and efficient means to assess mitogen- or antigen-induced proliferative capacity in normals. *Cytometry* 1996;26:305–310.

77. Chen JC, Davis BH, Leon MA, et al. Gamma radiation induces CD69 expression on lymphocytes. *Cytometry* 1997;30:304–312.

78. Bass DA, Parce JW, DeChatelet LR, et al. Flow cytometric studies of oxidative product formation by neutrophils. *J Immunol* 1983;130:190–195.

79. Crockard AD, Thompson JM, Boyd NA, et al. Diagnosis and carrier detection of chronic granulomatous disease in five families by flow cytometry. *International Archives of Allergy & Immunology* 1997;114:144–152.

80. Perticarari S, Presani G, Banfi E. A new flow cytometric assay for the evaluation of phagocytosis and the oxidative burst in whole blood. *J Immunol Meth* 1994;170:117–124.

81. Bray RA. Flow cytometry crossmatching for solid organ transplantation. *Methods in Cell Biology* 1994;41:103–109.

82. Bray RA. Flow cytometry in the transplant laboratory. *Annals of the New York Academy of Sciences* 1993;677:138–151.

83. Reed MH, Shapiro ME, Milford EL, et al. Interleukin 2 receptor expression of peripheral blood lymphocytes in association with renal allograft rejection. *Transplantation* 1989;48:361–366.

84. Siegel DL, Fox I, Dafoe CD, et al. Discriminating rejection from CMV infection in renal allograft recipients using flow cytometry. *Clin Immounol Immunopathol* 1989;51:157–171.

85. van Es A, Baldwin WM, Oljans PJ, et al. Expression of HLA-DR on T-lymphocytes following renal transplantation, and association with graft-rejection episodes and cytolomegalovirus infection. *Transplantation* 1984;37:65–69.

86. Stoolman LM. Lymphocyte subset monitoring after organ transplantation, In: Keren DF, ed. *Flow Cytometry in Clinical Diagnosis*. Chicago: ASCP Press, 1989:248–266.

87. Ortho Multicenter Transplant Study Group. A randomized clinical trial of OKT3 monoclonal antibody for acute rejection of cadaveric renal transplants. *N Engl J Med* 1985;313:337–342.

88. Ziabinger GJ, Ulrich W, Sonnleitner A, et al. Determination of serum OKT3 levels with flow cytometry. *Transplant Proc* 1989;21:989–991.

89. Hammond EA, Yowell RL, Greenwood J, et al. Prevention of adverse clinical outcome by monitoring of cardiac transplant patients for murine monoclonal CD3 antibody (OKT3) sensitization. *Transplantation* 1993;55:1061–1063.

90. Lim VL, Gumbert M, Garovoy MR. A flow cytometric method for the detection of the development of antibody to orthoclone OKT3. *J Immunol Methods* 1989;121:197–201.

91. Hesse CJ, Jutte NH. Detection of anti-idiotypic antibodies after OKT3 treatment by flow cytometry. *Transplantation* 1996;61:169.

92. McIntyre JA, Kincade M, Higgins NG. Detection of IGA anti-OKT3 antibodies in OKT3-treated transplant recipients. *Transplantation* 1996;61:1465-1469.

93. van den Berg AP, Twilhaar WN, van Son WJ, et al. Quantification of immunosuppression by flow cytometric measurement of intracellular cytokine synthesis. *Transplant International* 1998;1(suppl 11):S318–S321.

94. Rostaing L, Puyoo O, Tkaczuk J, et al. Differences in type 1 and type 2 intracytoplasmic cytokines, detected by flow cytometry, according to immunosuppression (cyclosporine A vs. tacrolimus) in stable renal allograft recipients. *Clinical Transplantation* 1999;13:400–409.

95. Popplewell L, Forman SJ. Allogeneic hematopoietic stem cell transplantation for acute leukemia, chronic leukemia, and myelodysplasia. Hematology—Oncology *Clinics of North America* 1999;13:987–1015, vi-vii.

96. Nieto Y, Shpall EJ. Autologous stem-cell transplantation for solid tumors in adults. Hematology — Oncology Clinics of North America. 1999;13:939–968, vi.

97. Passweg J, Gratwohl A, Tyndall A. Hematopoietic stem cell transplantation for autoimmune disorders. *Current Opinion in Hematology* 1999;6:400–405.

98. Siena S, Bregni M, Brando B, et al. Flow cytometry for clinical estimation of circulating hematopoietic progenitors for autologous transplantation in cancer patients. *Blood* 1991;77:400–409.

99. Gratama JW, Orfao A, Barnett D, et al. Flow cytometric enumeration of CD34+ hematopoietic stem and progenitor cells. *Cytometry* 1998;34:128–142.

100. Sutherland DR, Anderson L, Keeney M, et al. The ISHAGE guidelines for CD34+ cell determination by flow cytometry. *J Hematother* 1996;5:213–226.

101. Johnsen HE, Knudsen LM. Nordic flow cytometry standards for CD34+ cell enumeration in blood and leukapheresis products: report from the second nordic workshop. *J Hematother* 1996;5:237–245.

102. Sutherland DR and Keating A. The CD34 Antigen: Structure, Biology, and Potential Clinical Applications. *J Hematother* 1:115-129, 1992.

103. Gratama JW, Orfao A, Barnett D, et al. Flow cytometric enumeration of CD34+ hematopoietic stem and progenitor cells. *Cytometry* 1998;34:128–142.

104. Chen CH, Lin W, Shye S, et al. Automated enumeration of CD34+ cells in peripheral blood and bone marrow. *J Hematother* 1994;3:3-13.

105. Keeney M, Chin-Yee I, Weir K, et al. Single platform flow cytometric absolute CD34+ cell counts based on the ISHAGE guidelines. *Cytometry* 1998;34:61–70.

106. Davis BH, Bigelow N, Ball ED, et al. Utility of flow cytometric reticulocyte quantification as a predictor of engraftment in autologous bone marrow transplantation. *Am J Hemato* 1989;32:81–87.

107. Martin PJ, Hansen JA. Quantitative assays for detection of residual T cells of T-depleted human marrow. *Blood* 1985;65:1134–1140.

108. Lamb LS Jr, Abhyankar SA, Hazlett L, et al. Expression of CD134 (0X-40) on T cells during the first 100 days following allogeneic bone marrow transplantation as a marker for lymphocyte activation and therapy-resistant graft-versus-host disease. *Cytometry* 1999;38:238–243.

109. Yang YH, Hutchinson P, Littlejohn GO, et al. Flow cytometric detection of anti-neutrophil cytoplasmic autoantibodies. *J Immunol Meth* 1994;172:77–84.

110. Maher GM, Hartman KR. Detection of antineutrophil autoantibodies by flow cytometry: use of unfixed neutrophils as antigenic targets. *J Clin Lab Anal* 1993;7:334–340.

111. Hanson CA. Applications of flow cytometry in diagnostic pathology. *Am J Clin Pathol* 1989;91(suppl 1):S27-31.

112. Stockelberg D, Hou M, Jacobsson S, et al. Detection of platelet antibodies in chronic idiopathic thrombocytopenic purpura (ITP). A comparative study using flow cytometry, a whole platelet ELISA, and an antigen capture ELISA. *European Journal of Haematology* 1996;56:72–77.

113. Latorraca A, Lanza F, Moretti S, et al. Flow cytometric analysis of anti-platelet antibodies in idiopathic thrombocytopenic purpura. *Haematologica* 1994;79:269–272.

114. Freedman J, Hornstein A. Simple method for differentiating between HLA and platelet-specific antibodies by flow cytometry. *American Journal of Hematology* 1991;38:314–320.

115. Hall SE, Rosse WF. The use of monoclonal antibodies and flow cytometry in the diagnosis of paroxysmal nocturnal hemoglobinuria. *Blood* 1996;87:5332–5340.

116. Navenot JM, Bernard D, Harousseau JL, et al. Expression of glycosyl-phosphatidylinositol-linked glycoproteins in blood cells from paroxysmal nocturnal haemoglobinuria patients: a flow cytometry study using CD55, CD58 and CD59 monoclonal antibodies. *Leukemia & Lymphoma* 1996;21:143–51.

117. Lee LG, Chen CH, Chiu LA. Thiazole orange: a new dye for reticulocyte analysis. *Cytometry* 1986;7:508–517.

118. Davis BH, Bigelow NC. Flow cytometric reticulocyte quantification using thiazole orange provides clinically useful reticulocyte maturity index. *Arch Pathol Lab Med* 1989;113:684–689.

119. Rapi S, Ermini A, Bartolini L, et al. Reticulocytes and reticulated platelets: simultaneous measurement in whole blood by flow cytometry. *Clinical Chemistry & Laboratory Medicine* 1998;36:211–214.

120. Corberand JX. Reticulocyte analysis using flow cytometry. *Hematology & Cell Therapy* 1996;38:487–494.

121. George JN, Pickett EB, Saucerman S, et al. Platelet surface glycoproteins. Studies on resting and activated platelets and platelet membrane microparticles in normal subjects, and observations in patients during adult respiratory distress syndrome and cardiac surgery. *J Clin Invest* 1986;78:340–348.

122. Hakin RM, Schafer AI. Hemodialysis-associated platelet activation and thromocytopenia. *Am J Med* 1985;78:575–580.

123. Jy W, Mao WW, Horstman LL, et al. A flow cytometric assay of platelet activation marker P-selectin (CD62P) distinguishes heparin-induced thrombocytopenia (HIT) from HIT with thrombosis (HITT). *Thrombosis & Haemostasis* 1999;82:1255–1259.

124. Fitzgerald DJ, Roy L, Catella F, et al. Platelet activation in unstable coronary artery disease. *N Engl J Med* 1986;315:983–989.

125. Harker LA, Malpass TW, Branson HE, et al. Mechanism of abnormal bleeding in patients undergoing cardiopulmonary bypass. Acquired transient platelet dysfunction associated with selective alpha-granule release. *Blood* 1980;56:824–834.

126. Michelson AD. Flow cytometry: a clinical test of platelet function. *Blood* 1996;87:4925–4936.

127. McEver RP, Martin MN. A monoclonal antibody to a membrane glycoprotein binds only to activated platelets. *J Biol Chem* 1984;259:9799–9804.

128. Shattil SJ, Hoxie JA, Cunningham M, et al. Changes in the platelet membrane glycoprotein IIb-IIIa complex during platelet activation. *J Biol Chem* 1985;260:11107-11114.

129. Ault KA. Flow cytometric measurement of platelet function and reticulated platelets. *NY Acad Sci* 1993;677:293–308.

130. Shattil SJ, Cunningham M, Hoxie JA. Detection of activated platelets in whole blood using activation-dependent monoclonal antibodies and flow cytometry. *Blood* 1987;70:307–315.

131. Michelson AD, Furman MI. Laboratory markers of platelet activation and their clinical significance. *Current Opinion in Hematology* 1999;6:342–348.

132. Davis BH, Olsen S, Bigelow NC, et al. Detection of fetal red cells in fetomaternal hemorrhage using a fetal hemoglobin monoclonal antibody by flow cytometry. *Transfusion* 1998;38:749–756.

133. Navenot JM, Merghoub T, Ducrocq R, et al. New method for quantitative determination of fetal hemoglobin-containing red blood cells by flow cytometry: application to sickle-cell disease. *Cytometry* 1998;32:186–190.

134. Marcus SJ, Kinney TR, Shultz WH, et al. Quantitative analysis of erythrocytes containing fetal hemoglobin (F cells) in children with sickle cell disease. *Am J Hematol* 1997;54:40–46.

135. Campbell TA, Ware RE, Mason, M. Detection of hemoglobin variants in erythrocytes by flow cytometry. *Cytometry* 1999;35:242–248.

136. Reynolds WM, Evans PR, Wilson PJ, et al. Automated routine HLA-B27 typing by flow cytometry. *J Immunol Meth* 1996;197:1–5.

137. Chen JC, Davis BH, Bigelow NC, et al. Flow cytometric HLA-B27 typing using CD3 gating and molecules of equivalent soluble fluorochrome (MESF) quantitation. *Cytometry* 1996;26:286–292.

138. Coates E, Darke C. Routine HLA-B27 typing by flow cytometry: differentiation of the products of HLA-B2702, B2705 and B2708. *European Journal of Immunogenetics* 1998;25:29–37.

139. Hoffmann JJ, Janssen WC. HLA-B27 phenotyping with flow cytometry: further improvement by multiple monoclonal antibodies. *Clinical Chemistry* 1997;43:1975–1981.

140. Barlogie B, Drewinko B, Schumann J, et al. Cellular DNA content as a marker of neoplasia in man. *Am J Med* 1980;69:195–203.

141. Overton WR, Catalano E, McCoy JP. Method to make paraffin-embedded breast and lymph tissue mimic fresh tissue in DNA analysis. *Cytometry* 1996;26:166–171.

142. Joensuu H, Alanen KA, Klemi PJ, et al. Evidence for false aneuploid peaks in flow cytometric analysis of paraffin-embedded tissue. *Cytometry* 1990;11:431–437.

143. Price J, Herman, CJ. Reproducibility of FCM DNA content from replicate paraffin block samples. *Cytometry* 1990;11:845–847.

144. Eissa S, Khalifa A, Laban M, et al. Comparison of flow cytometric DNA content analysis in fresh and paraffin-embedded ovarian neoplasms: a prospective study. *British Journal of Cancer* 1998;77:421–425.

145. Duque RE, Andreeff M, Braylan RC, et al. Consensus review of the clinical utility of DNA flow cytometry in neoplastic hematopathology. *Cytometry* 1993;14:492–496.

146. Shankey TV, Kallioniemi OP, Koslowski JM, et al. Consensus review of the clinical utility of DNA content cytometry in prostate cancer. *Cytometry* 1993;14:497–500.

147. Bauer KD, Bagwell CB, Giaretti W, et al. Consensus review of the clinical utility of DNA flow cytometry in colorectal cancer. *Cytometry* 1993;14:486–491.

148. Hedley DW, Clark GM, Cornelisse CJ, et al. Consensus review of the clinical utility of DNA cytometry in carcinoma of the breast. Report of the DNA Cytometry Consensus Conference. *Cytometry* 1993;14:482–485.

149. Wheeless LL, Badalament RA, de Vere White RW, et al. Consensus review of the clinical utility of DNA cytometry in bladder cancer. Report of the DNA Cytometry Consensus Conference. *Cytometry* 1993;14:478–481.

150. Clinical practice Guidelines for the use of tumor markers in breast and colorectal cancer. *J Clin Oncol* 1996;14:2843–2877.

151. Clinical applications of flow cytometry: quality assurance and immunophenotyping of peripheral blood lymphocytes. National Committee for Clinical Laboratory Standards H42-A, 1993.

152. Loken MR, Brosnan JM, Bach BA, et al. Establishing optimal lymphocyte gates for immunophenotyping by flow cytometry. *Cytometry* 1990;11:453–459.

153. Stewart CC, Stewart SJ. Four color compensation. *Cytometry* 1999;38:161–175.

154. Kutok JL, Roma AO, Lemire SJ, et al. Four-color flow cytometric immunophenotypic determination of peripheral blood CD4+ T-lymphocyte counts: a comparison of validity and cost-effectiveness with a two-color method. *Am J Clin Pathol* 1998;110:465–470.

155. 1997 Revised Guidelines for Performing CD4+ T Cell Determinations in Persons Infected with Human Immunodeficiency Virus (HIV). MMWR 46:RR2, 1997.

156. Bertouch JV, Roberts-Thomson PJ, Bradley J. Diurnal variation of lymphocyte subsets identified by monoclonal antibodies. *Br Med J* 1983;286:1171–1173.

157. Thornthwaite JT, Rosenthal PK, Vazquez DA, et al. The effects of anticoagulant and temperature on the measurements of helper and suppressor cells. *Diagn Immunol* 1984;2:167–174.

158. Paxton H, Bendele T. Effect of time, temperature, and anticoagulant on flow cytometry and hematological values. *Ann New York Acad Sci* 1993;677:440–443.

159. Dzik WH, Neckers L. Lymphocyte subpopulations altered during blood storage. *N Engl J Med* 1983;309:435–436.

160. Weiblen BJ, Debell K, Valeri CR. Acquired immunodeficiency of blood stored overnight. *N Engl J Med* 1983;309:793.

161. DePoali P, Rentano M, Battisin S, et al. Enumeration of human lymphocyte subsets by monoclonal antibodies and flow cytometry: a comparative study using whole blood or mononuclear cells separated by density gradient centrifugation. *J Immunol Methods* 1984;72:349–353.

162. Slade HB, Greenwood HJ, Hudson JL, et al. Spurious lymphocyte phenotypes by flow cytometry from mononuclear cells prepared by ficoll-hypaque. *Pediatr Res* 1987;21:318A.

163. Stelzer GT, Marti G, Hurley A, et al. Consensus conference on the flow cytometric analysis of leukemia and lymphoma: report of the committee on the standardization and validation of laboratory procedures. *Cytometry* 1997;30:214–230.

164. Terstappen LW, Johnson D, Mickaels RA, et al. Multidimensional flow cytometric blood cell differentiation without erythrocyte lysis. *Blood Cells* 1991;17:585–602.

165. Lifson JD, Sasaki DT, Engleman EG. Utility of formaldehyde fixation for flow cytometry and inactivation of the AIDS associated retrovirus. *J Immunol Methods* 1986;86:143–149.

166. van Vugt MJ, van den Herik-Oudijk IE, van de Winkle JG. Binding of PE-CY5 conjugates to the human high-affinity receptor for IgG (CD64). *Blood* 1996;88:2358–2361.

167. Lanza F, Latorraca A, Moretti S, et al. Comparative analysis of different permeabilization methods for the flow cytometry measurement of cytoplasmic myeloperoxidase and lysozyme in normal and leukemic cells. *Cytometry* 1997;30:134–44.

168. Groeneveld K, te Marvelde JG, van den Beemd MW, et al. Flow cytometric detection of intracellular antigens for immunophenotyping of normal and malignant leukocytes. *Leukemia* 1996;10:1383–1389.

169. Farahat N, van der Plas D, Praxedes M, Morilla R, Matutes E, Catovsky D. Demonstration of cytoplasmic and nuclear antigens in acute leukaemia using flow cytometry. *J Clin Pathol* 1994;47:843-849.

170. Sartor M, Bradstock K. Detection of intracellular lymphoid differentiation antigens by flow cytometry in acute lymphoblastic leukemia. *Cytometry* 1994;18:119–122.

171. Cerra R, Zarbo RJ, Crissman JD. Dissociation of cells from solid tumors. *Methods in Cell Biol* 1990;33:1–12.

172. Hedley DW, Friedlander ML, Taylor IW, et al. Method for analysis of cellular DNA content of paraffin-embedded pathological material using flow cytometry. *J Histochem Cytochem* 1983;31:1333–1335.

173. Hedley DW. Flow cytometry using paraffin-embedded tissue: five years on. *Cytometry* 1989;10:229–241.

174. Herbert DJ, Nishiyama RH, Bagwell CB, et al. Effects of several commonly used fixatives on DNA and total nuclear protein analysis by flow cytometry. *Am J Clin Pathol* 1989;91:535–541.

175. Overton WR, McCoy JP. Reversing the effect of formalin on the binding of propidium iodide to DNA. *Cytometry* 1994;16:351-356.

176. Stephenson RA, Gay H, Fair WR, et al. Effect of section thickness on quality of flow cytometric DNA content determinations in paraffin-embedded tissues. *Cytometry* 1986;7:41–44.

177. Vindelov LL, Christensen IJ, Nissen NI. A detergent-trypsin method for the preparation of nuclei for flow cytometric DNA analysis. *Cytometry* 1983;3:323–327.

178. Tate EH, Wilder ME, Cram LS, et al. A method for staining 3T3 cell nuclei with propidium iodide in hypotonic solution. *Cytometry* 1983;4:211–215.

179. Braylan RC, Benson NA, Nourse V, et al. Correlated analysis of cellular DNA, membrane antigens and light scatter of human lymphoid cells. *Cytometry* 1982;2:337–343.

180. Holm M, Thomsen M, Hoyer M, et al. Optimization of a flow cytometric method for the simultaneous measurement of cell surface antigen, DNA content, and in vitro BrdUrd incorporation into normal and malignant hematopoietic cells. *Cytometry* 1998;32:28–36.

181. Toba K, Winton EF, Koike T, et al. Simultaneous three-color analysis of the surface phenotype and DNA-RNA quantitation using 7-amino-actinomycin D and pyronin Y. *J Immuno Meth* 1995;182:193–207.

182. McCoy JP, Krause JR, Hanley-Yanez K, et al. Three color analysis of surface marker, oncogene product, and DNA content in leukemia. *Cytometry* 1990;4(supp 1):29.

183. Shackney SE, McCoy JP, Burholt DR, et al. Sequential paraformaldehyde and methanol fixation for quantitative multiparameter cellular DNA measurement and immunofluorescence measurements of cell surface and intracellular proteins by flow cytometry (FCM). *Cytometry* 1990;4(supp 1):17.

184. Francis C, Connelly MC. Rapid single-step method for flow cytometric detection of surface and intracellular antigens using whole blood. *Cytometry* 1996;25(1):58–70.

185. McCoy JP, Carey JL, Krause JR. Quality control in flow cytometry for diagnostic pathology. I. Cell surface phenotyping and general laboratory procedures. *Am J Clin Pathol* 1990;93(suppl 1):S27-D37.

186. McCoy JP, Blumstein L, Donaldson MH, et al. Accuracy and cost-effectiveness of a one-tube, three-color method for obtaining absolute CD4 counts and CD4:CD8 ratios. *Am J Clin Path* 1994;101:279–282.

187. Keeney M, Gratama JW, Chin-Yee IH, et al. Isotype controls in the analysis of lymphocytes and CD34+ stem and progenitor cells by flow cytometry — time to let go! *Cytometry* 1998;34:280–283.

188. Borowitz MJ, Bray R, Gascoyne R, et al. US-Canadian Recommendations on the immunophenotypic analysis of hematologic neoplasia by flow cytometry: data analysis and interpretation. *Cytometry* 1997;30:236–244.

189. O'Gorman MR, Thomas J. Isotype controls—time to let go? *Cytometry* 1999;38:78–80.

190. Braylan RC, Atwater SK, Diamond L, et al. US-Canadian recommendations on the immunophenotypic analysis of hematologic neoplasia by flow cytometry: data reporting. *Cytometry* 1997;30:245–248.

191. Vogt RF, Marti GE, Schwartz A. Quantitative calibration of fluorescence intensity for clinical and research applications of immunophenotyping by flow cytometry. *Rev Biotechn Bioeng* 1994;1:147–175.

192. Henderson LO, Marti GE, Gaigalas A, et al. Terminology and nomenclature for standardization in quantitative fluorescence cytometry. *Cytometry* 1998;33:97–105.

193. Schwartz A, Marti GE, Poon R, et al. Standardizing flow cytometry: a classification system of fluorescence standards used for flow cytometry. *Cytometry* 1998;33:106–114.

194. Gratama JW, D'hautcourt JL, Mandy F, et al. Flow cytometric quantitation of immunofluorescence intensity: problems and perspectives. European Working Group on Clinical Cell Analysis. *Cytometry* 1998;33:166–178.

195. Hultin LE, Matud JL, Giorgi JV. Quantitation of CD38 activation antigen expression on CD8+ T cells in HIV-1 infection using CD4 expression on CD4+ T lymphocytes as a biological calibrator. *Cytometry* 1998;33:123–132.

196. Perfetto SP, Malone JD, Hawkes C, et al. CD38 expression on cryopreserved CD8+ T cells predicts HIV disease progression. *Cytometry* 1998;33:133–137.

197. Bauer KD. Quality control issues in DNA content flow cytometry. *Ann NY Acad Sci* 1993;677:59–77.

198. Zarbo RJ. Quality control issues and technical considerations in flow cytometric DNA and cell cycle analysis of solid tumors. In: Keren DF, Hanson CA, Hurtubise PE, eds. *Flow Cytometry and Clinical Diagnosis* Chicago: ASCP Press, 1994:425–469.

199. Gratner HG. Monoclonal antibody to 5-bromo and 5-iododeoxyuridine: a new reagent for the detection of DNA replication. *Science* 1982;218:474–475.

200. Bagwell CB. Clinical data analysis for flow cytometry. In: Keren DF, ed. *Flow Cytometry in Clinical Diagnosis* Chicago: ASCP Press, 1989:310–330.

201. Fox MH. Cell cycle models for analyzing flow cytometry DNA histograms. *Rev Biotechn Bioeng* 1994;1:111–145.

202. Kallioniemi OP, Visakorpi T, Holli K, et al. Automated peak detection and cell cycle analysis of flow cytometric DNA histograms. *Cytometry* 1994;16:250–255.

203. Dean PN. Methods of data analysis in flow cytometry. In: Van Dilla MA, Dean PN, Laerum OD, et al., eds. *Flow Cytometry: Instrumentation and Data Analysis*. New York: Academic Press, 1985:195–221.

204. Salzman GC, Wilkins SF, Whitfill JA. Modular computer programs for flow cytometry and sorting: the LACEL system. *Cytometry* 1981;1:325–336.

205. Code of Federal Regulations (10/99) Part 493 Laboratory Requirements.

206. Muirhead KA. Establishment of quality control procedures in clinical flow cytometry. *Ann NY Acad Sci* 1993;677:1–20.

207. Coon JS, Paxton H, Lucy L, et al. Interlaboratory variation in DNA flow cytometry. Results of the College of American Pathologists' Survey. *Archives of Pathology & Laboratory Medicine* 1994;118:681–685.

208. Rosner E, Willis V, Lambert R, et al. CD4+ T-cell testing practices as implications for training. *Cytometry* 1997;30:181–185.

209. Scornik JC, Bray RA, Pollack MS, et al. Multicenter evaluation of the flow cytometry T-cell crossmatch: results from the American Society of Histocompatibility and Immunogenetics-College of American Pathologists proficiency testing program. *Transplantation* 1997;63:1440–1445.

210. Hedley DW, Shankey VT, Wheeless, LL. DNA cytometry consensus conference. *Cytometry* 1993;14:471.

211. Braylan RC, Borowitz MJ, Davis BH, et al. US-Canadian consensus recommendations on the immunophenotypic analysis of hematologic neoplasia by flow cytometry. *Cytometry* 1997;30:213.

212. Clinical Applications of Flow Cytometry: Immunophenotyping of Leukemic Cells. National Committee for Clinical Laboratory Standards H43-A, 1993.

213. Methods for Reticulocyte Counting (Flow Cytometry and Supravital Dyes). National Committee for Clinical Laboratory Standards H44-A, 1993.

63
Cellular and Humoral Mediators of Inflammation

Karen James

- PROTECTIVE DETERRENTS TO INFLAMMATION
- NONSPECIFIC CELLULAR RESPONSE TO INFLAMMATION
- NONSPECIFIC HUMORAL RESPONSE TO INFLAMMATION

PROTECTIVE DETERRENTS TO INFLAMMATION

Part of "63 - Cellular and Humoral Mediators of Inflammation"

External Barriers

The skin and mucous membrane linings of the respiratory, urinary, and gastrointestinal tracts constitute the body's first line of defense as an intact barrier of epithelial cells. The epithelial barrier cells and the biochemical defense mechanisms in the secretions serve to protect from infections. Lysozyme (muramidase) decreases the viability of susceptible bacteria by cleaving the cell wall proteoglycan between the residues of N-acetylglucosamine and N-acetylmuramic acid. Stomach gastric acids inhibit infections with *Salmonella* spp., *Vibrio cholera*, and enveloped viruses. Secretory immunoglobulins can neutralize viruses and bacterial endotoxins and prevent bacterial attachment to mucosal cells.

Cell Adhesion Molecules (CAMs)

Cell Adhesion Molecules (CAMs) regulate cell-cell and cell-matrix adhesion, processes that are involved in host defense against infections, antigen recognition, and immune surveillance. There are at least six families of CAMs (1) (Fig. 63.1), based on their structural profiles or specific adhesion-promoting domains: (i) cadherins, (ii) immunoglobulinlike proteins, (iii) integrins, (iv) selectins, (v) cell surface mucins, and (vi) cell surface proteoglycans.

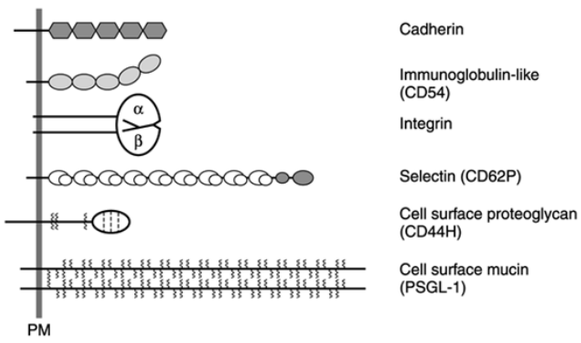


FIGURE 63.1. Schematic diagram of representative members of the cell adhesion molecule (CAM) families. The extracellular region of cadherins consists of five repeats. CD54 is shown here as a representative of the immunoglobulin (Ig)-like protein family. The characteristic structure of the Ig family is depicted here as ellipsoid. Integrins are drawn to accommodate the head and tail morphology revealed by electron microscopic studies. Human P-selectin consists of a globular lectin domain, an epidermal growth factor (EGF) domain, and nine short consensus repeats assembled as rod-shaped structure. Hema-topoietic CD44 has an N-terminal proteo-glycan tandem repeat stabilized by three pairs of cysteine residues, followed by a region containing multiple O-linked and chondroitin sulfate ||| attachment sites. P-selectin glycoprotein ligand (PSGL-1), a dimeric transmembrane glycoprotein with a large mucin ectodomain depicted as a rodlike structure, with arbitrary-based carbohydrate branches. Adapted with permission from Arnaout MA. Structural diversity of cell adhesion molecules and their role in inflammation. *Textbook of Rheumatology*, Kelley WN, Ruddy S, Harris ED, et al., eds., St. Louis, MO: W.B. Saunders, 1997:303–321.

Cadherins are crucial in embryonic stages of development to maintain initial intercellular adhesion and stability essential for morphogenesis. In adult tissues, cadherins maintain tight gap junctions and intercellular spacing. Cadherins bind homophilically to other cadherins in a temperature and calcium (Ca^{2+})-dependent binding reaction described as a “linear zipper” of cell adhesion.

The immunoglobulin (Ig) supergene family of CAMs, so called because all contain 60–100 amino-acid sequences common to immunoglobulins, mediate cell recognition events during morphogenesis, inflammation, homeostasis, and immunity. Many of the Ig supergene family proteins serve as receptors for parasitic or viral infections [e.g., CD4 is the major receptor for the human immunodeficiency virus (HIV)]. Another Ig member is $ß_2$-microglobulin, which disulfide bonded to an integral membrane protein.

Integrins are surface-membrane receptors consisting of one α-subunit noncovalently associated with a ß-subunit that spans the membrance and has a C-terminus binding site for cytoskeletal proteins. The ß-2 Leu-CAMs, first identified by monoclonal antibodies as the cluster designation (CD) 11 series: CD11a, CD11b, CD11c, and CD11d, are found on leukocytes and involved in chemotaxis, phagocytosis, adhesion, and migration across endothelial cells, lymphocyte proliferation, and natural killer (NK) cell and cell-mediated killing of target cells. Other ß-subgroup integrins (ß1, ß3, ß4, ß5, ß6, ß8) are found on other cell types and are involved in intracellular signaling events. Integrin-mediated signaling results in cell responses such as migration, proliferation, and gene induction. Leukocyte adhesion deficiency (LAD-1), a rare autosomal-recessive disorder where leukocytes are unable to migrate to tissue sites, is caused by a defect in ß2 integrins. LADs will be fully described in Chapter 65, Primary Immunodeficiency Disease.

Selectins bind directly to sialylated fucosylated lactose amines on proteins and glycosphingolipids. Three subgroups have been identified: L-selection, P-selectin, and E-selectin. L-selectin is expressed on prothymocytes, immature thymocytes, naïve T and B cells, monocytes and granulocytes where it appears to be responsible for leukocyte trafficking to peripheral lymph nodes and extravasation at inflammatory sites in nonlymphoid tissues. L-selectin is lost from the cell surface during activation and significant levels of the functional protein can be found in serum during active inflammation. P-selectin binding events are responsible for extravasation of leukocytes into inflamed tissues, phagocytic cell recruitment to sites of platelet activation, and phagocytosis of activated platelets. E-selectin is inducible on activated endothelial cells where it promotes binding of granulocytes and monocytes to activated endothelial cells. The synovial fluid of rheumatoid arthritis (RA) patients contains soluble adhesion receptors. Soluble (s)E-selectin levels are elevated in the peripheral blood as well as the synovial fluid of RA patients, associated with increases in erythrocyte sedimentation rate (ESR) and neutrophil infiltration into the synovium (2). Conventional therapy for RA may modify synovial inflammation by altering leukocyte adhesion.

Mucin segments on glycoproteins are responsible for cell-cell recognition and also modulate cell-cell adhesion, either promoting or impeding adhesion. N-acetyl galactosamine-containing O-linked oligosaccharides are found in cell surface mucins. Epithelial

mucins with specific repeats are present in carcinoma. Although many mucins are antiadhesive, when certain specific receptors (e.g., selectins) are present, mucins can promote rather than inhibit adhesion.

Proteoglycans are found in several forms: secreted, intracellular, and membrane-associated, and provide functions such as simple mechanical support, cell adhesion and migration, and cell differentiation and proliferation. CD44 proteoglycan, a major leukocyte hyaluronate receptor, is involved in lymphocyte homing to lymph nodes. Syndecans are a subfamily of heparin and chondroitin sulfate on cell surface proteoglycans expressed on epithelial cells where they serve as matrix receptors and coreceptors for certain growth factors. Sydecans also stabilize epithelial cell growth factors.

NONSPECIFIC CELLULAR RESPONSE TO INFLAMMATION

Part of "63 - Cellular and Humoral Mediators of Inflammation"

Granulocytes

Polymorphonuclear neutrophils (PMNs) are the largest population of leukocytes in the peripheral blood. These phagocytic cells contain distinct primary granules composed of proteases (cathepsin G, elastase, and proteinase 3); antibiotic proteins (myeloperoxidase, lysozyme, defensins, asurocidin, and bactericidal permeability-increasing protein [BPI]); and secondary granules that include lactoferrin, lysozyme, and human neutrophil lipocalin (HNL) (3). Cathepsin G is a proteolytic enzyme that can cleave complement factors, coagulation factors, and other proteins resulting in biologically active compounds and also can induce platelet aggregation. The proteolytic activity of Cathepsin G is neutralized by α-1 antichymotrypsin. Elastase and proteinase 3 are serine proteases with elasteolytic activities that contribute to tissue injury and may be involved with the glomerular injury in glomerulonephritis. Elastase and proteinase 3 activity can be inhibited by α-1 antitrypsin and α-2 macroglobulin. Antibiotic proteins in primary granules have microbicidal activities as well as other biologic functions. BPI/cationic antimicrobial protein (CAP57) is an effective killer of gram-negative bacteria and can bind to lipopolysaccharide (LPS) and neutralize some of the endotoxin effects on neutrophils. Lactoferrin, an iron-binding protein, is believed to interfere with the iron metabolism of microbes. Lysozyme is found in both primary and secondary granules, has bactericidal activities, and can be quantitated in plasma to measure monocyte/macrophage turnover in monocytic leukemia. HNL binds hydrophobic molecules (e.g., lipids and vitamins), and is measurable in serum and other body fluids as a specific marker of neutrophil activation to distinguish between bacterial and viral infections.

The mechanisms by which PMNs are involved in the nonspecific immune response include: (i) attachment to the damaged epithelium, (ii) amoeboid movement, (iii) emigration through the blood vessels, (iv) directed movement toward the particles to be engulfed, (v) ingestion of the bacteria, (vi) degranulation, (vii) oxidative metabolism (respiratory burst), and (viii) digestion of the foreign material.

Adherence of PMNs to the vascular endothelial cells occurs relative to membrane changes after exposure to chemoattractants that facilitate adherence through events associated with CAMs. Locomotion or ameboid movement, in conjunction with adherence to the walls of the capillaries, both are necessary for diapedesis to occur. During diapedesis (emigration), the pseudopods of the PMN are inserted between the junctions of endothelial cells. This enables the leukocyte to move from the inside of the blood vessel, through the capillary walls, between the endothelial cells, into surrounding tissues.

Chemotaxis is the directed migration of phagocytic cells mediated primarily by fluid-phase components of the complement system, particularly C5a. Other molecules known to be chemotactic include by-products of the coagulation or fibrinolytic pathways such as arachidonic acid metabolites and platelet activating factor (PAF), certain cytokines (interleukin [IL]-8), and various bacterial products. Chemoattractants form a gradient whereby PMNs move toward the increasing concentration of the substance.

Phagocytosis, the ingestion of particles, occurs when the external cell wall of the PMN adheres to and completely surrounds the offending bacteria or other particle. The foreign substance becomes encapsulated with a layer of inside-out membrane, a phagosome. Complement-activation products (C3b/iC3b) and immunoglobulins (IgG and IgA) can function as opsonins because PMNs and other phagocytic cells have receptors on their surface (Table 63.1) that specifically recognize those molecules coating bacteria or other particles. Neutrophils and eosinophils express two kinds of complement receptors (CR): CR1 (CD35) that interact with C3b and C4b, and CR3 (CD11b/CD18) that bind iC3b (cleaved and inactivated C3b). Three types of FcγR (IgG Fc receptors) are found on neutrophils: FcγRI (CD64), FcγRII (CD32), and FcγRIII (CD16). FcγRI are induced by the cytokine, interferon (IFN)-γ to mediate phagocytosis of monomeric IgG. FcγRII requires two or more molecules of IgG to induce phagocytosis, the respiratory burst, and degranulation. FcγRIII interacts with dimeric IgG to induce granule secretion without phagocytosis or respiratory burst. FcαR (IgA Fc receptor) are found in a low-affinity state on neutrophils and eosinophils, but can be upregulated to high-affinity receptors by the colony stimulating factors (CSF), GM-CSF (granulocyte/monocyte) and G-CSF (granulocyte). The FcR and CR selectively increase the type, rate, and quantity of particle uptake. During phagocytosis, O_2 and glucose are consumed, primarily through the hexose monophosphate shunt.

TABLE 63.1. MONONUCLEAR PHAGOCYTE CELL SURFACE RECEPTORS AND MARKERS

CD Designation	Molecule Function/Characteristic	Monocyte	Macrophage	PMN
	Immunoglobulin Receptors			
CD64	FcγRI, High affinity IgG receptor	++	++	+
CD32	FcγRII, Low-affinity IgG receptor	++	++	++
CD16	FcγRIII, Low-affinity IgG receptor	++	++	++
CD23	FcεRIIa,b, Low-affinity IgE receptor	--	--	++
	Complement Receptors			
CD11b	CR3. Binds C3bi	++	--	++
CD35	CR1. Binds C3b, C3bi, C3c, C4b	++	--	++
	Cell Adhesion Molecules (CAM)			
CD11a	LFA-1 α chain; p180	++	+	++
CD11c	CR4. P150,95 α chain	++	+	++
CD18	95 kD glycoprotein, linked to CD 11a,b,c	++	++	++
CD49f	Laminin receptor; VLA-α6 chain	++	??	--
CD29	VLA-ß chain	++	++	++
CD44	LECAM-III, hyaluronic acid receptor	++	++	++
CD58	LFA-3; binds CD2 molecule on T cells	++	++	++
CD36	Thrombospondin receptor; 85 kD glycoprotein	--	++	--
CD54	ICAM-1	+	++	+
CD102	ICAM-2	+	++	--
CD31	PECAM	++	+	++

CD, cluster of designation; LFA, leukocyte-function associated antigen; kD, kilodalton; VLA, very-late antigen; VCAM, vascular cellular adhesion molecule; LECAM, leukocyte endothelial cellular adhesion molecule; ICAM, intercellular adhesion molecule; CR, complement receptor; Ig, immunoglobulin; PECAM, platelet endothelial cell adhesion molecule.
Adapted from Thomas R, Wong R, Lipsky PR. *Textbook of Rheumatology*. W.B. Saunders (pub), Kelley WN, Ruddy S, Harris ED, et al. eds. St Louis, MO: W.B. Saunders, 1997:135.

Degranulation, the extracellular release of granule contents, can occur before (during chemotactic migration), during ("regurgitation while feeding"), or after phagocytosis (at cell death). The contents of the primary granules are released mainly into the phagocytic vacuoles, but the constituents of the secondary granules are often mobilized to the plasma membrane and released to the external environment. The granule contents are responsible for much of the tissue damage of inflammation.

During the respiratory burst that occurs as the phagocytosed particle is digested, the O_2-radicals formed are dismutated to H_2O_2. The mechanism by which many types of bacteria are killed following ingestion by phagocytes requires H_2O_2 and myeloperoxidase. Defects in these killing mechanisms result in a heterogeneous group of inherited disorders, chronic granulomatous disease, that will be discussed thoroughly in Chapter 65, Chronic Granulomatous Disease (and Related Disorders). Bystander tissues also can be damaged by the reactive oxygen species that might contribute to neutrophil-induced tissue damage at vascular sites of immune complex deposition, lung injury, or Crohn's disease, for example. See Venge, Bergstrand, and Hakansson, 1997 (3) for a more thorough review of the pathologic effects attributable to oxidative metabolism.

Eosinophils arise from a common progenitor with PMNs, are found in small quantities (1% to 4%) in peripheral blood, and are primarily found in tissues. Eosinophils are less efficient than PMNs at phagocytosis. Eosinophil granules are rich in acid phosphatase and peroxidase, but do not contain lysozyme. Eosinophil granules contain eosinophilic cationic protein (ECP), eosinophil protein X (EPX), and major basic protein (MBP), which are bactericidal and toxic to helminthic parasites (3). Eosinophils also produce leukotrienes and PAF. Functions attributed to eosinophils include clearance of parasites, ingestion of immune complexes, and antagonizing the effects of mediators released from basophils, mast cells, and platelets during inflammatory

reactions. Eosinophilia is associated with parasitic infections and allergic reactions.

Basophils and mast cells (MCs) originate from hematopoietic stem cells, but basophils complete their differentiation in the bone marrow while mast cells differentiate in peripheral tissues. Both release biochemical substances (mediators) that increase both vascular permeability and smooth muscle contraction to augment the inflammatory response. Basophil granules contain biogenic amines (histamine), proteoglycans (heparin and chondroitin sulfate E), neutral proteases (tryptase, chymase, cathepsin G, carboxypeptidase) and several sulfidopeptide leukotrienes (4). Histamine causes vascular dilatation and leakage by stimulating the H_1 receptor on vascular smooth muscle and endothelium. Basophils respond to IgE-containing immune complexes that bind to their IgE receptors to stimulate degranulation. The primary function of the basophil is to amplify and disseminate the reactions that begin with mast cells at the site of the entry of antigen. No specific diseases have been associated with basophilia.

Two distinct types of MCs have been described with different protease compositions of their granules. MC_T, containing only tryptase, is predominant in lung alveoli and small intestinal mucosa. MC_{TC} contains chymase, carboxypeptidase, cathepsin G, and tryptase, and are found in the skin, gastrointestinal submucosa, and synovium. Immunologic reactions, involving IgG or IgE binding to their respective receptors on the surface of mast cells or basophils, trigger degranulation and mediator release. Nonimmunologic mechanisms such as infections of the skin and mucous membranes, mechanical trauma (e.g.,surgical incisions), C5a, ECP, certain bacterial products, various chemokines, and other agents (e.g., opiates), also can stimulate mast cells to degranulate. The pharmacologic responsiveness of mast cells is dependent upon the tissue source (e.g., disodium cromoglycate inhibits the secretory response of lung-derived mast cells, but not skin or intestinal mast cells).

Platelets

Platelets play a critical role in coagulation, but also are involved in inflammation. Platelet *dense* granules contain mediators such as serotonin and adenine nucleotides that are released during platelet aggregation. Platelet *alpha* granules contain coagulation factors (fibrinogen, plasminogen, factor V), adhesion molecules (fibronectin, von Willebrand factor), plasma proteins (albumin and IgG), platelet-specific proteins [platelet factor 4 (PF-4)], ß-thromboglobulin and mitogenic proteins [platelet-derived growth factor (PDGF) and transforming growth factor-ß (TGF-ß)] (5). Platelets interact with PMNs and mutually enhance their activation. TGF-ß is a major product of platelets during the inflammatory process, augmenting the adhesion of leukocytes by inducing CAM receptors. TGF-ß also induces FcγRIII receptors on monocytes to promote phagocytosis, and stimulates the production of certain cytokines [IL-1, tumor necrosis factor (TNF), PDGF, and others]. TGF-ß also functions as an immunosuppressant for T and B lymphocyte growth and effector functions.

Mononuclear Phagocytes

Blood and bone marrow monocytes and tissue macrophages are mononuclear phagocytes that reside in every tissue and organ, known as the mononuclear phagocyte system. Their functions include removing debris such as old or injured erythrocytes, leukocytes, platelets, bacteria, antigen-antibody complexes, and degenerated or damaged cell membranes. Tissue macrophages include Küpffer cells of the liver; alveolar macrophages; splenic macrophages; histiocytes of the lymph nodes, peritoneum, and other areas; bone osteoclasts, nervous system microglial cells, and synovial A and C cells (6).

Mononuclear phagocytes express a wide variety of surface markers (Table 63.1). The primary role of FcR is to facilitate phagocytosis, but FcR also are involved in cell-mediated cytotoxicity, tumor immunity, and macrophage-T cell antigen presentation. Complement receptors (CR) bind different portions of C3 and C4 molecules to facilitate phagocytosis in synergy with FcR. CAMs mediate adhesion of leukocytes to endothelial cells and are involved in migration of macrophages to sites of tissue inflammation. Mononuclear phagocytes also have receptors for most of the cytokines (to be discussed later in this chapter), and for certain hormones.

Although mononuclear phagocytes can respond to chemoattractants, they primarily respond to soluble factors released from T lymphocytes (cytokines). Monocytes found in peripheral blood are not as efficient at phagocytosis and killing of bacteria as are PMNs, and their mechanisms of killing are not as well understood. Monocytes and macrophages secrete a wide variety of biologically active factors including peptide hormones (cytokines) that influence the activities of lymphocytes and other cells, complement components, coagulation factors, proteolytic and other enzymes (e.g., lysozyme), enzyme and cytokine inhibitors, CAMs and other binding proteins, bioactive lipids and other low-molecular-weight substances (6).

Phagocytic functions described previously for PMNs also occur with mononuclear phagocytes. Surface receptors enable the phagocytic cells to distinguish between normal cells and damaged cells or foreign bodies. Pinocytosis is another function of mononuclear phagocytes whereby soluble or fluid phase compounds are interiorized and concentrated. Phagocytosed and pinocytosed substances are digested intracellularly via an enzymatic process using lysosomal hydrolases to degrade internalized microorganisms or other ingested materials.

The functions of phagocytosis, pinocytosis, and intracellular digestion facilitate the antimicrobial and antiviral functions of mononuclear phagocytes. Reactive oxygen intermediates (antioxidants) are an important antimicrobial function for oxygen-dependent microorganisms. Oxygen-independent mechanisms include removal of iron by iron-binding proteins, direct actions of cationic proteins, and digestion by lysozyme and lysosomal hydrolases. Although certain cytokines produced by macrophages (IL-4, TGF-ß, IFN) suppress viral replication, ironically, macrophages also can provide a reservoir of protection for internalized viruses (e.g., HIV). Macrophages also play a role in tumor immunity in a non-phagocytic manner, by lysing neoplastic cells by a contact-dependent mechanism, i.e., antibody-coated tumor cells bind to Fc receptors on activated macrophages initiating killing mechanisms.

Although PMNs are end-stage cells that die after phagocytosis and degranulation, mononuclear phagocytes are stimulated (activated) by the phagocytic processes to become microbicidal

or tumoricidal, primarily by IFNγ secreted by T cells, but also by IL-2, IL-4, IL-13, and GM-CSF. Following phagocytosis, monocytes and macrophages become activated secretory cells that can proliferate locally within tissues.

Antigen presentation is one form of immunoregulation, an important function of a subset of macrophages called antigen-presenting cells (APC), that are regulated by the expression of class II MHC molecules on their surface. Class II MHC molecules can be upregulated by IFNγ, IL-4, IL-13, and GM-CSF and downregulated by IFNα, E series prostaglandins, TGF-ß, IL-6, IL-10, α-2 macroglobulins, TNF-α, and other molecules. APCs present antigens to T cells, which serve to stimulate the specific immune response.

Natural Killer Cells

Natural killer (NK) cell-mediated cytotoxicity, although not MHC-restricted, now is known to be regulated by the expression of class I MHC molecules expressed on target cells. Different subsets of NK cells express receptors that recognize MHC molecules, delivering inhibitory signals to NK cells to avoid eliminating normal self cells (7). Other receptors on NK cells trigger NK cells to eliminate allogeneic or autologous cells expressing "abnormal" MHC-molecules. Since NK cells also kill MHC-negative target cells, certain receptors might recognize non-MHC ligands. These unique NK receptors are summarized in Table 63.2.

TABLE 63.2. UNIQUE HUMAN NATURAL KILLER CELL RECEPTORS AND LIGANDS

CD Designation	Common Name	Ligand	Molecular Wt.	Receptor-Ligand Interaction Characteristic
		Inhibitory Receptors		
Killer inhibitory receptors – Ig superfamily				
CD158	NKIR	HLA-Cw3,4	58 kD	Inhibition of NK cell cytotoxicity
undesignated	NKB1	HLA-Bw4	70kD	Influence recognition mediated by NK cells
undesignated	undefined	HLA-A3,11	140 kD	Inhibition of NK cell cytotoxicity
Lectin superfamily				
CD161	NKR-P1	undefined	undefined	Inhibit or activate NK cell clones cytotoxicity
CD94	NKG2A-E	HLA-A,B,C	39/43 kD	Affect recognition HLA-A,B,C transfected targets
CD94	NKG2C	HLA-G	undefined	Protects trophoblast from maternal NK activity
		Activating Receptors		
CD158 a,b	KARAPS	HLA-C	50 kD	Activation of NK cell cytotoxicity
CD11/CD18	ß-2 integrins	ICAM-1,-2	150/95 kD	Tumor cell destruction by NK cells

CD, cluster of designation; KIR, killer inhibitory receptor; HLA, human leukocyte antigen; kD, kilodaltons; KARAPS, killer-activating receptor-associated proteins; ICAM, intracellular adhesion molecules.
Adapted from Thomas R, Wong R, Lipsky PE. Monocytes and macrophages. In: Kelley WN, Ruddy S, Harris ED, et al. eds. *Textbook of Rheumatology*, St Louis MO: W.B. Saunders, 1997:128–145 and Moretta A, Parolini S, Castriconi R, et al. Function and specificity of human natural killer all receptors. *Eur Immunogen* 1997;24:455–468.

NK cells are large granular lymphocytes (LGL), most of which express the surface antigen phenotype $CD3^-$ (T cell receptor), $CD2^+$ (E rosette receptor), $CD56^+$, $CD8^{+/-}$, $CD16^+$ (IgG FcR), $CD45^+$ (leukocyte common antigen), CD11a/$CD18^+$ [lymphokine function-associated antigen-1 (LFA-1)], CD11b/$CD18^+$ (Mac-1), and CD11c/$CD18^+$ (p150,95). NK cells do not recirculate between blood and lymph, but are found in the spleen, liver, and lungs (8). Subsets of LGL include NK cells, killer (K) cells, and lymphokine-activated killer (LAK) cells. Morphologically, these cells resemble lymphocytes, but function more like activated macrophages. These cells are termed LGL because they have azurophilic granules in their cytoplasm and have a high cytoplasm:nucleus ratio. The stages of effector function for LGL include: (i) target cell binding, when physical contact is made between the effector cells and the target cells; (ii) programming for lysis, during which the cytoskeletal components and Golgi apparatus of the effector cell move within the cytoplasm to the area of physical contact with the target cell; (iii) secretion of factors by the effector cell (e.g., NK cytotoxic factor, IL-1, and granule cytolysin); (iv) the cell-independent phase of the lytic event where soluble factors complete the killing process.

NK cells recognize and kill tumor cells *in vitro*. Similar to the activation of macrophages, NK cells can be activated by cytokines released by T helper cells when they encounter foreign antigens: all three classes of interferons (IFN α, ß, γ) and the interleukins (IL-2, IL-7, IL-12). Morphologic changes induced by these activators include a significant increase in the number of cytoplasmic granules. Unlike T and B lymphocytes, activation of NK cells does not induce immunologic memory. NK activity is suppressed by corticosteroids, hepatitis B surface antigen, and ligand-binding to the IgM Fc receptor site. NK cells rapidly and efficiently secrete several cytokines when activated: IFNγ, TNF-α, IL-1, IL-3, and GM-CSF.

Killer (K) cells cannot be distinguished from NK cells by their morphologic appearance or by their surface markers. K cells are characterized by their ability to bind antibody-coated target cells (via their IgG Fc receptors) and subsequently lyse those cells. These target cells can be erythrocytes, bacteria, parasites, or tumor cells. K cells appear to represent a particular stage of NK cell development or activation.

LAK cells can be generated *in vitro* using purified or recombinant IL-2. LAK cells have many of the same cell surface markers as NK cells and apparently are previously unstimulated NK cells. The LAK activity reflects the potent ability of IL-2 both to stimulate cytotoxic activity and to expand the population(s) of LGL effector cells. *In vitro* generated LAK cells are used in treatment of some forms of cancer.

Endothelial Cells

Endothelial cells line the cavities of the heart, the blood vessels, and the lymphatic system, forming an interface between blood and tissues. By responding to vasodilatory mediators, endothelial cells contract and relax, forming intercellular gaps that allow for passage of fluids between the tissues and blood vessels. Vasoactive amines such as histamine, serotonin, bradykinin, C3a, C5a, PAF, and leukotrienes can cause increased permeability of the endothelial cells resulting in vascular leakage. Endothelial retraction also can be induced by certain cytokines. Leukocyte-endothelial interactions may injure endothelial cells and cause increased vascular permeability. Immune-complex mediated injury to endothelial cells accompanied by complement activation allows the exudation of leukocytes as well as plasma proteins. Activation of endothelial cells results in synthesis and secretion of CAMs including E-selectin, ICAM-1, and VCAM-1 as well as a variety of cytokines, e.g., IL-1, IL-6, IL-8, PDGF, and TGF-ß(9). Lymphocyte homing to peripheral lymph nodes is controlled by CAMs on lymphocytes and endothelial cells. Angiogenesis is governed by endothelial cells and their extracellular matrix. Endothelial cells also contribute to T cell activation.

Recent evidence suggests that endothelial cell dysfunction is involved in atherosclerosis. Factors possibly involved in endothelial cell dysfunction in atherosclerosis include elevated low density lipoproteins (LDL); hypertension; diabetes mellitus; elevated homocysteine levels; certain infectious diseases; free radicals resulting from cigarette smoking; and combinations of these and other causes (10). The inflammatory response to these factors results in adhesiveness of the endothelium, deposition and activation of platelets and phagocytic cells, leading to the release of hydrolytic enzymes, cytokines, chemokines, and growth factors that induce further damage. When the artery can no longer dilate to compensate for the lesion formed, the flow of blood is altered leading to myocardial infarction (10).

NONSPECIFIC HUMORAL RESPONSE TO INFLAMMATION

Part of "63 - Cellular and Humoral Mediators of Inflammation"

Acute-Phase Response

When the body is injured, it responds by increasing the hepatic synthesis of a number of plasma proteins. The systemic acute-phase response helps to ensure survival during the period immediately following injury. The systemic response must help to achieve the same goals as the localized inflammatory response, i.e., to contain or destroy infectious agents, to remove damaged tissue, and to repair the affected organ.

Fever was one of the first acute-phase responses recognized. Fever may occur following many types of inflammatory stimuli including non-infectious states. Fever reflects the effects of endogenous pyrogens (IL-1, IL-6, TNFα, IFNγ), which elevate the set point of the hypothalmic center for body temperature. Another long recognized, but variable, acute-phase response is an increase in the granulocyte count in the blood. This initially reflects release from the storage pool and later reflects increased production by the bone marrow. A principle reason for the leukomoid response to inflammation is to provide a source of lysosomal enzymes.

Lysosomal Enzymes

During phagocytosis and following the death of phagocytes, lysosomal contents of granules are released to enhance inflammation and provide microbicidal activity. Much of the tissue damage of inflammation is from destruction of “innocent bystander” cells by lysosomal enzymes released during the inflammatory process.

C-Reactive Protein

C-reactive protein (CRP) was recognized in 1930 because of its ability to precipitate with the C-polysaccharide extract of pneumococcus (11). Although CRP is distinct from antibody, many parallels between the two molecules exist. CRP will react with its substrate to cause lattice formation and precipitation. CRP can promote passive agglutination of red blood cells coated with binding substrate. An example of the similarity between CRP and immunoglobulins is the initiation of the complement cascade through C1 activation by complexed CRP (analogous to antibody-antigen complexes). Opsonization for ingestion by phagocytes can result from complement activation by CRP.

Unlike antibody, CRP is produced by hepatocytes. CRP is a pentamer with five identical, noncovalently linked subunits. The binding of CRP to C-polysaccharide or other phosphocholine-containing compounds is calcium-dependent. Approximately one phosphate is bound to each of the 5 CRP subunits, requiring two Ca^{2+} ions per phosphate binding site.

CRP may be considered a primitive form of an antibody molecule, with specificity for components found in cell membranes of microorganisms (bacteria, fungi) and for damaged membranes of cells. When complexed to a binding specificity, CRP activates complement via the classical pathway to opsonize and clear microorganisms and serves as a regulatory protein in alternative pathway activation (12). Complexed CRP binds to LGL and to macrophages and can activate these cells to be tumoricidal (13, 14).

Although CRP has many similarities to antibodies, deficiency states of CRP have not been found. CRP is present in microgram per liter quantities in normal humans, but elevates dramatically and rapidly in the presence of bacterial infections where levels >100 mg/L are common. With resolution of those infections, the CRP levels decrease within a few days, reflecting the half-life measured in hours (Fig. 63.2). Moderate CRP levels of 10 to 100 mg/L are found in chronic inflammatory conditions such as systemic lupus erythematosus (SLE), malignancies, congestive heart failure, and pregnancy. Viral infections, by contrast, do not manifest a CRP response unless there is superimposed bacterial infection.

CRP is an especially useful marker of inflammation to monitor patients after surgery. In uncomplicated cases, the CRP levels peak between 48 and 72 hours and return to normal by day 7. If the postoperative course is complicated by inflammation or sepsis, the CRP levels remain elevated or increase (15). For CRP to be a valuable tool in assessing inflammatory complications, the test must be available 24 hours/day, 7 days/week. Recently, CRP has been proposed as predictor of coronary heart disease (CHD) (16, 17). For CRP to be used effectively to predict CHD, the test must be able to detect the protein at low levels (1 to 10 mg/L) as well as the high

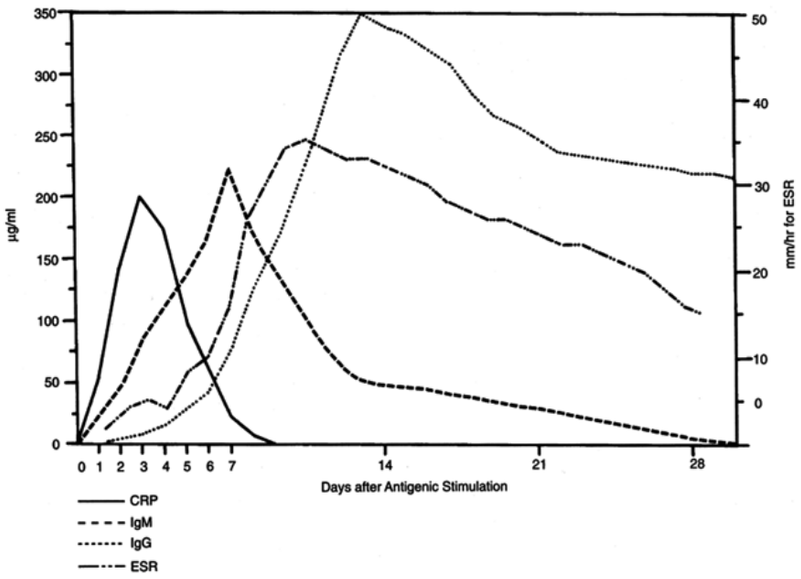


FIGURE 63.2. Kinetics of CRP and ESR responsiveness. (From James K. Immunoserology of infectious diseases. *Clin Microbiol Rev* 1990;3:135–152.)

levels more commonly seen in acute infectious process (100 to 400 mg/L). The technology is available to perform CRP as a random access assay and to detect low levels, but to-date, has not yet been routinely established in clinical laboratories to meet the clinical needs.

Erythrocyte Sedimentation Rate

Measuring the erythrocyte sedimentation rate (ESR) has become a commonly used index of inflammation, most useful in assessing chronic inflammation because the ESR elevates much more slowly than does CRP. The ESR measurement is directly proportional to the plasma level of fibrinogen, haptoglobin, or other acute phase reactants (18). Immunoglobulin elevations (e.g., macroglobulinemia, myeloma) also affect *rouleaux* formation and cause the ESR to rise. Alterations in erythrocyte size or shape (e.g., macrocytes, microcytes, sickle cells) also influence the ESR.

In contrast to CRP, the ESR elevation is slower and much more prolonged (Fig. 63.2), reflecting the significantly slower indirect stimulus of fibrinogen synthesis by hepatocytes. The magnitude of increase of fibrinogen (two to fourfold) also is significantly different from that of CRP (elevates several hundredfold). Certain chronic inflammatory diseases such as SLE have very elevated ESRs, but normal to only slightly elevated CRP levels. Rheumatoid arthritis, in contrast, demonstrates elevations in both ESR and CRP. Although the ESR has certain applications as an indicator of inflammation, the test cannot be performed STAT and should be replaced by random access quantitation of CRP by immunoassay for most of the current uses of the test.

Complement

The complement system consists of 16 components involved in three separate pathways of activation. Five proteins unique to the classical pathway include the trimolecular complex of C1 (C1q, C1r, C1s), C4 and C2. Three proteins unique to the alternative pathway include factor B, factor D and P (properdin). Two components are unique to the lectin pathway: mannose-binding lectin (MBL) and MBL-associated serine protease (MASP). Six components participate in all three pathways: C3, C5, C6, C7, C8 and C9. Because the components were named in the order discovered, the sequence of activation is not in numerical order, but the components interact in a specific cascading sequence (Fig. 63.3). All three pathways can be divided into three units (recognition, activation, and membrane attack) to simplify the understanding of complement activation (19).

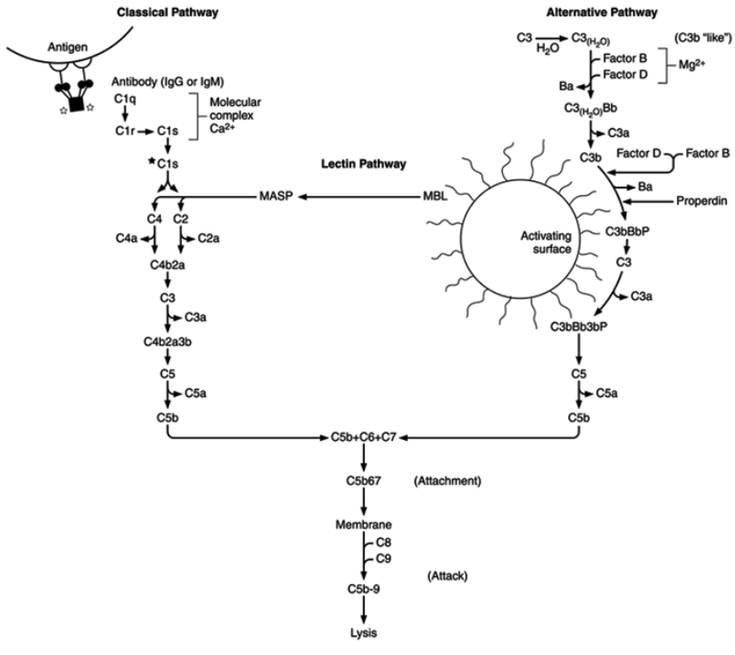


FIGURE 63.3. Complement pathways.

Classical Pathway Recognition Unit

The C1q molecule contains a collagenous region with six globular head groups. When specific antibody interacts with its corresponding

antigen, binding sites are exposed on the Fc region of the antibody molecule for the globular head groups of C1q. Two molecules of IgG or CRP, or one molecule of IgM, are required for the binding of C1q. In circulation and in the presence of calcium, the collagen portion of C1q is surrounded by two molecules each of C1r and C1s. When C1q binds to the Fc region of antibody or CRP, a conformational change occurs in C1q. This change in C1q causes the proenzyme C1r to become the enzymatically active C1r. The substrate for the enzyme C1r is C1s, which then is cleaved to become the serine esterase, C1s.

Classical Pathway Activation Unit

The active enzyme of C1s cleaves C4 and C2 in a magnesium-dependent reaction. C4b and C2a combine to form an active enzyme C4b2a, the classical pathway C3 convertase, with C4a and C2b as by-products. The enzymatically active C4b2a complex can cleave many molecules of C3 into C3a and C3b. The C3b can then either form a covalent bond with the antigen or with bystander surfaces (e.g., erythrocytes) in immune adherence, or can bind to C4b2a to form C4b2a3b, an enzyme with specificity for C5. The final enzymatic step of the classical complement pathway is the cleavage of C5 into C5a and C5b by the C5 convertase, C4b2a3b. At this point, the classical pathway and the alternative pathway converge.

All Pathways Membrane Attack Unit

C5b binds to one molecule of C6 to form a stable bimolecular complex, C5b6. If C7 is present, the trimolecular complex C5b67 is formed, which binds hydrophobically to a membrane. Once C5b67 is bound, C8 can attach to form a functional transmembrane channel. Up to 6 molecules of C9 then surround the

lesion to prevent the channel from being resealed. C9 accelerates lysis, but is not essential for the lytic event.

Lectin Pathway Recognition and Activation Unit (20)

MBL binds to specific carbohydrates at the surface of microorganisms and triggers activation via the mannose residues on bacterial surfaces. MASP then cleaves C4 and C2 enzymatically to form the active enzyme C4b2a.

Alternative Pathway Recognition Unit

Activation of the alternative pathway requires an activating surface not protected by sialic acid. Substances known to provide an activation surface are bacterial cell walls, bacterial lipopolysaccharide (LPS), fungal cell walls, certain virus-infected cells, and rabbit erythrocytes. The “activating surface” actually is a protective surface, protecting spontaneously hydrolyzed C3 (nonenzymatically cleaved into C3a and C3b) from being inactivated by the control proteins (21). Hydrolyzed C3 becomes C3b-like. In the presence of Factor D and Mg^{2+}, this C3b-like molecule can cleave Factor B into Ba and Bb. Ba becomes a by-product while Bb binds to the C3b to form an alternative pathway C3 convertase, C3bBb. By itself, C3bBb is a very unstable molecule and is quickly inactivated by control proteins unless it binds to an activating surface.

Alternative Pathway Activation Unit

When protected by an activating surface and stabilized by P (properdin), the C3bBbP enzymatic complex can cleave additional molecules of C3. If a second C3b molecule is inserted into the C3 convertase to become C3bBb3bP, this becomes a C5 convertase, which can cleave C5 into C5a and C5b.

Membrane-Associated Complement Proteins (Receptors and Regulators) (22)

Complement Receptor Type 1

CR1 binds C3b and C4b. These receptors are present on erythrocytes, PMNs, monocytes, B cells, NK cells, some T cells, eosinophils, mast cells, renal podocytes, and follicular dendritic cells. CR1 is involved in the clearance of immune complexes and also accelerates the decay of C3 and C5 convertases. CR1 has the cluster designation CD35.

Complement Receptor Type 2

CR2 binds C3d; C3d,g; and iC3b. They are present on B cells, NK cells, and follicular dendritic cells. CR2 is involved in regulation of B cell functions such as retention of memory and induction of antibody. CR2 is the receptor for Epstein-Barr virus. CR2 has the cluster designation CD21.

Complement Receptor Type 3

CR3 binds iC3b and has the cluster designation CD11b/CD18. CR3 is found on PMNs, monocytes, macrophages, NK cells, and cytotoxic T cells. CR3 is the specific receptor for phagocytosis.

Complement Receptor Type 4

CR4 also binds iC3b, but has the cluster designation CD11c/CD18. It is found on the same cells as CR3.

Membrane Cofactor Protein

MCP is involved in the breakdown of C3b and C4b on cell surfaces as a cofactor, but does not have decay accelerating activity. MCP is present on most cell types except erythrocytes and has the cluster designation CD46.

Decay Accelerating Factor

DAF accelerates the decay of C3 and C5 convertases, but does not have cofactor activity. It is present on most cell types, including erythrocytes. DAF has the cluster designation CD55.

Homologous Restriction Factor

HRF binds C8 and C9 to prevent insertion of the membrane attack complex into the cell membrane. It is found on erythrocytes and is restricted to homologous species.

Membrane Inhibitor of Reactive Lysis

MIRL binds C5b-8 and prevents formation of the membrane attack complex by preventing polymerization of C9. MIRL is found on erythrocytes, renal cells, and most other cell types and has the cluster designation CD59.

Biologic Consequences of Complement Activation

Amplification

C3b can be generated by C3 convertase from either the classical or lectin pathway (C4b2a) or the alternative pathway (C3bBbP). This provides a feedback loop that uses the alternative pathway components (B, D, P) to amplify the activation of the C3 through C9 components of activation and membrane attack.

Anaphylatoxin

The cleavage of C4, C3, and C5 results in the release of the biologically active peptides C4a, C3a, and C5a. These anaphylatoxins mediate inflammation by inducing the release of histamine from basophils and mast cells, by causing smooth muscle to contract, and by increasing vascular permeability.

Immune Adherence

Immune adherence is the covalent bonding between the cleaved form of C3 (C3b) and nearby soluble immune complexes or particulate surfaces. The portion of the C3b that does not adhere is exposed and available for binding to the CR1 for C3b or C4b on human erythrocytes, B cells, NK cells, monocytes, renal podocytes, eosinophils, or mast cells. B cells and macrophages also have CR2s for C3d, which is formed by cleaving C3b into C3c and C3d (Fig. 63.4). One biologic purpose for immune adherence

would be to facilitate removal of soluble circulating immune complexes (CIC) by first binding to erythrocyte CR1, then interacting with CR3, stimulating phagocytosis of the erythrocyte-bound CIC.

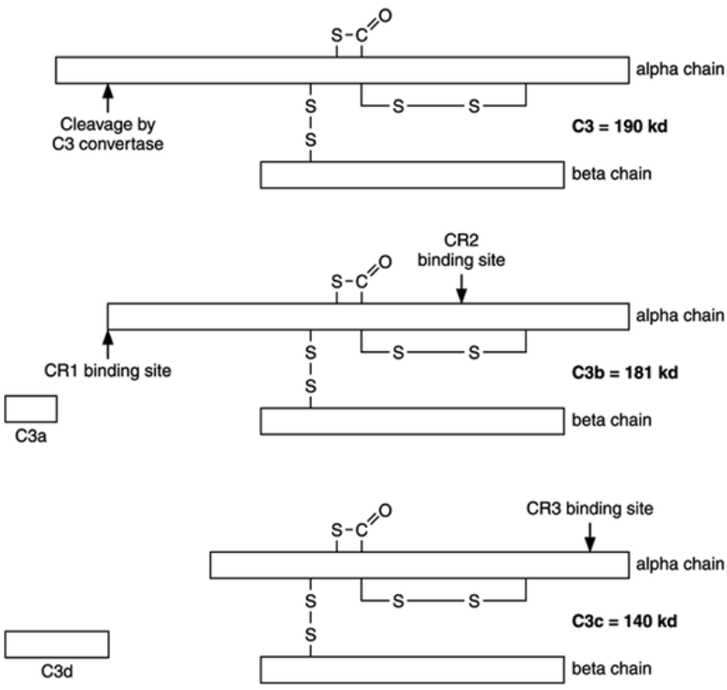


FIGURE 63.4. C3 Molecular structure.

Opsonization and Phagocytosis

Deposition of C3b opsonizes a surface, for more effective phagocytosis by PMNs and/or monocytes. The CR3s on these phagocytic cells bind to the exposed C3b on the surface of the particle. The membrane of the phagocytic cell surrounds the opsonized particle, the cell membrane fuses together thereby phagocytosing the particle. The cross-linking of CR1 on phagocytic cells stimulates the secretion of lysosomal enzymes, involved in killing of microorganisms and contributing to tissue damage at inflammatory sites.

Chemotaxis

The by-product resulting from the cleavage of C5 is C5a, a potent chemoattractant as well as an anaphylotoxin. C5a induces the directed migration of neutrophils and monocytes into the area of inflammation.

Kinin Activation

The fragment of C2 (C2b), which is released during cleavage by C1s, interacts with plasmin to produce kininlike activity. The biologic activity of C2b results in smooth muscle contraction, mucous gland secretion, increased vascular permeability, and pain.

Lysis or Direct Killing

In the laboratory, the activity of complement is studied by measuring the degree of lysis of sheep erythrocytes (ShE) that occurs. Lysis, however, plays a relatively minor biologic role. One example of lysis as the biologic consequence of complement activation would be an antibody-mediated transfusion reaction. The lytic function of complement also appears to be necessary for host defense against *Neisseria* spp.

Control Mechanisms of Complement Activation

If any or all of the biologic consequences of complement activation were to go uncontrolled, the effects of even minor inflammatory processes involving activation of either complement pathway would be potentially devastating. The body, however, does not leave a reaction uncontrolled. The first means of control is the extreme lability of activated complement components. If an activated enzyme does not combine with its substrate within milliseconds, the activity is lost or markedly decayed. “Innocent bystander” cells in the vicinity of activated complement would be rapidly destroyed if the activated components were not so highly labile. Additionally, there also are several inhibitors or inactivators of specific reactions or products involved in the complement cascade.

C1 Inhibitor (C1Inh)

C1Inh forms an irreversible complex with both C1r and C1s, which blocks their enzymatic activities and dissociates them from C1q. The hereditary or acquired deficiency of this protein results in uncontrolled activation of the classical pathway. Control proteins that exercise their activity at the level of C3 or later are still functioning, retarding the amplification loop and other biologic consequences of C3-C9 activation. C1s, in the absence of C1Inh, continues to cleave C4 and C2 unchecked, resulting in release of C2b kininlike activity and C4a anaphylatoxin activity.

ß1H (H) and C3b Inactivator (I)

The most important biologic consequence of complement activation is the feedback loop amplification mediated by C3b. Proteins H and I serve to tightly control the enzymes that cleave C3 and C5. H accelerates the decay of the alternative pathway C3 convertase by dissociating Bb from the enzyme, while I inactivates C3b and C4b. H and I are both involved in cleaving C3b into its hemolytically inactive form, C3bi, which is further cleaved into C3c and C3d. Fluid phase C3b is rapidly inactivated by H and I. Consequently, activation of the alternative pathway is dependent upon the presence of a protective (activating) surface which shelters C3b from these two control proteins.

C4 Binding Protein (C4BP)

C3b inactivator (I) can also cleave and inactivate C4, but requires an accessory protein, C4BP.

Anaphylatoxin Inactivator

Carboxypeptidase controls the effects of C4a, C3a, and C5a by removing a single amino acid, a carboxyterminal arginine. Cleavage of this amino acid destroys the anaphylatoxin activity of these peptides.

Properdin (P)

The above control mechanisms are all inhibitors. Properdin is an enhancer. While not required for the activation sequence of the alternative pathway, P stabilizes the C3 and C5 convertases to prolong their activity.

C3 Nephritic Factor (NF)

NF is a pathological enhancing protein. NF is an IgG antibody with specificity for the alternative pathway C3 convertase (23). NF binds to the C3 convertase in such a way that it prevents inactivation by the control proteins H and I. When NF is present, C3 activation proceeds uncontrolled, thereby markedly depleting C3.

Measuring Complement in the Clinical Laboratory

The complement components usually quantitated are C3 and C4; C3 because of its pivotal location and role in both the classical and alternative pathway, and C4 to distinguish classical from alternative pathway activation. All complement components can be quantitated, however, components other than C3 and C4 rarely are required and are best performed by research laboratories. The only reason to measure any component other than C3 and C4 would be if a component deficiency is suspected. The screening method of choice for a component deficiency is the hemolytic assays.

Hemolytic complement assays are measured as CH_{50} (hemolytic complement 50% lysis point), which is a fluid phase assay, or CH_{100} (hemolytic complement 100% lysis), which is performed in agar. In both systems, either the classical pathway or the alternative pathway can be measured by varying the indicator cells. For the classical pathway hemolytic assays, sheep erythrocytes (ShE) are sensitized with antibody (ShEA). For the alternative pathway hemolytic assays, rabbit erythrocytes (RaE) are used without antibody. In the fluid phase assay, the ShEA (or RaE) are suspended in buffer containing the appropriate divalent cations (Ca^{2+} for the classical pathway or Mg^{2+} for the alternative pathway). Because complement components are normally present in excess levels from what is needed for lysis of either pathway, the CH_{100} would be less sensitive to half levels (heterozygous deficiencies) of certain components than the fluid phase assay would be. Because complement deficiencies are relatively rare, and half-levels are only relevant in complicated disease processes, CH_{100} can replace CH_{50} in most situations. The solid phase assay is much easier to perform, detects consumption of components in disease states, and does detect homozygous component deficiencies.

Neither CH_{50} nor CH_{100} are useful to monitor disease activity because their endpoint (lysis) is very insensitive to changes in component levels and can occur even with significant depletion of most components. It takes very few C2 and C4 molecules to cause lysis and yield a “normal” CH_{50} level. Quantitation of C3 and C4 are much more sensitive measures of complement consumption *in vivo* than are the hemolytic assays. The patterns of complement consumption in various disease states are shown in Table 63.3.

TABLE 63.3. INTERPRETATION OF LABORATORY TESTS FOR COMPLEMENT ACTIVATION

↓ CH_{50}, ↓ C3, ↓ C4	↓ CH_{50}, ↓ C3, Normal C4
Classical Pathway Activation	*Alternative Pathway Activation*
Active SLE	Acute post-strep GN
Serum sickness	Membranoproliferative GN
Immune complex disease	Severe bacteremia
Chronic active hepatitis	C3 nephritic factor
Subacute bacterial endocarditis	C3, H, or I deficiency
↓ to Normal CH_{50}, Normal C3, ↓ C4	↓ CH_{50}, Normal C3, Normal C4
Mild Classical Pathway Activation	*Defective Hemolytic Function*
Cryoglobulinemia	Improperly handled sera
Vasculitis	Coagulation-associated depletion
C1Inh deficiency	
C4 deficiency	Component deficiency (e.g., C2, C5-C8)

Complement components are acute-phase reactants and increase in concentration during inflammation. Consequently, the acute-phase increases can offset any consumption *in vivo* and result in normal hemolytic and normal quantitative levels in the presence of subclinical disease. C4 is the most sensitive indicator of *in vivo* complement consumption. Because C3 levels are 10-fold

greater than C4 levels, it takes significantly more consumption of C3 to offset an acute phase response than it does for C4.

Unfortunately, some of the rheumatology literature still considers CH_{50} as the assay of choice to monitor complement activation in disease. This is attributable to a lack of understanding by commercial companies and their users of the need to use monospecific antisera to C3c or another unique domain on the C3 molecule (Fig. 63.4). Some commercially available C3 assays use antisera that react with multiple domains on the C3 molecule and therefore detect breakdown products of C3 as well as the intact molecule, resulting in falsely elevated (or falsely normal) levels in active stages of disease. Falsely elevated levels are not as pronounced with C4 because the concentration is lower.

Specimen handling is critical for accurate measurement of complement components, and especially for the functional assays. Complement components (many of which are enzymes) are extremely labile and are easily activated in the test tube or inactivated due to adverse storage conditions. Although not considered a "standard of practice," specimens for quantitation of C3 and C4 AND for functional assays should be collected as plasma with EDTA as the anticoagulant. EDTA chelates the available Ca^{2+} and Mg^{2+}, preventing activation of either pathway *in vitro*. The plasma should be frozen until tested to prevent breakdown by proteolytic enzymes as well as preserve the labile complement components. The buffer systems used to test for functional activity in either the CH_{50} or the CH_{100} assay contain an excess of Ca^{2+} and Mg^{2+}, more than enough to support either the classical or alternative pathway without these divalent cations from serum. The amount of EDTA present in the plasma is insufficient to chelate the excess divalent cations provided by the buffer systems. Collection of complement specimens as EDTA plasma prevents the artificial activation of complement *in vitro* associated with the coagulation pathway in certain patient sera (24).

Disease resulting from deficiencies of complement components and control proteins are thoroughly discussed in Chapter 65, Primary Immunodeficiency Diseases.

Cytokines and Cytokine Receptors (CR) (25,26,27,28,29,30,31,32,33,34,35,36,37,38 and 39)

Cytokines are small polypeptide or glycopeptide molecules (mol. wt. 5 to 70,000 daltons) whose primary functions are to mediate infectious and inflammatory processes. Cytokines act at very low concentrations as pleiotropic hormones, overlapping suppressive functions and stimulatory events. The inflammatory cytokines (IL-1α, IL-1ß, TNF-α, IFN-α, -ß, -γ) are involved in the pathogenesis of many infectious and autoimmune diseases. The antiinflammatory cytokine (IL-10) profoundly suppresses macrophages, inhibiting the synthesis of proinflammatory cytokines. Cytokines can be categorized loosely as interleukins, interferons, cytotoxins, chemokines, and growth and stimulatory factors. Fig. 63.5 illustrates some of the functions and interactions

of cytokines. Table 63.4 is a partial list of cytokines, the cells producing them, their target cells, and their primary functions.

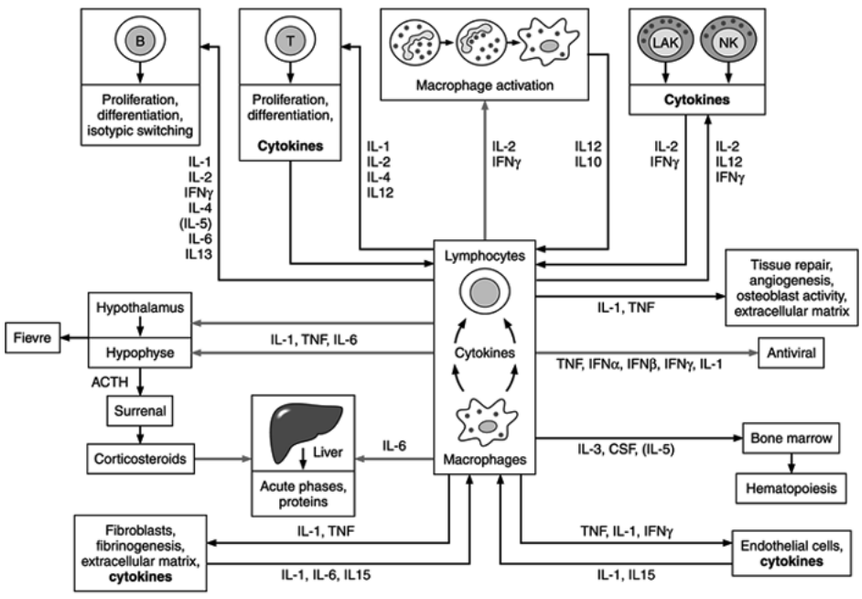


FIGURE 63.5. Selected functions of cytokines. Reproduced with permission from Fridman WH, Tartour E. Cytokines and cell regulation. *Mol Aspects Med* 1997;18:3–90.

Cytokine receptors (CyR) are noncovalently associated with intracellular tyrosine kinases. CyRs can be classified based on their subunit composition (25) (i) homodimers where two receptor molecules bind one molecule of ligand (e.g., hGH-R, EPO-R, G-CSF-R); (ii) two different subunits (α and ß) where the α subunit functions primarily in ligand binding while the ß subunit functions as a signaling subunit (e.g., IL-6R, GM-CSF-R, IL-3R, IL-5R); and (iii) three different subunits (α and ß and γ) where one of these subunits is common to several different receptors (e.g., IL-2R, IL-4R, IL-7R, IL-9R, IL-13R, and IL-15R).

Certain of the CyRs are found in the peripheral circulation as soluble cytokine receptors (sCyR). They are synthesized either by proteolytic cleavage or shedding (sIL-1R, sIL-6R, sTNF-R) or by differential splicing leading to the production of an altered receptor lacking the transmembranous and cytoplasmic regions (sIL-4R, sIL-5R, sIL-6R, sIL-7R, sIL-9R) (26). sCyR may modulate the host response to the cytokine by binding to the cytokine to block interaction with cell-bound CyR, either reversibly when low affinity or irreversibly whereby the cytokine-sCyR complex is rapidly removed from circulation (Fig. 63.6).

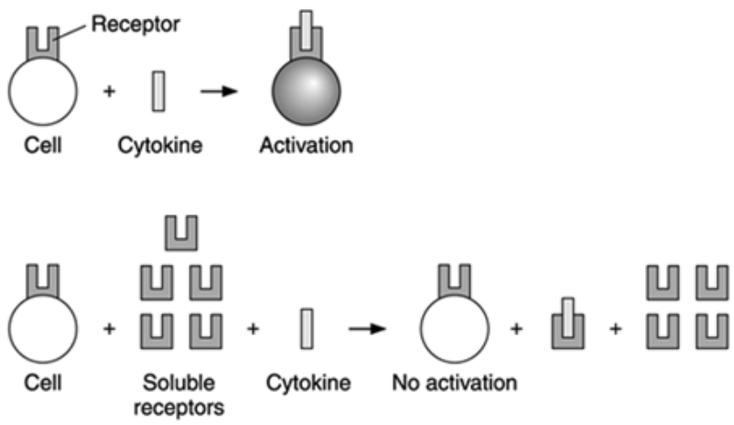


FIGURE 63.6. Cytokine receptors and soluble cytokine receptors. Reproduced with permission from Remick DG. Applied molecular biology of sepsis. *J Critical Care* 1995;198–212.

Interleukins and Interleukin Receptors

Most interleukins are synthesized by activated T-lymphocytes, however IL-1 and several other interleukins are also secreted by

TABLE 63.4. A PARTIAL LIST OF HUMAN CYTOKINES[27,28]

Cytokine	Cells Producing	Target Cells	Primary Functions
IL-1α	Macrophages, B cells, NK	T, B, NK cells, eosinophils, hepatocytes, synovial cells, many other cell types	Induces fever, acute phase protein synthesis by hepatocytes, hypotension, anorexia, phagocytosis, hematopoiesis, hormones (insulin, ACTH, cortisol)
IL-1ß	cells neutrophils, synovial cells, many other cell types		
IL-1ra	Hepatic cells, monocytes/ macrophages, PMNs, endothelial cells, fibroblasts, many other cell types	Cells expressing IL-1R	Inhibits the effects of IL-1α and IL- 1ß by specifically blocking the IL-1 receptor on target effector cells.
IL-2	T cells	T, B, NK cells	Growth factor for target cells
IL-3	T cells	Stem cells, mast cells	Hematopoietic growth factor
IL-4	T cells, B cells, stromal cells	T cells, B cells, hematopoietic precursors	T & B cells growth factor, induces isotype shift in B cells (IgE and IgG_4), regulates antibody production
IL-5	T cells, mast cells	Eosinophils, B cells, hematopoietic Precursors	Eosinophil growth factor, bone- marrow cells maturation and differentiation
IL-6	Macrophages, T cells, hepatocytes, endothelial cells, fibroblasts, many other cell types including neoplastic cells	Hepatocytes, megakaryocytes, bone marrow precursors, and many other cell types	Stimulates production of acute-phase proteins by hepatocytes, promotes immunoglobulin secretion by activated B cells, induces maturation of megakaryocytes and other bone marrow precursors
IL-7	Bone marrow and thymic stromal cells	Pre-T and pre-B cells	Promotes the maturation of T & B cells, enhances NK cell function
IL-8	Monocytes, T cells, many other cell types	PMNs, T cells	Activates PMNs, chemotactic for PMNs, T cells, basophils
IL-9	T cells	T cells, erythroid progenitors, mast cells, megakaryocytes	Hematopoietic growth factor
IL-10	T cells, B cells, macrophages	T cells, B cells, mast cells, monocytes	Regulates T cell and macrophage function, coactivates B cells for secretion of IgA, IgG_1, and IgG_3; induces IL-1ra
IL-11	Stromal cells	Hematopoietic stem cells, megakaryocytes	Supports proliferation of hematopoietic precursors
IL-12 (NKCF)	Monocytes, B cells	T cells, NK cells	Activates the cytolytic potential of T cells and NK cells
IL-13	T cells	B cells, NK cells, monocytes	Similar to IL-4, promotes B cell growth and differentiation for production of IgE and IgG_4
IL-14	T cells	Hematopoietic stem cells	Hematopoietic growth factor
IL-15	Monocytes, fibroblasts, endothelial cells, almost all nucleated cell types	T cells, NK cells	Similar to IL-2, growth factor for target cells
IFN-α1/α2	T cells, B cells, monocytes/macrophages	Macrophages, NK cells, B cells, many other cell types	Activating factor, MHC I and II modulcation, antiviral activity
IFN-ß	Fibroblasts, many virus infected cells	Macrophages, NK cells, many other cell types	Similar to IFN-α, activating factor
IFN-γ	T cells, NK cells, endothelial cells	Monocytes/macrophages, fibroblasts, smooth muscle cells, many other cell types	Stimulates vasoconstriction of smooth muscle cells, modulates MHC I & II expression, inhibits cell growth
TNF-α	Macrophages, PMNs, T cells, NK cells	Lymphocytes, PMNs, eosinophils, endothelial cells, fibroblasts, other cells types, virus-infected cells	Cytotoxic to virus-infected cells, induces IL-1 and IL-6 in many cells, activates lymphocytes, granulocytes, endothelial cells
TNF-ß	T cells	Similar to TNF-α	Similar to TNF-α, lymphotoxin, cytotoxic to virus-infected cells
Lymphotoxin	T cells	Similar to TNF-α	Similar to TNF-α
MCP	Monocytes/macrophages, fibroblasts, endothelial cells, B cells, smooth muscle cells	Monocytes/macrophages	Chemokine for monocytes; regulates expression of adhesion molecules on the surface of macrophages
RANTES	T cells, platelets, renal epithelium, mesangial cells	Monocytes, T cells, eosinophils, basophils	Chemokine for monocytes and eosinophils; enhances basophil histamine release
MIP-1 α,ß	T cells, B cells, monocytes, mast cells, fibroblasts	Monocytes, T cells, PMNs, eosinophils (MIP-α)	Chemokine for target cells; activates monocytes
GM-CSF	T cells, fibroblasts, endothelial cells, macrophages, mast cells, eosinophils	Stem cells, T cells, monocytes, granulocytes	Activation and differentiation
G-CSF	Monocytes/macrophages, fibroblasts, endothelial cells, T cells, PMNs	PMNs and precursors	Activation and differentiation of granulocytes
M-CSF	Monocytes/macrophages, fibroblasts, endothelial cells	Monocytes and precursors	Activation and differentiation of monocytes
TGF-α	Macrophages, keratinocytes, pituitary cells, brain cells	Fibroblasts, epithelial cells	Activates fibroblasts, angiogenic
TGF-ß	Megakaryocytes/platelets, monocytes/macrophages, lymphocytes, fibroblasts, many other cell types	B cells, osteoblasts, fibroblasts, endothelial cells, NK cells, others	Induces isotype shift in B cells, activates osteoblasts, inhibits endothelial cell proliferation, inhibits growth of NK cells, T and B cells
PDGF	Platelets, endothelial cells, and other cell types	Fibroblasts, smooth muscle cells, NK cells, endothelial cells, epithelial cells	Activates fibroblasts to synthesize collagen, vasoconstrictor for smooth muscle cells, inhibits NK cells, stimulates granule release from PMNs and monocytes
MCAF	Monocytes/macrophage, fibroblasts, neoplastic cells and other cell types	Monocytes/macrophages	Regulates expression of adhesion molecules on macrophages, chemotactic for monocytes
SCF	Bone marrow stromal cells, endothelial cells, fibroblasts	Multipotential progenitor cells	Activation and differentiation of target
MIF	T cells	Monocytes/macrophages	Activates targets, inhibits migration

IL, interleukin; NK, natural killer; PMN, polymorphonuclear neutrophils; NKCF, natural killer cytotoxic factor; IFN, interferon; MHC, major histocompatibility complex; TNF, tumor necrosis factor; RANTES, regulated on activation, normal T expressed and secreted; MIP, macrophage inflammatory protein; CSF, colony stimulating factor; TGF, transforming growth factor, PDGF, platelet-derived growth factor, MCAF, monocyte chemotactic and activating factor (also called MCP, monocyte chemotactic protein); SCF, stem cell factor; MIF, migration inhibitory factor.

monocytes/macrophages. Many other cell types produce interleukins including fibroblasts, endothelial cells, PMNs, NK cells, B cells, hepatocytes, and others. To date, 17 biologically potent signaling molecules have been designated as interleukins (27, 28): IL-1α, IL-1ß, IL-1ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12 (NKCF), IL-13, IL-14, IL-15. Several of these molecules share biologic activities.

The IL-1 family of cytokines includes IL-1α, IL-1ß, and IL-1 receptor antagonist (IL-1ra). Although IL-1α and IL-1ß have distinct amino-acid sequences, the two molecules are structurally related, interact with the same receptors, are not distinguishable based on their biologic properties, and are collectively referred to as IL-1. IL-1 induces fever, neutrophilia, anorexia, increased levels of hormones (insulin, ACTH, cortisol), and synthesis of acute phase proteins by hepatocytes. The changes in endothelial cells induced by IL-1 are directly related to the pathologic lesions in vascular tissues that occur during inflammation. These changes include increasing the adhesiveness of the endothelial cells promoting leukocyte adherence to endothelial cells by inducing the expression of the cell adhesion molecules, ICAM-1, VCAM-1 and E-selectin. IL-1 is the proliferation-induction factor for many cells. IL-1ra binds to IL-1 receptors (IL-1R), but does not induce an intracellular response. Instead, it blocks the binding of IL-1α and IL-1ß, thereby inhibiting their proinflammatory effects.

There are two types of IL-1R, type I (CDw121a) that mediate most of the activities, and type II (CDw121b) that binds and inactivates IL-1 (effectively functioning as a decoy), decreasing the concentration of IL-1 available for activating type I receptors. IL-1R cannot distinguish between IL-1α and IL-1ß, but does not recognize other cytokines. Both type I and type II IL-1R are found on T cells and many other cell types. Soluble IL-1R (sIL-1R), types I and II, are found in relatively high levels (50 to 200 pM) in healthy individuals, with 2- to 3-fold increases in severe disease (29). The inhibitory effects of IL-1ra are decreased by sIL-1R, type I, but enhanced by sIL-1R, type II.

IL-2 is a growth factor for lymphocytes and NK cells. It is synthesized and secreted by T cells (Th0, Th1) that have been activated in the presence of antigen-presenting cells (macrophages) and an appropriate antigen. Once activated and producing IL-2, these T cells then develop IL-2 receptors (IL-2R) for the feedback loop which promotes proliferation of the T cells. IL-2 stimulates the proliferation of T, B, and NK cells, also activating NK cells to achieve their cytotoxic potential. IL-2 induces the synthesis by T cells of IFN, IL-4, IL-5, IL-6, and GM-CSF.

IL-2R exist as three classes of receptors distinguished by their ligand-binding affinities: IL-2Rα (CD25) a low affinity 55-kDa polypeptide chain; IL-2Rß (CD122) an intermediate affinity 75-kDa (p75) receptor; and γc, a high affinity receptor which is common to several other cytokine receptors (to date: IL-4, IL-7, IL-9, and IL-15) (30). Binding of IL-2 by all three receptor moieties is required to induce a proliferative signal in T cells. Activated T cells can release their IL-2R into solution, however, sIL-2R has low affinity for IL-2 and the presence of sIL-2R does not inhibit IL-2 reactions (31). The primary reason for measuring sIL-2R is as a marker for T-cell activation.

IL-3 is a potent hematopoietic growth factor, secreted by activated T cells. This cytokine promotes the development of granulocytes, macrophages, megakaryocytes, mast cells, erythroid cells, T cells, and B cells. It is not required for normal hemopoiesis, but sustains or amplifies the hemopoietic response during immune reactions. The principle action of IL-3 appears to be cellular ontogenesis before differentiation has taken place. IL-3R (CD123) is a heterodimer composed of an α-chain specific for IL-3R and a $ß_c$-chain common to GM-CSFR and IL-5R. IL-3R is expressed on a subset of T cells. To date, sIL-3R has not been reported.

IL-4 induces isotype shifts in B cells involved in IgE and IgG_4 production and inhibits the synthesis of IL-1α, IL-1ß and IFNγ by IL-2 activated NK cells. IL-4R is a member of the IL-2 receptor family, all of which share the γ_c chain. Two forms of IL-4R (CDw124) exist, classical and alternative. Both are heterodimers. The classical form consists of an IL-4R α-chain and the γ_c chain, predominantly found on hematopoietic cells, while the alternative form includes an IL-4R α-chain and an IL-13R α-chain is expressed on nonhematopoietic cells. sIL-4R has been found in biological fluids.

IL-5 promotes eosinophil differentiation and function. It also is involved in B cell responses and in bone marrow cell maturation and differentiation. IL-5R (CD125) is a heterodimer composed of an α-chain specific for IL-5R (predominantly expressed on eosinophils) and a $ß_c$-chain, common to GM-CSFR and IL-3R (more widely expressed). The α-subunit is required for ligand binding specificity while the ß-chain is responsible for the binding affinity. To date, sIL-5R has not been reported.

IL-6 is produced by many cell types including neoplastic cells. It stimulates hepatocytes to secrete acute phase proteins and also promotes the secretion of Ig by activated B cells. It induces the maturation of megakaryocytes and shortens the resting phase in hematopoietic progenitor cells. Plasmacytomas and plasma cells from patients with multiple myeloma produce and respond to IL-6; thus IL-6 may be an autocrine growth factor for malignant plasma cells. IL-6R (CD126), is a heterodimer with a ligand-specific α-chain, that shares the gp130 molecule as a common ß-chain transducing signal with IL-11R and several other molecules not discussed in this chapter (28). Circulating sIL-6R not only can bind IL-6 while in soluble form, but also can activate cells that express gp130 (but are devoid of IL-6R) that normally would not respond to IL-6 (28).

IL-7 promotes the growth of mature, activated T cells. IL-7R (CDw127) is a member of the IL-2 receptor family. In *in vitro* studies, bacterial invasion induced IL-7R expression in a colonic epithelia cell line, a process that was inhibited by cytochalasin D, a specific inhibitor of actin polymerization (32). IL-7/IL-7R interactions appear to be critical during early thymocyte development. To date, sIL-7R has not been reported.

IL-8, a chemokine (described further below) activates PMNs and is chemotactic for neutrophils, T cells, and basophils where it also stimulates release of histamine. IL-8R (CDw128) is α-chemokine receptors (CXCR). There are two IL-8R, IL-8RA (CXCR-1) and IL-8RB (CXCR-2) on the surface of PMNs. These IL-8R are decreased in several types of infections. To date, sIL-8R has not been reported.

IL-9 is an hematopoietic growth factor. IL-9R (CD not designated) is a member of the IL-2 receptor family. Three functions have been attributed to IL-9/IL-9R interaction: growth stimulation, protection against dexamethasone-induced apoptosis,

and cell proliferation of target cells. To date, sIL-9R has not been reported.

IL-10 regulates T cell and macrophage function and coactivates B cells for secretion of IgA, IgG_1, and IgG_3 and inhibits the production of IFN-γ by T cells, monocytes/macrophages, and NK cells. It suppresses NK cell functions and inhibits tumor cytotoxicity by human monocytes and alveolar macrophages. The IL-10 gene coding sequence is highly homologous to the Epstein-Barr virus (EBV) BCRF1 open reading frame, which might provide insight into viral tumorigenesis (28). IL-10R (CD not designated) is a member of the interferon receptor family (includes IFNα, IFNß, IFNγ, and IL-10). Ligand binding and activation of IL-10R results in the phosphorylation of signal transducer and activator of transcription 1 (STAT1) and STAT3 proteins. To date, sIL-10R has not been reported.

IL-11 supports the proliferation of hematopoietic cells much the same as IL-6. IL-11R (CD130) is a member of the IL-6R family, which shares the gp130 molecule as the common ß chain transducing signal. Soluble IL-11R/IL-11 complexes are capable of mediating an IL-6-type signal in cells that are devoid of IL-6R.

IL-12 (NKCF) is a cytotoxin. IL-12R (CD not designated) is composed of two subunits, designated ß1 and ß2. Mutations in the ß1 subunit have been associated with recurrent infections with intracellular bacteria. To date, sIL-12R has not been reported.

IL-13 functions similarly to IL-4. IL-13R (CD not designated) is a member of the IL-2R family with similar functions to IL-4R. Blockade of IL-13 activities by sIL-13R resulted in the complete reversal of allergen-induced airway hyperresponsiveness to a variety of stimuli and may have other effects on the immunopathogenesis of allergic asthma.

IL-14 is a hematopoietic growth factor. IL-14R (CD not designated) has not been characterized, to date.

IL-15 has similar functions to IL-2. IL-15R (CD not designated) is a heterotrimeric complex composed of IL-2Rß and γ chains in combination with a unique α-chain. IL-15 can bind to components of the IL-2R even though it has no sequence homology with IL-2, however, IL-15 can substitute for IL-2 under most conditions. To date, sIL-15R has not been reported.

Interferons and Interferon Receptors

Interferons originally were described as antiviral substances, but are now known to be important regulators of the immune response in general (33, 34). IFN-γ (type II IFN), in conjunction with TNF and in the presence of bacteria or virus, activates macrophages stimulating the production of IL-12 and IFN-α and -ß (type I IFNs). Type I IFNs, by increasing the expression of MHC class I antigens, enhance NK cell activity to virus-infected cells.

IFN-R have been identified for the two types of IFN; type I (α, ß, ω) and type II γ. Type I IFN-R consists of two chains, IFN-αR1, a full chain, and IFN-αR2 exists in soluble, short and long forms. Type II IFN-R consists of two transmembrane chains, IFN-γR1 that binds the IFN-γ ligand and IFN-γR2 that is required for signal transduction. In both type I and type II IFN-R, ligand binding activates the Janus family kinases (Jak1 and Jak2), which phosphorylate the R1 chain to serve as the recruitment site for STAT (signal transducers and activators of transcription) (35). Soluble type II IFN-R has been described.

Cytotoxins and Receptors

TNF activity was first described as a monokine that could cause tumor cell death (necrosis). Subsequently, TNF was shown to be homologous to cachectin, the mediator of cachexia accompanying parasitic infections (36). Some biologic properties of TNF are quite similar to those of IL-1; i.e., pyrogenic, induces shock, stimulates hepatocytes to synthesize acute phase proteins, induces prostaglandin E2 and collagen synthesis. Despite the similarities, TNF and IL-1 are distinct and receptor binding to each ligand is only displaced by the specific cytokine. The beneficial effects of TNF are most relevant as protective functions in parasitic, certain fungal and viral infections. The detrimental effects of TNF/cachectin include induction of cachexia, a state of anorexia and wasting, which occurs over a period of months.

TNF is known to be a major cytokine involved in inflammatory reactions. IL-1 and TNF are induced during inflammatory reactions where they, in turn, increase the expression of adhesion molecules on endothelial cells. Vascular permeability is increased, facilitating cell and plasma protein extravasation. IL-1 and TNF then are directly involved in tissue damage, degrading cartilage, facilitating proteolysis of muscle cells, destroying bone at the site of the local inflammatory reaction (27). TNF-α and IL-1 exhibit synergistic effects when used together. TNF works in conjunction with IFN to promote antiviral activity.

Natural killer cytotoxic factor (NKCF) has now been designated as IL-12. This cytotoxin activates NK cells and LAK cells to enhance their lytic capacity. It is released from LGL after LGL bind to their susceptible target cells. NKCF then binds to the target cell, leading to target cell death.

Lymphotoxin is 27% to 32% homologous to TNF-ß, the two substances compete for the same cellular receptor, and are functionally indistinguishable (37). Lymphotoxin, TNF-ß, and IL-1 all can stimulate bone resorption. Lymphotoxin can be released from myeloma cells and may be involved in the hypercalcemia and osteolytic lesions associated with multiple myeloma.

There are many members of the TNF-R family, including TNF-R1, TNF-R2, CD27, CD30, CD40, Fas, and others. These receptors all have cysteine rich pseudorepeats, each containing about six cysteines and 40 amino acids. Two distinct TNF receptors, TNF-R1 (the primary TNF receptor on most cells) and R2 are representative of the TNF-R family. TNF-R1 is responsible for signaling cytotoxicity and induction of several genes, while TNF-R2 is capable of signaling proliferation of primary thymocytes, significantly reducing the TNF concentration required for cell killing by regulating the rate of TNF association with TNF-R1. TNF receptor triggering results in three types of biological responses: necrosis, apoptosis (to be discussed later in this chapter), or proliferation. The TNF receptor family contains death receptors (TNF-R1, Fas, DR3, DR4, DR5) and a cytoplasmic death domain, capable of inducing apoptosis. Various biologic fluids (e.g., urine, serum from normal humans, septic, cancer and autoimmune patients) contain sTNF-R, particularly during active phases of the diseases. The effects of sTNF-R appear to be modulatory rather than inhibitory (31).

Chemokines and Receptors

Chemokines are very small (8 to 10 kd) chemoattractant proteins with 20% to 70% amino-acid sequence homology. There are at least four families of chemokines, only two of which have been extensively studied to date: α-chemokines that have one amino acid separating the first two cysteine residues (CXC), and ß-chemokines where the two cysteine molecules are adjacent (CC) (38). IL-8 is an α-chemokine. Monocyte chemoattractant protein (MCP), RANTES (regulated on activation, normal T expressed and secreted), and macrophage inflammatory protein-1 (α and ß) are ß-chemokines. For a complete discussion of chemokines, see Reference 38.

Four human CXC chemokine receptors (CXCR1 – CXCR4), eight human CC chemokine receptors (CCR1 – CCR8) have been identified. CCR1 and CCR2 are constitutively expressed on monocytes, but CCR2 can be downregulated by lipopolysaccharide, making the cells unresponsive to MCP-1 (38). The chemokine receptor, DARC (Duffy antigen receptor for chemokines) found on erythrocytes and endothelial cells, can bind both CXC and CC chemokines and may function as a sink for chemokines, clearing them from circulation.

Growth and Stimulatory Factors and Receptors

Granulocyte/monocyte colony stimulating factor (GM-CSF) and IL-3 stimulate progenitor cells to produce hematopoietic cells of multiple lineages. Once IL-3 and GM-CSF stimulate the microenvironment, stroma cells and macrophages produce the other two CSFs (G-CSF and M-CSF) to support self-renewal and survival of stem cells. GM-CSF-R (CDw116) is a member of the IL-3, IL-5, GM-CSF receptor family. G-CSF-R is a member of the IL-6 receptor family.

Transforming growth factor (TGF-ß) induces an isotype shift in B cells in conjunction with IL-10. It also inhibits the growth of NK cells, T cells, and B cells. Platelet-derived growth factor (PDGF) is mitogenic for many cell types; i.e., it stimulates fibroblast collagen synthesis, causes smooth muscle cells to vasoconstrict. It also stimulates the release of granules from PMNs and monocytes. Monocyte chemotactic and activating factor (MCAF) regulates the expression of CAMs on macrophages. Stem cell factor (SCF) activates and differentiates multipotential stem cells.

TGF-ß transduces signals through two different serine/threonine kinase receptors, TB-R1 and TB-R2. TB-R2 is a primary binding protein for the ligands and TB-R1 is an effector protein that determines the specificity of the signals (39). TBR have been found to act as tumor suppressor genes in various tumors.

Cytokine Antibodies

Natural antibodies to cytokines in human sera have been reported to date against TNF-α, IL-1α, IL-1ß, IL-2, IL-6, IL-8, IL-10, IFN-α, IFN-ß, IFN-γ, GM-CSF, chemokines, and nerve growth factor (NGF). These autoantibodies have been shown to bind to naturally occurring cytokines as well as recombinant cytokines. High avidity antibodies might interfere with the functions of their corresponding cytokines. The following cautions have been identified: (i) cytokine autoantibodies can interfere with the quantitation of cytokines in biologic fluids, (ii) existing autoantibodies could neutralize cytokine therapy, (iii) monitoring of certain immunoinflammatory diseases by quantitation of specific cytokines might be efficacious, and (iv) some pooled human IgG preparations contain high avidity cytokine autoantibodies (40). No differences in the levels of cytokine autoantibodies were detectable in studies comparing healthy individuals and RA patients (41).

Cytokines in Disease

Although mouse models (where knockout or transgenic mice can be used to delete specific cytokine gene or cytokine receptor genes), and *in vitro* studies have been useful in identifying the role of cytokines in immune regulation and pathologic states, the role of cytokines in human disease is much more complicated. Inflammatory diseases (infections, sepsis, autoimmune diseases, allergic respiratory disease [ARD]) have been the target of most studies regarding the effects of cytokines in humans. IFN-α2 and -ω have been shown *in vitro* to inhibit HIV replication. *In vivo* high serum levels of IFN appear to decrease viral burden during the acute stage of HIV infection, but late in the infection, these cytokines and sTNF-R are associated with disease progression. TNF, IL-1, IL-6, and IFN levels have been quantitated in HIV infections and inconsistently shown to be elevated (42).

Sepsis results from excessive systemic host inflammatory response to infection, largely mediated by cytokines. Cytokines that are elevated in human sepsis are TNF-α, IL-1ß, IL-6, and IL-8. Human trials of cytokines in sepsis (using anti-TNF-α or IL-1ra to block the effects of their respective cytokines) have been disappointing (43). The timing of administration and dose may be critical and once the cytokine cascade has been activated, attempts to intervene may be insufficient to improve outcomes. There is always the danger that blocking cytokine production may inhibit other critical host defense mechanisms.

Inflammation, associated with abnormal cytokine production, is a prominent feature of several autoimmune diseases, e.g., joint disease in patients with RA, destruction of pancreatic islet cells in insulin-dependent diabetes, altered intestinal mucosa in inflammatory bowel disease, and myelin sheath destruction in multiple sclerosis. IL-1 and TNF-α induce cartilage degeneration promoting synovial inflammation in RA (44). Neutralizing the effects of TNF-α with monoclonal antibodies has shown some therapeutic promise, but the deleterious effects of an antibody response to the mouse protein complicates the process. Neutralizing the effects of TNF-α with sTNF-α-R, a natural inhibitor of the cytokine, has reduced inflammation or prolonged survival. Use of IL-1ra to block IL-1 activities in several disease states (sepsis and graft vs. host disease) showed a dose-dependent decrease in IL-6 and decreased mortality (29). When IL-1ra was given to RA patients, there was a dose-dependent reduction in the number of swollen joints and new bone erosions and a decrease in CRP and ESR. Corresponding studies using sIL-1R did not show beneficial effects. IL-4, IL-10, and IL-13 function as antiinflammatory molecules. Administration of IL-10, in RA is in process (44). IL-4 and IL-13 are other candidates for treatment in RA to downregulate the synthesis of proinflammatory cytokines, however, the safety and efficacy has not yet been determined.

ARD has increased in prevelance and severity in recent years. In the mouse model, IL-13 plays a critical role in the development of both allergic and nonallergic asthma. Blocking IL-13 effects with recombinant IL-13 reversed AHD in mice and resulting in a decrease in total serum IgE (45). These findings suggest a possible therapeutic approach for treating asthma in humans.

There is increasing evidence that at least some of the beneficial effects of administering intravenous immunoglobulin (IVIG) to patients with immunoinflammatory conditions might be from naturally occurring cytokine antibodies (27, 41). Cytokine antibodies do not appear to be pathogenic since they are frequently present in healthy humans. They are IgG molecules, and as such, may be responsible for IgG FcR-mediated activation of killer cells and potentiate tumor cell destruction (41).

It is likely that chemokines play a role in inflammatory diseases, causing the accumulation and activation of leukocytes in tissues. Chemokine receptors are thought to play a role in HIV infection because high-risk individuals who have a polymorphic variant of CCR5, when exposed to HIV-1, remain uninfected (38). Additionally, cells from individuals who are homozygous for this variant cannot be infected with HIV-1 *in vitro*, while cells from individuals heterozygous for the variant have a much slower rate of progression of HIV-1 infection than other individuals. Several herpesviruses express chemokine-receptor homologues, thought to be involved in infectious processes (38).

It is possible that normal cytokine balances are disrupted in some human inflammatory diseases, but the circumstances have not been definitively proven. Likewise, the use of cytokines as therapeutic agents is an area currently under intense exploration. Cytokines and cytokine receptors still hold promise of therapeutic potential.

Lipid Mediators

Arachidonic acid (AA), the precursor of eicosanoids, is liberated from membrane phospholipids (phospholipases A_2 [PLA_2]) during inflammatory reactions. AA is metabolized to prostaglandins (PG) and thromboxanes (TXA) by the cyclooxygenase (COX) pathway and leukotrienes (LT) by the 5-lipoxygenase pathway (Fig. 63.7) (46). PGE_2 and PGI_2 are responsible for edema and vasodilation that contribute to the pain experienced at the site of inflammation. LTs increase local permeability and are chemoattractants for PMNs, which contributes to the phagocytic cell infiltration at the site of inflammation. TXA_2 and $PGF_{2\alpha}$ function as vaso- and bronchoconstricting agents. Platelet activating factor (PAF) contributes to this inflammatory activity by aggregating platelets, activating leukocytes, promoting AA metabolism, increasing microvascular permeability and decreasing blood pressure. PAF is also a potent mediator of shock.

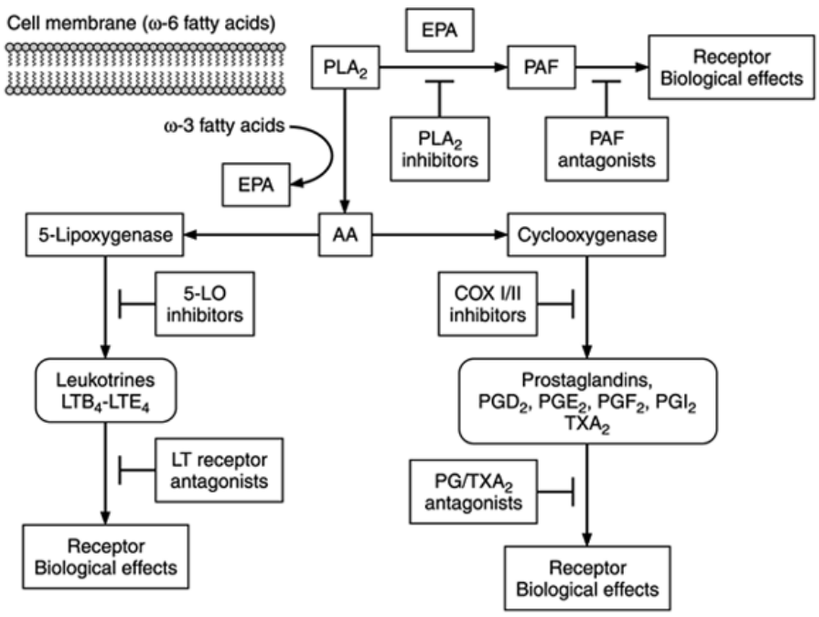


FIGURE 63.7. Lipid mediator pathways. Reproduced with permission from Heller A, Koch T, Schmeck J, et al. Lipid mediators in inflammatory disorders. *Drugs* 1998;55:487–496.

Lipid Mediator Inhibitors

Two forms of COX pathways have been elucidated: the constitutive COX-1, and the cytokine-induced COX-2. Substances that inhibit the COX-1 or -2 pathways include nonsteroidal antiinflammatory drugs (NSAIDs) that inhibit the conversion of AA to prostaglandins and thromboxanes (Fig. 63.7). COX inhibitors are very effective systemic antiinflammatory agents, but COX-1 inhibitors tend to have more gastrointestinal toxicity than do COX-2 inhibitors (46). COX-2 inhibitors also have

been shown to suppress the growth of colorectal cancer. In ARD, supplementing COX-2 inhibitors with the aerosolized vasodilator PGI_2 controls inflammation and the symptoms of ARD.

LT receptor antagonists or 5-lipoxygenase pathway inhibitors have been effective in controlling inflammation and obstructive airway disease in asthma. Synthetic PAF receptor antagonists have been used in septic shock, rheumatoid arthritis, and asthma without significant benefit. Interestingly, dietary supplementation with ω-3 fatty acids has shown beneficial effects in a variety of inflammatory diseases. It is thought that the eicosapentaenoic acid (EPA) released from ω-3 fatty acids competes with AA resulting in decreased production of lipid mediators of inflammation (46).

Immune Complexes

Immune complexes are formed whenever antigen and antibody combine. In normal individuals, immune complexes are effectively and efficiently removed by PMNs and phagocytic mononuclear cells. The dynamics of immune complex formation are continuously modified relative to the concentration of antigen or antibody present. The size of the immune complexes in circulation is determined by the relative ratio of antigen to antibody and the valence of the antigen (large antigens have more antigenic determinants) and the antibody (IgG has two combining sites, IgM has 10). Antigen excess produces very small circulating immune complexes. Antibody excess produces very large immune complexes that are efficiently removed by the phagocytic system. When antigen and antibody are present in more balanced ratios (equivalence), they localize in vessels and mesangial areas. Complement is activated while the immune complexes are in circulation, depositing C3b on their surface, which facilitates localization or deposition in blood vessels and renal glomeruli. C3b deposited on basement membranes provides the “activating surface” necessary to initiate the activation loop of the alternative complement pathway. Complement activation releases chemotactic factors that attract PMNs. During phagocytosis, the PMNs extrude the contents of their granules, which are cytotoxic for endothelial cells. Thus immune complexes initiate as well as perpetuate the inflammatory process. Immune complex diseases will be discussed in Chapter 65.

INFLAMMATION

The “cardinal signs” of inflammation include redness, swelling, heat, pain, and loss of function. Inflammation represents an orderly series of events which protect the host by destroying foreign invaders, eliminating the debris, and repairing any damage to host tissue. Cytokines, cytokine receptors, and lipid mediators are all involved in the inflammatory process. IL-1 and TNF function synergistically as proinflammatory molecules involved in the initiation and progression of inflammation, locally and systemically. Locally, at the site of production, IL-1 and TNF activation stimulates hydrolysis of phosphatidylchone, ultimately releasing AA precursors of lipid-derived mediators. IL-1 and TNF also induce chemokines (IL-8). COX inhibitors reduce the IL-1 and TNF-induced production of PGE_2. Corticosteroids inhibit the transcription of IL-1, TNF, and most cytokines.

The inflammatory response is generally self-limiting, primarily because of a negative feedback biological process via antiinflammatory cytokines. IL-4, IL-10, IL-13, and TGFß suppress the synthesis of IL-1, TNF, and other cytokines and increase the production of IL-1ra that competes with IL-1 for IL-1R.

The four stages of inflammation are: increased vascular permeability, emigration of neutrophils, emigration of mononuclear cells, and cellular proliferation.

Vascular Permeability

The microcirculation (capillaries, arterioles, and venules) is involved in the vascular phase of the inflammatory response. When an injury occurs, blood rushes to the affected area (hyperemia), facilitated by localized dilatation of capillaries. The dilatation is a direct result of the effects of histamine released from mast cells in the skin that has been damaged. Histamine and kinins also increase vascular permeability by causing the smooth muscle of the endothelial cells to contract, creating gaps between the cells through which fluids (transudates) and cells (diapedetic PMNs) can pass. As transudation continues, the blood flow slows or stops completely (stasis) caused by hemoconcentration when fluids are lost through transudation, leaving red blood cells clumped. In contrast to thrombosis, stasis is reversible; when blood flow is restored, the clumps of red cells disperse. If the inflammation is severe, microthrombi form and irreversible platelet aggregation occurs.

Emigration of Granulocytes

PMN migration from the peripheral blood to inflamed tissues involves three steps: (i) selectin-mediated interactions with endothelial cells, (ii) integrin-dependent adhesion to endothelial cell surfaces expressing CAMs, and (iii) chemokine-regulated migration of leukocytes across the vessel wall (47). Damage to the endothelial cells causes changes in the surface membranes that facilitate adherence of granulocytes, primarily PMNs. Chemokines and C5a attract additional PMNs until the entire area becomes pavemented by leukocytes. Emigration of PMNs into the area of inflammation results in active phagocytosis of microorganisms and any other foreign material. With phagocytosis, degranulation occurs. The contents of these granules are directly toxic to surrounding tissues.

Emigration of Mononuclear Phagocytes

The third stage of the inflammatory response is migration of mononuclear phagocytes into the affected area. These activated macrophages begin to synthesize cytokines, particularly IL-1 and IL-6, which are responsible for further inflammatory reactions such as fever, synthesis of acute phase proteins, and attraction of T and B lymphocytes to the area of inflammation.

Cellular Proliferation and Repair

IL-1 stimulates IL-2 synthesis and IL-2R generation. Interaction of IL-2 with IL-2R results in proliferation of lymphocytes and fibroblasts, which begins within 18 hours and peaks by 72 hours. During proliferation, fibroblasts produce acidic mucopolysaccharides

to neutralize the effects of chemical mediators released by damaged mast cells and basophils. Resolution and repair are the final stages of the normal inflammatory process.

Apoptosis And The Immune System

The term apoptosis is derived from the Greek word that means “a falling off” such as leaves fall off of trees (48). This derivation is applied to the first definition “fragmentation of a cell into membrane-bound particles that are then eliminated by phagocytosis.” The second definition “programmed cell death,” a tightly regulated event in normal development and homeostatis, is the more common medical usage of the term. Apoptosis is a physiologic process in all multicellular organisms, and in animals is involved in the development, regulation, and functioning of the immune system. When the process of programmed cell death malfunctions, immunodeficiency disease, autoimmune diseases, or lymphoid malignancies can result.

Stages of Apoptosis

Programmed, physiologic cell death has been described in four stages: (i) the stimulus that provokes the apoptotic event, (ii) detection and transduction of the signal that activates the processes involved in apoptosis, (iii) the effector phase that involves activation of specific proteases, and (iv) the postmortem stage where the cellular chromatin condenses and its DNA is degraded (49, 50). The signaling process may be initiated through surface receptors (various members of the TNF-R family, (e.g., CD95), TNF-R1, TNF-R2, CD30, CD40) or may originate intracellularly as a result of action of a drug, toxin, or radiation damage. Once the signal is detected, the mechanisms are activated and this message is sent to the apoptotic effector machinery. The key components of the effector phase are the cysteine proteases (IL-1ß converting enzyme or ICE, CPP32, and others) that are activated during apoptosis. There are also regulatory mechanisms for the apoptotic process, the most studied of which is bcl-2. The expression of bcl-2 inhibits apoptosis in many cell types by inhibiting activation of the cysteine proteases. In the final postmortem stage, the apoptotic cells are efficiently phagocytosed by macrophages and other inflammatory cells.

Developmental Apoptosis

During development, the majority of T and B lymphocytes (90% to 95%) in the thymus and the bone marrow die because they fail to receive signals to survive, the apoptotic pathway is activated, and the cells die by neglect (50). In the thymus, cells that bear the T-cell receptors (TCR) and recognize self major histocompatability (MHC) antigens are positively selected while a subset of cells that express high-affinity MHC are negatively selected (Fig. 63.8) (51). Less currently is known about the development of B cells in the bone marrow, but it is thought that surface immunoglobulin (sIgM and sIgD) functions similarly to the TCR on T cells, whereby self-reactive B cells are clonally deleted by apoptosis. Current theory proposes that certain autoimmune diseases are the result of failure in programmed cell death of self-reactive T and/or B cells.

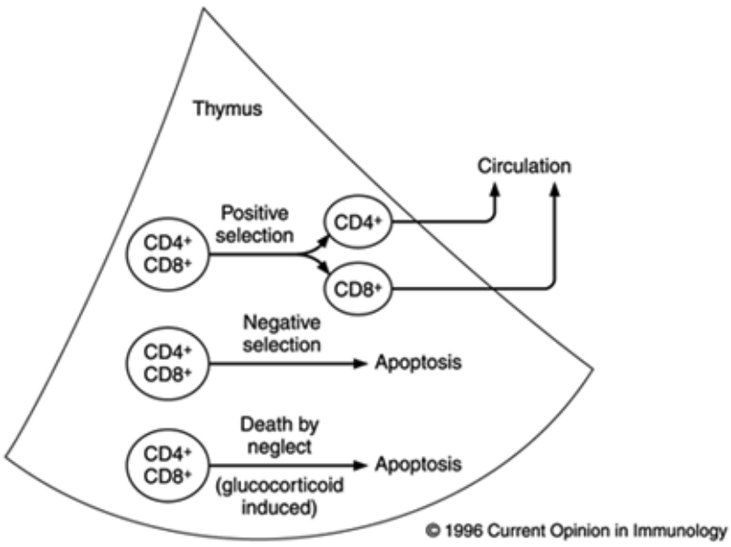


FIGURE 63.8. Developmental apoptosis. Reproduced with permission from Osborne BA. Apoptosis and the maintenance of homeostatis in the immune system. *Curr Opin Immunol* 1996;8:245–254.

Homeostasis of Peripheral Lymphocytes

Lymphocytes that develop in the primary lymphoid organs traffic to the secondary lymphoid organs where they encounter a multitude of antigens and accessory signals (e.g., cytokines) and become activated. The activated lymphocytes proliferate and differentiate, producing antibodies (B cells) or secreting cytokines or killing target cells (T cells). Cells that are no longer needed are removed by activation-induced cell death (ACID) (50, 51). ACID appears to be mediated by binding of its associated ligand to receptors that are members of the TNF-R family, especially CD95 (Fas/APO-1). It is thought that cells that generate low affinity antibodies or antibody directed to self-antigens are removed by apoptosis in normal humans. Cytotoxic T lymphocytes (CTL) induce apoptosis in target cells by two distinct pathways:

primarily through the action of granzyme B and perforin, or secondarily through a CD95/CD95L (receptor-ligand interaction) pathway.

Role of Apoptosis in Diseases of the Immune System

Certain body sites that are protected from immune interaction (immune privilege), e.g., the eye and the testis, have been shown to express high levels of CD95L, believed to contribute to preventing inflammation in these delicate tissues (51). Experiments of nature illustrate the effects of defective apoptosis: the rare autoimmune lymphoproliferative disease where mutations in CD95 inhibit apoptosis, X-linked severe combined immune deficiency (SCID) where a defect in IL-2R results in increased apoptosis of the lymphocyte population, and Chediak-Higashi syndrome where CTLs are unable to degranulate and destroy target cells.

Apoptosis is a defense mechanism against viral-infected cells. When a cell detects a virus and activates the intrinsic apoptotic mechanisms or when a CTL recognizes a virally infected cell, the cytopathic effects are defensive apoptosis. The death of CD4+ T cells in HIV is known to occur by apoptosis. Uninfected, bystander cells also can be eliminated by apoptotic mechanisms. Fulminant hepatic necrosis following infection with hepatitis B or hepatitis C is an example of apoptosis gone awry. It is hoped that apoptotic protease inhibitors might one day combat graft rejection by blocking CTL killing, prevent CD4 cell death in AIDS, or minimize hepatic cell damage in viral hepatitis. Bcl-2 antagonists might be used to treat follicular lymphoma or eliminate lymphocytes mediating autoimmune diseases. Radiation, certain chemotherapeutic agents, and steroids currently used for immunosuppression and/or tumor depletion, exert their effects by activating apoptosis.

REFERENCES

1. Arnaout MA. Structural diversity of cell adhesion molecules and their role in inflammation. In: Kelley WN, Ruddy S, Harris ED, et al. eds., *Textbook of Rheumatology*, St. Louis, MO: W.B. Saunders, 1997;303–321.

2. McMurray RW. Adhesion molecules in autoimmune disease. *Sem Arthritis Rheum* 1996;25:215–233.

3. Venge P, Bergstrand H, Hakansson L. Neutrophils and eosinophils. In: Kelley WN, Ruddy S, Harris ED, et al., eds. *Textbook of Rheumatology*, 1997;146–160.

4. Schwartz LB. The mast cell. In: Kelley WN, Ruddy S, Harris ED, et al. eds. *Textbook of Rheumatology* St. Louis, MO: W.B. Saunders, 1997;161–175.

5. Carr ME. Platelets. In: Kelley WN, Ruddy S, Harris ED, et al. eds. *Textbook of Rheumatology* St. Louis, MO: W.B. Saunders, 1997;pp 176–181.

6. Thomas R, Wong R, Lipsky PE. Monocytes and macrophages. In: Kelley WN, Ruddy S, Harris ED, et al. eds. *Textbook of Rheumatology*, St. Louis, MO: W.B. Saunders, 1997;128–145.

7. Moretta A, Parolini S, Castriconi R, et al. Function and specificity of human natural killer cell receptors. *Eur J Immunogen* 1997;24:455–468.

8. Rolstad B, Seaman WE. Natural killer cells and recognition of MHC Class I molecules: new perspectives and challenges in immunology. *Scand J Immunol* 1998;47:412–425.

9. Cotran RS, Briscoe DM. Endothelial cells in inflammation. In: Kelley WN, Rudd S, Harris ED, et al., eds. *Textbook of Rheumatology*. St. Louis, MO: W.B. Saunders, 1997;183–198.

10. Epstein FH. Atherosclerosis—an inflammatory disease. *New Engl J Med* 1999;340:115–126.

11. Tillett W, Francis T. Serological reactions in pneumonia with nonprotein somatic fraction of pneumococcus. *J Exp Med* 1930;52:561–571.

12. Suankratay C, Mold C, Zhang Y, et al. Complement regulation in innate immunity and the acute-phase response: Inhibition of mannan-binding lectin-initiated complement cytolysis by C-reactive protein (CRP). *Clin Exp Immunol* 1998;113:353–359.

13. James K, Baum L, Adamowski C, et al. C-reactive protein antigenicity on the surface of human lymphocytes. *J Immunol* 1983;131:2930–2934.

14. Deodhar SD, James K, Chiang T, et al. Inhibition of lung metastases in mice bearing a malignant fibrosarcoma by treatment with liposomes containing human C-reactive protein. *Cancer Res* 1982;42:5084–5088.

15. Fischer CL, Gill C, Forrester MG, et al. Quantitation of “acute phase proteins” postoperatively: Value in detection and monitoring of complications. *Am J Clin Pathol* 1976;66:840–846.

16. Ridker PM, Cushman M, Stampfer MJ, et al. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N Engl J Med* 1997;336:973–979.

17. Haverkate F, Thompson SG, Pyke SDM, et al. Production of C-reactive protein and risk of coronary events in stable and unstable angina. *Lancet* 1997;349:462–466.

18. Jandl JH. Blood and bloodforming tissues. In: *Blood* Boston: Little, Brown and Company 1987:1–48.

19. James K. Complement: activation, consequences, and control. *Am J Med Technol* 1982;48:735–742.

20. Matsushita M, Fujita T. Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. *J Exp Med* 1992;176:1497–1502.

21. Pangburn MK, Schreiber RD, Muller-Eberhard HJ. Formation of the initial C3 convertase of the alternative complement pathway: acquisition of the C3b-like activities of spontaneous hydrolysis of the putative thioester in native C3. *J Exp Med* 1981;154:856–867.

22. Giclas PC. A glossary of complement nomenclature. *AMLInteractions* 1997;2:8–10; 3:10–11.

23. Davis AE, Ziegler JB, et al. Heterogeneity of nephritic factor and its identification as an immunoglobulin. *Proc Natl Acad Sci USA* 1977;74:3980–3983.

24. Inai S, Kitamura H, Fujita T, et al. Differences between plasma and serum components in patients with chronic liver disease. *Clin Exp Immunol* 1976;25:403–409.

25. Gonda TJ, D'Andrea RJ. Activating mutations in cytokine receptors: Implications for receptor function and role in disease. *Blood* 1997;89:355–369.

26. Remick DG. Applied molecular biology of sepsis. *J Critical Care* 1995;10:198–212.

27. Bendtzen K, Hansen MB, Ross C, et al. Cytokines and autoantibodies to cytokines. *Stem Cells* 1995;13:206–222.

28. Fridman WH, Tartour E. Cytokines and cell regulation. *Mol Aspects Med* 1997;18:3–90.

29. Dinarella CA. Role of pro- and antiinflammatoary cytokines during inflammation: Experimental and clinical findings. *J Biol Regulators & Homeostatic Agents* 1997; 11:91–103.

30. Karnitz LM, Abraham RT. Interleukin-2 receptor signaling mechanisms. *Adv Immunol* 1996;61:147–199.

31. Dayer JM, Arend WP. Cytokines and growth factors. In: Kelley WN, Ruddy S, Harris ED, et al., eds. *Textbook of Rheumatology*. St. Louis, MO: W.B. Saunders, 1997:267–285.

32. Yamada K, Shimaoka M, Nagayama K, et al. Bacterial invasion induces interleukin-7 receptor expression in colonic epithelial cell line, T84. *Eur J Immunol* 1997;3546–3460.

33. Belardelli F. Role of interferons and other cytokines in the regulation of the immune response. *AMPIS* 1995;103:161–179.

34. Ramshaw IA, Ramsay AJ, Karupiah G, et al. Cytokines and immunity to viral infections. *Immunol Reviews* 1997;159:119–135.

35. Pestka S. The interferon receptors. *Sem Oncology* 1997;24:S9-18–S9-40.

36. Beutler B, Cerami A. The common mediator of shock, cachexia and tumor necrosis. *Adv Immunol* 1988;42:213–231.

37. Ruddle NH. Tumor necrosis factor and related cytotoxins. *Immunol Today* 1987;8:129–130.

38. Luster AD. Mechanisms of disease: Chemokines — Chemotactic cytokines that mediate inflammation. *N Eng J Med* 1998;338:436–445.

39. Miyazono K. TGFß receptors and signal transduction. *Int J Hematol* 1997;65:97–104.

40. Bendtzen K, Hansen MB, Ross C, et al. High-avidity autoantibodies to cytokines. *Immunology Today* 1998;19:209–212.

41. Hansen MG. Human cytokine autoantibodies. Characteristics, test procedures and possible physiological and clinical significance. *APMIS* 1996;59(suppl):1–33.

42. Thea DM, Porat R, Nagimbi K, et al. Plasma cytokines, cytokine antagonists, and disease progression in African-American women infected with HIV-1. *Ann Intern Med* 1996;124:757–762.

43. Blackwell TS, Christman JW. Sepsis and cytokines: current status. *Br J Anaesth* 1996;77:110–117.

44. Isomäki P, Punnonen J. Pro- and anti-inflammatory cytokines in rheumatoid arthritis. *Ann Med* 1997;29:499–507.

45. Wills-Karp M, Luyimbazi J, Xu X, et al. Interleukin-13: Central mediator of allergic asthma. *Science* 1998;282:2258–2261.

46. Heller A, Koch T, Schmeck J, et al. Lipid mediators in inflammatory disorders. *Drugs* 1998;55:487–496.

47. Tetta C, Marianoo F, Buades J, et al. Relevance of platelet-activating factor in inflammation and sepsis: mechanisms and kinetics of removal of extracorporeal treatments. *Am J Kidney Dis* 1997;30:S57–S65.

48. *Dorland's Illustrated Medical Dictionary* 28th ed., St. Louis, MO: W.B. Saunders, 1994.

49. Vaux DL, Strasser A. The molecular biology of apoptosis. *Proc Natl Acad Sci* 1996;93:2239–2244.

50. Ekert PG, Vaux DL. Apoptosis and the immune system. *Brit Med Bull* 1997;53:591–603.

51. Osborne BA. Apoptosis and the maintenance of homeostasis in the immune system. *Curr Opin Immunol* 1996;8:245–254.

64

Monoclonal Gammopathies

David F. Keren

A monoclonal gammopathy is the protein product of a single clone of plasma cells. Although the term “monoclonal gammopathy” often has been used to designate the proteins produced by the malignant cells in multiple myeloma and other B-cell lymphoproliferative lesions, such as Waldenström's macroglobulinemia, a monoclonal gammopathy also may be the product of a “regulated” cell. Patients with profound infections or considerable polyclonal autoimmune stimulation often have marked expansion of one or a few (oligo-) clones. When one clone predominates, a picture resembling a monoclonal gammopathy from a B-cell lymphoproliferative lesion may result. Thus, the presence of a monoclonal protein (M-protein) alone does not indicate that a neoplastic process is present. Factors to consider when evaluating an M-protein include the quantity of the M-protein, the fluid in which it is found (serum versus urine), and the chemical type of the M-protein (light chains will more likely be found in the urine than in the serum). The term M-protein is used in the literature to describe the monoclonal gammopathy. Although it has been variously used for “monoclonal component, myeloma protein, and macroglobulin protein,” the recently published guidelines for evaluating monoclonal gammopathies recommend the use of the term “M-protein” (1,2 and 3).

The clinical investigator must be able to detect even small M-proteins. Formerly, small M-proteins were not thought to be of clinical significance. Indeed, as discussed below, most M-proteins less than 1,000 mgidl, are part of the monoclonal gammopathy of undetermined significance (MGUS) category. However, some of them represent the products of B-cell lymphoproliferative lesions such as chronic lymphocytic leukemia and well-differentiated lymphocytic lymphoma. Conditions often overlooked when a quantitatively small M-protein is detected include amyloidosis, peripheral neuropathies, and nonsecreting, or, more typically, hyposecreting myeloma.

This chapter reviews the new monoclonal gammopathy guidelines, and the major disease states that result in the production of a monoclonal gammopathy. The methodologies for detecting them are evaluated and a strategy for assessing M-proteins is given.

GUIDELINES FOR EVALUATING MONOCLONAL GAMMOPATHIES

Part of "64 - Monoclonal Gammopathies"

In May 1998, new National Consensus Guidelines for the Clinical and Laboratory Evaluation of Patients with Monoclonal Gammopathies were adopted (1). The Guidelines conference was cosponsored by The Clinical Center of the National Institutes of Health, The American Society of Hematology, The American College of Rheumatology, and The College of American Pathologists.

The conference focused on three questions:

1. Which patients should be evaluated for the presence of a monoclonal (M) protein?
2. What types of specimens should be studied, and how often should they be used to follow patients with M-proteins?
3. Which laboratory methods should be used to detect, identify, and follow M-proteins?

Which Patients Should be Evaluated for the Presence of an M-Protein?

M-proteins are found in the serum and/or urine of patients with a wide variety of clinical conditions. The occurrence and clinical symptoms of these conditions vary widely (Table 64.1). For instance, approximately 13,000 new cases of multiple myeloma are reported annually in the United States, whereas, the estimated occurrence of MGUS has been as high as 750,000. Patients with multiple myeloma have an average survival of 3 years, while individuals with MGUS have the average life expectancy for their age group (about 10 years). Clearly, the significance of M-proteins varies widely and the mere presence of an M-protein must be evaluated in the clinical context of that patient (2).

TABLE 64.1. CONDITIONS ASSOCIATED WITH THE PRESENCE OF AN M-PROTEIN

Condition	Clinical Features
Multiple myeloma	Back pain, osteolytic lesions, fatigue, elevated sedimentation rate, nephrotic syndrome, infections, anemia, elevated calcium
Waldenström's macroglobulinemia	Fatigue, elevated sedimentation rate, dizziness, anemia, hyperviscosity
B-Lymphoproliferative disorders	Fatigue, anemia, lymphadenopathy, lymphocytosis
Amyloid (AL), light chain deposition disease	Nephrotic syndrome, congestive heart failure, carpal tunnel syndrome
Neuropathy and M-protein	Peripheral sensory and/or motor neuropathy
Monoclonal gammopathy of undetermined significance (MGUS)	No symptoms related to the M-protein

TABLE 64.2. CONDITIONS WITH RELATIVELY SMALL QUANTITIES OF M-PROTEIN IN SERUM

Condition	Associated Feature
Light chain disease	15% of cases of multiple myeloma are light chain disease
Amyloid (AL)	20% have multiple myeloma or Waldenström's macroglobulinemia
Heavy chain disease	Alpha chain disease is rare in United States
	Gamma chain disease is associated with B-lymphoproliferative disorders
M-protein associated neuropathy	Must do Immunofixation of serum to rule out
Cryoglobulin type II	Associated with Hepatitis C
Solitary plasmacytoma	Tiny or no M-protein in serum
B-Lymphoproliferative disorder	Often with overall gamma suppression
MGUS	The most common cause of serum M-proteins

Traditionally, the size of the M-protein has been helpful in determining the significance for the patient. Typically, a large quantity of M-protein (greater than 2 G/dL) is more likely to be associated with myeloma, whereas, a small quantity of M-protein may be associated with MGUS. Yet, several important conditions associated with small quantities of M-proteins may be overlooked if these M-proteins are not recognized (Table 64.2). This is why it is important to use a logical sequence of testing and high quality laboratory methods to detect M-proteins.

Because of the high prevalence of MGUS, it is not useful to screen patients who do not have signs or symptoms consistent with the other conditions listed in Table 64.1 and Table 64.2 with serum and urine protein electrophoresis. In order for a positive result to have a significant predictive value for the individual patient, serum and urine protein evaluations must be performed selectively. The degree of clinical suspicion is also useful in determining how far to proceed with the testing process.

What Types of Specimens Should be Studied, and How Often Should They be Used to Follow Patients with M-Proteins?

The initial evaluation of a patient with clinical features suggesting the presence of an M-protein should include a serum and urine electrophoresis using high-resolution methods (Fig. 64.1). It is important to include both urine and serum in this initial evaluation, because in some patients the M-protein will only be located in one or the other fluid. Alternatively, in some cases, only a small amount of M-protein will be found in the serum while a massive amount of monoclonal free light chain (Bence Jones protein) will be present in the urine.

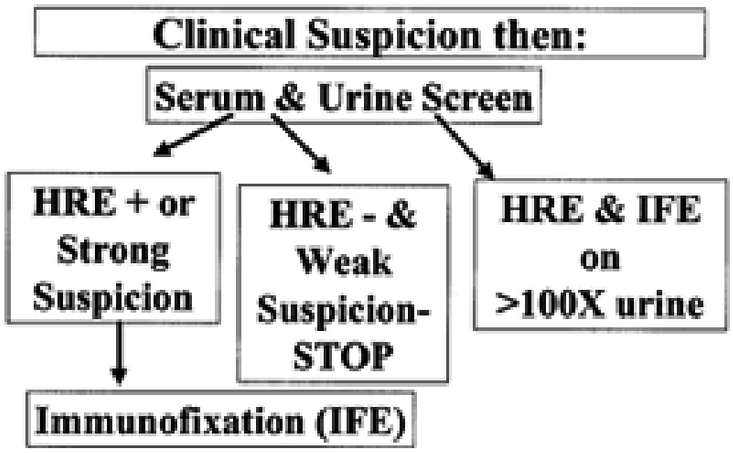


FIGURE 64.1. Specimens and evaluation for detecting monoclonal gammopathies.

If the serum protein electrophoresis is negative, and there is a low index of suspicion, one may wish to look for other causes of the symptoms before performing an immunofixation (IFE). However, if the clinical suspicion is high, an IFE should be performed even when no M-protein is seen on the screening electrophoresis. If an M-protein is found on the screening electrophoresis, an IFE must be performed to identify the M-protein. This identification serves several purposes. First, it rules out a false positive due to some other protein that may mimic an M-protein on serum protein electrophoresis (Table 64.3). Second, because IFE is more sensitive than screening electrophoresis, it may identify a second M-protein (often a monoclonal free light chain that accompanies the intact molecule in the serum). Third, it identifies the M-protein for comparison with subsequent samples from the patient. If there is no change in the migration of the identified M-protein on subsequent samples, there is no need to perform further IFE studies. Repeat IFE studies are only needed when the electrophoretic migration changes on subsequent samples because this may indicate the development of a second clone, or the appearance of monoclonal free light chains in addition to the intact M-protein.

TABLE 64.3. PROTEINS THAT MAY MIMIC AN M-PROTEIN

Fibrinogen
C3 Variant
Transferrin variant
Beta-1 lipoprotein
Antibiotic band on capillary zone electrophoresis

Although one may detect monoclonal free light chains (Bence Jones proteins) in the serum, they mainly are found in the urine. Urine for evaluating the presence of monoclonal free light chain should be a 24-hour sample. False negatives are most likely to occur on spot samples of urine. Therefore, a negative result on a spot urine sample cannot be used to rule out the presence of monoclonal free light chains.

All urine samples suspected of containing an M-protein must be concentrated at least 100X by the laboratory and must be tested by *both* electrophoresis and IFE. False negatives may occur with dipsticks, sulfosalicylic acid tests and the old heat-acid precipitation test. There is an *incorrect* general belief that the sulfosalicylic acid test is an adequate screen. The monoclonal gammopathy panel of experts was unanimous that it is not. All panel members reported seeing several cases of monoclonal free light chain missed by the sulfosalicylic acid method.

Once an M-protein has been identified, how should one follow it? This depends upon the clinical situation. Depending on the population of patients selected for M-protein evaluation, most cases will be MGUS. However, one must rule out all of the relevant conditions listed in Table 64.1 and Table 64.2 before classifying a patient as MGUS. For instance, congestive heart failure from

amyloid (AL) may be suspected because of a small quantity of lambda M-protein in the serum or urine. Once those conditions are ruled out, patients classified as MGUS need have serum and urine quantifications of the M-protein only annually, assuming no change in the clinical or laboratory features. Kyle has found that about 20% to 25% of patients with MGUS followed for 20 to 35 years will develop clinically significant B-lymphoproliferative disorders including myeloma, Waldenström's macroglobulinemia, B-cell leukemia, and B-cell lymphoma (3, 4).

Patients with multiple myeloma and Waldenström's macroglobulinemia who are receiving therapy need to be followed every 1 to 2 months with serum and urine electrophoresis. The laboratory needs to report the quantity of the M-protein as a guide to the effectiveness of ongoing therapy. It is wasteful to repeat the IFE on samples with no change in the migration of the M-protein by serum protein electrophoresis. A repeat IFE does not provide useful information. Indeed, repeating IFE on a serum with a previously characterized M-protein and no change in migration pattern is an unfortunate waste of time and money. The panel was unanimous that the clinician needs to know what has happened to the quantity of the M-protein since the last evaluation, not to reidentify a known molecule.

For serum, the best measurement of the M-protein is achieved by performing high-resolution electrophoresis and measuring the amount of protein below the spike (Fig. 64.2). Sometimes, when the M-protein migrates in the beta region, this may be a problem if the quantity of the M-protein is small and it comigrates with either transferrin or C3. In practice, however, most clinically significant M-proteins either migrate in the gamma region, or are large enough in quantity that the interference in measurement by transferrin or C3 is relatively trivial. For those few cases, however, quantification of the M-protein (usually IgA or IgM) will provide adequate follow-up (5).

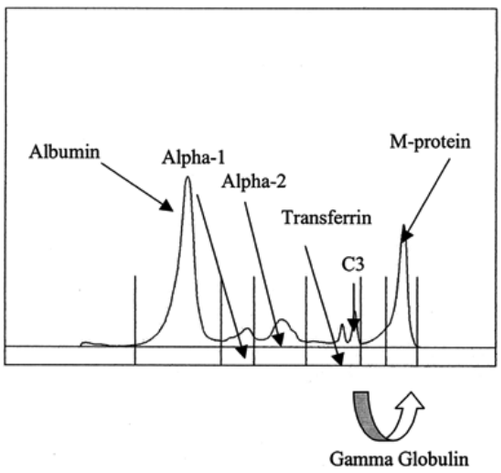


FIGURE 64.2. Capillary zone electropherogram with M-protein. M-proteins should be measured by integrating the area under the peak.

For the urine, it is important to follow the monoclonal free light chains and not the intact M-protein immunoglobulin that may be found in the urine. The prognosis and threat to the kidney is mainly due to the amount of monoclonal free light chain and not due to the intact M-protein.

The panel agreed that the term “Bence Jones protein” has been somewhat confusing in this regard. Bence Jones protein only refers to the monoclonal free light chains, not to the intact monoclonal immunoglobulins. Because of this, the consensus panel recommends use of the more specific term “monoclonal free light chain” instead of the older term “Bence Jones protein.” The panel recognized that because of the long use of the term “Bence Jones” protein, it would be a slow process to remove it. However, the confusion caused by mistaking intact M-proteins with monoclonal free light chains is of enough significance to warrant this effort.

Monoclonal free light chains may damage renal tubules, and are associated with amyloid (AL), as well as light-chain deposition disease. To follow the quantity of monoclonal free light chain in the urine, obtain a 24-hour sample of urine. The laboratory needs to perform electrophoresis and determine the amount of protein associated with the monoclonal free light chain peak. If there has been any change in migration of the monoclonal free light chain, the laboratory needs to perform IFE to determine if a new peak is another form of monoclonal free light chain (dimer, tetramer) or a new intact M-protein peak. Results are expressed as grams of monoclonal free light chain/24 hours.

Which Laboratory Methods should be used to Detect, Identify, and Follow M-Proteins?

The panel was unanimous that the laboratory should use a high-resolution electrophoretic method in the initial evaluation of serum and urine. A high-resolution method is one that permits a crisp separation of transferrin and C3 in the beta region (Fig. 64.3). This gives the interpreter the best view of subtle, but potentially clinically significant M-proteins (6, 7).

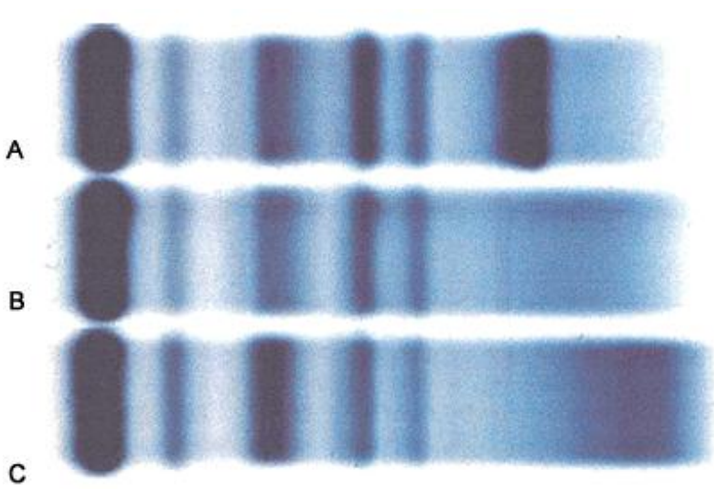


FIGURE 64.3. High-resolution electrophoresis with M-protein (top), normal (middle), and polyclonal patterns. The gel or CZE pattern should provide crisp separation of the beta-1 (transferrin) band from the beta-2 (C3) band.

There is a wide variation in the quality of electrophoretic gels available. The importance of using high-resolution electrophoresis methods was demonstrated on a national level in a College of American Pathologists (CAP) Survey Sample (8). Less than one third of laboratories using a low-resolution electrophoretic system could detect a subtle, but potentially clinically significant M-protein in the CAP Survey, while over 95% of laboratories using high-resolution methods were able to detect this M-protein.

In addition to gel-based methods, capillary zone electrophoresis also is available in a high-resolution format. The crisp separation of the beta region components by high-resolution capillary zone electrophoresis is evident in the examples shown in Fig. 64.2 and Fig. 64.3.

All urine must be evaluated by both high-resolution electrophoresis and by IFE on the initial evaluation for the presence of an M-protein. The urine must be concentrated at least 100X prior to study. It is important to do both tests initially to be certain that an obvious M-protein seen on the high-resolution electrophoresis of the urine is not missed because of a technical error, or to an antigen excess effect. IFE must be able to detect the presence of kappa and lambda monoclonal free light chains. When monoclonal free light chains are identified, they must be quantified on 24-hour samples to follow the patients.

The presence of a monoclonal free light chain in the urine is more ominous and uncommon than the presence of a small M-protein in the serum. This is because the tubules are capable of reabsorbing up to 1 gram of protein in 24 hours. Therefore, before the monoclonal free light chains are detected in the urine, they must be present in a relatively large amount to exceed the tubular reabsorptive capacity. Therefore, when one detects a monoclonal free light chain in the urine, one must be particularly mindful of the more subtle conditions such as amyloid (AL) and light chain deposition disease that may be associated with those M-proteins.

DISEASE STATES

Part of "64 - Monoclonal Gammopathies"

Myeloma

Multiple myeloma is the most common malignant cause for the presence of an M-protein in serum or urine. This disease is because of a malignant proliferation of a clone of B lymphocytes that manifests predominantly as plasma cells. Note that multiple myeloma often is considered to be a condition of malignant plasma cells. In fact, the malignancy involves the entire B-cell lineage but manifests mainly as malignant plasma cells. When one examines the bone marrow from individuals with multiple myeloma, pre-B cells with the same idiotype as the monoclonal protein can be found. Usually, patients with multiple myeloma present with involvement of the bone marrow. The plasma cells synthesize and secrete large quantities of the monoclonal whole immunoglobulin. These plasma cells are like inefficient factories because they often produce excessive numbers of monoclonal free light chains (Bence Jones protein). Patients with multiple myeloma present with a wide variety of clinical symptoms. Bone pain often is a presenting feature because the bone marrow involvement creates lytic lesions in the ribs, vertebrae, skull, and long bones. The plasma cells produce an osteoclast-activating factor that results in recruitment of osteoclasts that dissolve the bone in these locations. As a result of the destruction of the bony elements, these individuals may present with a so-called pathological fracture. These are fractures of bones resulting from relatively trivial trauma. Because the lytic lesions weaken the cortex, it is more susceptible to breakage (9,10 and 11).

Patients with multiple myeloma also present with infectious diseases. This results from a marked suppression of the production of polyclonal immunoglobulins in these individuals. When one examines the serum protein electrophoresis, one usually finds a large M-protein in the gamma or the beta region.

Note, however, that a spike may occur anywhere from the a region through the y region in monoclonal gammopathy. In multiple myeloma, the normal production of y-globulin is suppressed so that the background of normal antibodies is markedly decreased. As a result, many patients with multiple myeloma suffer from pyogenic bacterial infections that are normally defended by opsonizing immunoglobulins.

Other myeloma patients present with weakness resulting from anemia. The anemia tends to develop later in the course of the disease and is relatively nonspecific. However, examination of the peripheral blood from patients with multiple myeloma may show a blue hazy background due to the excessive amount of protein present in the serum. Other key laboratory parameters in these individuals include Rouleaux formation, an increase in the serum alkaline phosphatase, and calcium from the destruction of bone. Note that the elevated calcium usually appears relatively late in the course of this disease. Early on, one usually will not find evidence of hypercalcemia.

The excessive production of free light chains (Bence Jones protein) in these patients appears mainly in the urine. Where they occur depends on their structure. The light chains may be present as monomers, dimers, or tetramers. The monomers are small molecular weight (22,000 daltons) and readily pass through the glomerular basement membrane into the urine. Dimers weigh approximately 44,000 daltons and pass into the glomerular filtrate less readily than monomers. Therefore, patients with dimers may have a small spike in the serum and most of the Bence Jones protein in the urine. Lastly, patients with tetramers will have virtually no Bence Jones protein in the urine, while a spike is seen in the serum. This is because the weight of the tetramer is too great to pass through a normal glomerular basement membrane. The finding of Bence Jones proteins is clinically significant for the patient. Individuals with large amounts of Bence Jones proteins tend to do worse clinically than individuals who lack Bence Jones proteins. The main reason for this is that the Bence Jones proteins are reabsorbed by the proximal convoluted tubules in the kidneys. There, depending on their immunochemical makeup, they may damage the proximal tubules. Significant damage may occur and produce myeloma kidney. In this, the proximal tubule cells form syncytia, and casts consisting of monoclonal protein are found within the tubule lumen. Renal failure is an uncommon but real cause of demise in some patients with multiple myeloma (12,13,14,15,16,17,18 and 19).

The classification of multiple myeloma is not merely based on finding a monoclonal protein in serum. Indeed, some investigators worry that they may not be able to detect every possible case of monoclonal gammopathy no matter how extensive a screening

program they set up. Clearly, there are cases of nonsecreting or hyposecreting myelomas that will be missed. There are other unusual cases of myeloma in which a relatively small amount of the monoclonal protein is produced and hides under a normal beta region component such as transferrin (simulating an iron-deficiency anemia pattern) or under C3 (simulating a subacute inflammatory response pattern). The laboratorian must be aware that serum protein electrophoresis, though an excellent screening method, is not the only criterion used for the diagnosis of multiple myeloma. These individuals need to be assessed by virtue of their clinical symptoms (bone pain, anemia, infectious disease problems, renal disease, neuropathies, etc.) (11, 20). Some unusual variants of multiple myeloma may be more difficult to diagnose. One of these is POEMS syndrome. This stands for the 5 to 7% of cases of multiple myeloma that present mainly with a peripheral neuropathy, organomegaly (mainly the liver and spleen), endocrinopathy (often thyroid dysfunction or diabetes mellitus), M-protein, and skin lesions (usually increased skin pigmentation). These individuals have the interesting skeletal findings of osteosclerotic lesions rather than the usual osteolytic lesions. Note that both osteosclerotic and osteolytic lesions may be present simultaneously. Sometimes, multiple myeloma may be present with amyloidosis (amyloid AL) (21,22 and 23).

Immunoglobulin Isotypes in Multiple Myeloma

The isotypes of immunoglobulins seen in multiple myeloma largely reflect their prevalence in normal serum. IgG myeloma proteins predominate. They are almost always monomers weighing 160,000 daltons. It is rare that IgG myeloma proteins produce significant problems with hyperviscosity (see Waldenström's Macroglobulinemia, below). IgE myeloma is extremely rare and can occur as monomers or as polymers with variable molecular weight. Because IgE molecules tend to self-aggregate, they, like IgM, may cause problems with hyperviscosity symptoms. IgA myeloma may be difficult to diagnose on occasion because IgA tends to migrate in the P region. Therefore, early in its course, the other beta region bands may mask the IgA myeloma. C3, transferrin, or fibrinogen from an inadequately clotted sample of blood or an inadvertently examined sample of plasma may obscure the presence of an IgA myeloma M-protein. Also, IgA myeloma bands tend to be more diffuse because of the heavy glycosylation of IgA myeloma proteins. IgD myelomas are quite uncommon, representing about 1% of the total cases of multiple myeloma. The light chain types usually reflect the percentage of kappa and lambda present in the serum. Two thirds of IgG and IgA myelomas are of the kappa light chain type and one third are of the lambda type. However, an exception occurs with IgD myeloma, for which the ratio of kappa to lambda light chains is 1:9.

Although some anecdotal reports have illustrated minor differences in clinical course if one has kappa as opposed to lambda light chains, no significant difference has been documented in more extensive studies. Therefore, the typing of light chains is merely useful for classifying the patients but does not alter the prognosis.

IgD myelomas also may be missed because the amount of M-protein may be relatively small (24). If one is using the older five-band electrophoresis techniques, a normal alpha 2 or beta region band may hide the small IgD component. Also, as discussed later, when radial immunodiffusion plates perform quantification, at least two dilutions of the patient's serum must be performed to avoid antigen excess. In antigen excess, the ring of precipitation may be so small that it would go unnoticed or be washed away. Indeed, this is one of the reasons why the recently released monoclonal gammopathy guidelines recommends not using radial immunodiffusion to quantify the immunoglobulins.

It is uncommon for multiple myeloma to occur with IgM. That is not to say that it does not occur.

Sometimes one can reach that incorrect conclusion from the literature. However, only a handful of cases of multiple myeloma with IgM have been reported. This is because patients with IgM M-proteins usually have a different clinical disease called Waldenström's macroglobulinemia. This author has seen only rare examples of multiple myeloma with IgM. One of them was a defective IgM product that contained only the CH1 and variable portion of the IgM chain together with a single light chain. It appeared mainly in the patient's urine. It is important to anticipate such unusual events, but overall, IgM myeloma is about as rare as IgE myeloma (25).

Alpha heavy-chain disease is extremely uncommon in the western world. In the Middle East and Mediterranean regions, an unusual disease called a heavy-chain disease occurs. The tissue distribution of the plasma cells in this lesion roughly parallels that for the normal distribution of IgA along the gastrointestinal tract and other mucosal surfaces. This disease is different from IgA multiple myeloma. In alpha heavy-chain disease, the individuals usually have a slow progression of their disease during the early phase. Cures have been reported with antibiotic treatment at this stage. Also, the disease is particularly difficult to diagnose in the clinical laboratory. Often, serum samples do not show obvious abnormal bands on serum protein electrophoresis, although one may find a rather broad band in the beta region. Because no reaction is seen with the light-chain antisera, the assumption is made often that a monoclonal gammopathy is not present. Note that with IgA myeloma, light chains are often difficult to demonstrate. A technique termed immunoselection has been used to assist in this diagnosis. In this technique, antibodies against light chains are mixed into the agarose (similar to setting up a radial immunodiffusion). Then, a standard immunoelectrophoresis is carried out. Since intact immunoglobulin molecules contain either kappa or lambda light chains, they must precipitate with the anti-kappa and anti-lambda in the sample well. The only molecules to migrate from the well will be those that lack light chains (hence the alpha chain disease molecule). The distance migrated from the well correlates with this concentration. Another way to determine alpha-chain disease is always to quantify IgG, IgA, IgM, and light chains by nephelometry or turbidimetry and to perform serum protein electrophoresis simultaneously. Patients with alpha chain disease will have increased concentrations of IgA. By subtracting the sum of the light chain-containing immunoglobulins from the IgG +IgA +IgM, a relatively large positive number would, be obtained (normally the number should be near zero). This is because the excess alpha chains do not have detectable light chains. Therefore, one would suspect the possibility of heavy-chain disease from this calculation and perform an immunoselection. The diagnosis

of alpha-chain disease may be missed in the western world because these cases are so rare that immunoselection technique is not a standard procedure in clinical laboratories in this country (25).

The more common form of heavy-chain disease seen in the Western world is Franklin's disease (gamma heavy-chain disease). Most of these patients have a form of B-cell lymphoma or leukemia. The patients may present with a generalized lymphadenopathy, hepatosplenomegaly, pleural effusions, and ascites. Similar to alpha chain disease, an immunoselection method may be optimal to confirm this diagnosis (25). We prefer to use the Sun technique, a modified immunoselection procedure (26).

Biclonal gammopathies can be somewhat confusing. These conditions have expansion of two B lymphocyte clones rather than one. Actually, there is no special significance to this finding. Individuals who clinically have multiple myeloma have no different disease process if they have one clone or two clones. To confirm a true biclonal gammopathy, one must either demonstrate two different light chain types, which are by definition two distinct clones, or perform DNA rearrangement studies to document this distinction. If one has two clones with the same light chain, DNA studies would be the only way to determine if they were from one clone or two. Because there is no clinical significance to making this determination, it is not cost-effective. About 5% of patients with monoclonal gammopathies have a biclonal gammopathy pattern. Note that the presence of two different heavy chain types with the same light chain type does not imply necessarily a biclonal gammopathy. During maturation of B-lymphocytes, there is often a stage at which a switch occurs from one heavy-chain isotype to another. That is why some B lymphocytes will have both a mu chain and an alpha or gamma chain simultaneously on their surface membrane. Therefore, the finding of two different heavy chain types with the same light chain isotype does not establish the existence of a biclonal gammopathy. As mentioned earlier, however, making the distinction is of no clinical significance (25).

Lastly, some patients with multiple myeloma present with profound hypogammaglobulinemia. Although there have been several theories as to why this profound hypogammaglobulinemia often occurs in multiple myeloma, the pathogenesis is still uncertain. Further, not all individuals with multiple myeloma will present with a hypogammaglobulinemia. Some studies have shown that peripheral blood lymphocytes from patients with multiple myeloma have a poor response to B-cell mitogens. In other studies, a marked decrease has been found in the number of normal peripheral blood B lymphocytes in patients with multiple myeloma. This implies the existence of some sort of suppressive influence on B lymphocyte maturation. However, it has not been clearly demonstrated that suppressor T lymphocytes are the cause of this change. An increase in the number of circulating T helper cells with the suppressor-inducer phenotype (CD4+, CD45R+) has been described. Patients with multiple myeloma also have an increase in the T helper cells that have the immunoglobulin-secreting helper capacity phenotype (CD4+, CD45R-). Clearly, the two helper and suppressor cells can function to effect the immunoglobulin containing cells in these patients. Patients with T helper lymphomas, for instance, will often have a polyclonal increase in gamma globulins. However, one must be careful to avoid oversimplification with these functions. For example, in individuals with acquired immunodeficiency syndrome (AIDS) the T helper (CD4+) lymphocytes have been destroyed by infection with human immunodeficiency virus (HIV). Yet, a polyclonal immunoglobulin increase is consistently found in the serum of these patients. Further, transient M-proteins may be found in patients with other infectious processes such as subacute bacterial endocarditis (25, 27,28 and 29).

Waldenström's Macroglobulinemia

Patients with Waldenström's macroglobulinemia suffer from complications of hyperviscosity from the presence of large amounts of circulating IgM. IgM almost always circulates as a pentamer (19S). Therefore, when one develops a monoclonal proliferation of cells producing IgM, the increased concentration of large proteins in the blood disturbs the normal circulation hemodynamics, resulting in the hyperviscosity syndrome. This causes a variety of neurologic complaints, cardiac insufficiency, and vascular insufficiency throughout the body. In addition to the hyperviscosity syndrome, Waldenström's macroglobulinemia differs from multiple myeloma in other respects. These patients do not develop the lytic skeletal lesions seen in multiple myeloma. They do tend to present with symptoms of weakness and fatigue because of the combination of hyperviscosity and anemia. Clinically, these patients have a more protracted course than individuals with multiple myeloma. Usually, an IgM level of greater than 20 g/L is needed to produce the hyperviscosity syndrome. Keep in mind that IgM is not the only immunoglobulin that will produce the hyperviscosity syndrome. On occasion, IgA also will produce a hyperviscosity syndrome when secreted in large amounts. This is because IgA tends to self-aggregate and IgA can occur in serum as a multimer. IgG rarely has been reported to cause a hyperviscosity syndrome, and this author has seen one case of hyperviscosity syndrome from K light-chain disease (25, 30).

The clinical laboratory can provide several supportive bits of information for this clinical diagnosis. On examination of the peripheral smear, Rouleaux formation is characteristic of this disease. It is not pathognomonic for this condition, as it also may be seen in multiple myeloma. The bone marrow shows an infiltration by lymphoplasmacytoid cells. If one performs immunohistochemical examination, the cells would contain the monoclonal IgM component. Obviously, the serum viscosity usually is increased in these patients. Serum protein electrophoresis is an excellent screening test for this condition. The usual case of Waldenström's macroglobulinemia will show a monoclonal spike that is located close to the origin. Although this is not requisite, the bulky nature of the molecule together with the relative lack of charge at the pH 8.6 normally used for serum protein electrophoresis causes the IgM gammopathy to linger near the origin. Quantification of immunoglobulins discloses large amounts of IgM, usually greater than 20 g/L, and normal amounts of IgG and IgA. Performing immunofixation or immunoselection following serum protein electrophoresis best identifies the gammopathy. Although can identify large M-proteins by measuring IgG, IgA, IgM, kappa and lambda together

with high-resolution electrophoresis, small M-proteins are not readily identified by this technique. Immunoelectrophoresis will identify most M-proteins, but is slower than immunofixation and often more difficult to interpret. Therefore, the recent guidelines recommend replacement of immunoelectrophoresis with immunofixation (1, 5, 31,32,33 and 34).

Monoclonal Gammopathy Of Undetermined Significance

In older textbooks, the term for this section would have been “benign monoclonal gammopathies.”

That name is incorrect. The term benign implies that we can predict the biological behavior of the neoplasm. In pathology this means that a benign lesion will not invade further nor metastasize with potentially lethal consequences. We cannot make such a confident statement about a monoclonal gammopathy that we identify in the serum. Kyle correctly coined the term MGUS to categorize these individuals until we have a better understanding of the significance of the lesions that we detect. A restricted band on serum protein electrophoresis in the gamma region that is identified as a monoclonal gammopathy implies that it has been the product of a single clone or plasma cells. This is correct. However, the nature (benign versus malignant) of those plasma cells and their B-lymphocyte precursors cannot be inferred from the presence of the gammopathy. Even relatively large gammopathies, up to 3,000 mg/dL, may represent a product of a benign lymphoplasmacytic lesion, while some cases of multiple myeloma will produce tiny or undetectable amounts of monoclonal component (3).

In his series, Kyle found that 70% of patients with monoclonal gammopathy had MGUS as their working diagnosis. By this term, he referred to individuals who have a monoclonal component demonstrated in the serum, but who lack other key features for diagnosing a malignant monoclonal gammopathy. These patients are not anemic, they do not have lytic lesions, and they do not have elevated serum calcium or serum enzymes. They do not have a prominent plasmacytosis in the bone marrow, although they may have as many as 10% plasma cells in the bone marrow. About 20% of patients with MGUS over a 10-year period will progress to develop a malignant B-cell lymphoproliferative lesion, usually multiple myeloma. Some of the other cases will develop into chronic lymphocytic leukemia, well-differentiated lymphocytic lymphoma, amyloidosis, and rarely other forms of B-cell proliferative lesions. The vast majority of patients with MGUS do not develop any other overt clinical process (4).

The majority of patients with MGUS do not have a distinctive pathological process that can be detected.

Obviously, somewhere in their bodies they harbor a clone producing this gammopathy. Some may merely represent a response to a particular infection. Others may be part of an autoimmune process in which one clone against the autoantigen predominates. In many other cases, however, we have no idea if some antigen is driving the clone or whether it is a self-driven, neoplastic clone. MGUS patients have quantities of M-protein that range from as little as 300 mg/dL to greater than 3,000 mg/dL. Some of these patients have suppression of their normal immunoglobulin synthesis as shown by decrease in the normal gamma-globulin component by serum protein electrophoresis. This feature alone, however, does not portend a negative prognosis. Usually, these patients do not have significant amounts of Bence Jones protein, although small amounts may be detected. Therefore, in any patients having a monoclonal gammopathy detected in serum in our laboratory, we routinely recommend that a urine sample be obtained to detect Bence Jones protein. We believe that currently, the most sensitive technique for this determination is a combination of serum protein electrophoresis and immunofixation. We also recommend that the clinician follow this patient with a serum protein electrophoresis every 6 to 12 months to determine if the process is evolving or regressing. Of course, routine hematologic and chemical analysis should be performed. If these are in the normal range, and there are no clinical symptoms of bone pain, further studies such as bone marrow examination are probably useless (2).

B-Cell Neoplasms

A variety of B-cell neoplasms have been associated with monoclonal gammopathies. These neoplasms are discussed in detail elsewhere in this text. However, it is useful to note their association with monoclonal gammopathies, as they may provide the clinician with further evidence to do more aggressive, invasive diagnostic procedures. It used to be thought that monoclonal gammopathies were rare in cases of chronic lymphocytic leukemia or well-differentiated lymphocytic lymphoma. In fact, the majority of these patients have a monoclonal gammopathy in either their serum or their urine when high-resolution electrophoresis and immunofixation are combined to study these samples. Monoclonal gammopathies also have been demonstrated in patients with Burkitt's lymphoma and B-cell ALL. Clearly, serum protein electrophoresis is not the ideal way to diagnose these conditions. Nevertheless, the finding of a monoclonal band in the serum or urine of a patient with malaise and lymphadenopathy with splenomegaly, for example, might prompt the clinician to perform a lymph node biopsy (24, 35, 36).

Amyloidosis

About 4% patients with plasma cell proliferation will have amyloid deposition in their tissues. There are two major biochemical types of amyloid. One consists of an immunoglobulin light chain and is termed AL. The second is composed of protein A and is termed AA. AA is probably a breakdown product of the much larger normal serum molecule amyloid A-related protein. When either type of amyloid deposits in tissues, a dysfunction of the involved organs occurs, depending on the location of the amyloid deposits. For instance, in patients with AL amyloid, there is frequently deposition in the heart, gastrointestinal tract, tongue, blood vessels, tendon, skin, and peripheral nerves. The clinical picture that these individuals present with depends on which site is involved. For example, patients with amyloid deposition primarily in the tongue will present with macroglossia. Those who have deposition within the heart will present with symptoms of congestive heart failure.

Patients with amyloidosis and monoclonal gammopathy usually

do not have bone pain or osteolytic lesions. It may be appropriate to biopsy the rectum, or involved sites to determine the presence of amyloid. Tissues with extensive involvement of amyloid stain dark blue with iodine, hence the term amyloid (meaning starchlike). The definitive diagnosis of amyloidosis requires positive staining with Congo red, giving a characteristic blue-green birefringence when polarized light is used. These tinctorial qualities relate to the structural beta-pleated sheet formation that occurs with either AL or AA. Monoclonal gammopathies may be found in both the serum and urine of patients with amyloidosis (23, 25).

Methodologies

Immunologic Reaction in Gel Diffusion

The classic precipitation reaction is the basis for immunofixation and immunoelectrophoresis. As discussed earlier in this section, antibody-antigen interactions are highly concentration dependent. Therefore, when an antibody is placed in one well and diffuses toward the antigen placed in another well, a precipitation reaction occurs somewhere between the two wells. The distance from the well where the band forms depends upon the concentration and the size of the reactants. When the antigen and antibody are of approximately equal size and concentration, the band will occur midway between the two wells. These reactions are highly specific and are therefore, useful to identify the specific immunoglobulin involved in a monoclonal gammopathy (25).

Serum Protein Electrophoresis

The older five-band electrophoretic patterns have been replaced in recent years by high-resolution agarose or acetate electrophoresis patterns. These are vastly superior for evaluating serum for the presence of a monoclonal gammopathy. They permit distinction of subtle monoclonal bands that are present in the alpha 2 and beta regions, as well as the easy-to-see bands that are present in the y region. The older five-band gel methods did not permit facile discrimination of bands in the alpha 2 and beta regions. Several versions of high-resolution electrophoresis are commercially available in both agarose and acetate form. Any of the systems is preferred over the older acetate form. The older acetate gels would be adequate screens if one were to perform immunofixation along with them. However, this is expensive, time-consuming, and redundant in most cases (25, 37). The current monoclonal guidelines recommend that a technique that provides crisp resolution of the beta-1 and beta-2 region bands (transferrin and C3) (1).

Most, but not all clinically significant monoclonal gammopathies are readily detected electrophoresis that provide crisp definition of the beta and gamma regions. They contain either a band in the alpha-2, beta, or gamma region, a distorted band from binding of a normal component by the monoclonal gammopathy, or hypogammaglobulinemia (the latter must always be evaluated with immunofixation and urine examination for Bence Jones protein) (24, 25).

When a sample is sent for monoclonal gammopathy evaluation, a high quality electrophoresis is performed together with either immunofixation, immunosubtraction, or immunoglobulin quantification if an M-protein is seen. The presence of any abnormal band or hypogammaglobulinemia on the electrophoresis gel results in the performance of an immunofixation on that sample. If the electrophoresis demonstrates even minor distortions in the beta region an immunofixation study will be performed. This reflects the fact that IgA and IgM M-proteins may be present in the beta region and the presence of transferrin, C3, and beta-1 lipoprotein in this region may obscure a small M-protein (25, 5). Although the identification of the M-protein in most cases can be evaluated without resorting to immunofixation if high-resolution electrophoresis is performed with quantification of IgG, IgA, IgM and kappa and lambda, most laboratories perform immunofixation to identify the M-protein. The newer immunofixation procedures are quick, efficient and some have features of automation that dramatically decrease the workload (5).

Urine electrophoresis and immunofixation should be performed on all patients who are suspected of having myeloma, Waldenström's macroglobulinemia or amyloid AL. Light-chain disease may have a normal or mildly hypogammaglobulinemic pattern on serum protein electrophoresis. This is why urine must be examined along with serum in all of these cases as recommended by the new guidelines (1, 5).

Immunofixation

In evaluating patients for a monoclonal gammopathy, one must put the entire clinical picture together: the history that is provided, and other laboratory parameters (hematocrit, total protein, evidence of lytic lesions, serum calcium) to make the correct diagnosis (38,39,40,41,42,43 and 44). Immunofixation is performed by repeating the serum protein electrophoresis with the patient sample run in several tracts. After the electrophoretic step, each tract is overlaid with a specific antiserum against IgG, IgA, IgM, kappa or lambda (Fig. 64.4). The patient sample applied to each tract is diluted to achieve a concentration of approximately 100 mg/dL prior to electrophoresis. This optimizes the immunoprecipitation reaction. If too large a quantity of patient serum is used, an antigen excess effect will be seen. If too small a concentration is used, rarely, an antibody excess effect will be seen or the resulting reaction may be too small to be seen. Therefore, one must know the concentration of specific immunoglobulins prior to performing immunofixation in order to use the correct concentration of serum. Following electrophoresis, the specific antiserum is overlaid onto the tract and allowed to incubate for 30 minutes. After this step, the sample is washed so that unprecipitated proteins will be removed. Thereafter, the sample is stained with a sensitive dye. Coomassie blue or Crystal violet serves well in this regard. After the sample is dried, a permanent record is formed. We use immunofixation to detect subtle gammopathies where there is not an obvious band by serum protein electrophoresis, but where there is a lingering clinical suspicion of the patient's harboring an M-protein (45,46,47,48,49 and 50).

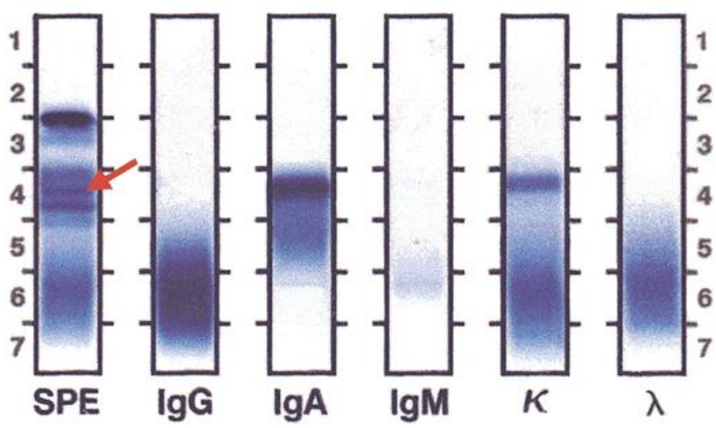


FIGURE 64.4. Immunofixation with an IgA kappa M-protein. Note the small M-protein band (arrow) in the SPE lane. The IgA and kappa restrictions are obvious.

Capillary Zone Electrophoresis

In capillary zone electrophoresis (CE), serum is aspirated into narrow (50 μ) capillary tubes that have a strong net negative

charge. This negative charge provides a strong endosmotic flow, which propels the serum proteins to the cathode. At the cathodal end, protein fractions are estimated by measuring the optical density at 215 nm (peptide bond absorbance) and an electropherogram is produced (Fig. 64.2). This is a highly automated version of electrophoresis because there is no staining and destaining. The technologist puts reagents on the instrument, reviews the electropherograms, and performs maintenance on the instrument. About 40 samples can be processed in an hour. The instrument allows the operator to convert the electropherograms into virtual gel images that resemble the high-resolution gel patterns with crisp separation of the beta-1 and beta-2 regions. However, one of the drawbacks of this methodology is the relatively short gamma region compared to the high-resolution gel-based systems. Artifacts also have been reported in this system because of non-M-protein molecules that absorb at 215 nm. For instance, radiocontrast dyes may produce a restriction that resembles an M-protein. This is another reason why all putative M-protein spikes must be characterized (whether seen on CE or traditional gel-based methods). CE is relatively new for general application, but has been gaining application in larger laboratories that wish to have an automated, high-resolution system (51).

Immunosubtraction

Immunofixation may not be performed on a CE system. However, advantage may be taken of a technique called immunosubtraction (ISUB) that was described several years before CE was even used. With ISUB one mixes the serum with beads coated with specific antibody. After, an incubation is performed for each major heavy-chain isotype and kappa and lambda. Identification of the M-protein is accomplished by its removal using the appropriate antibody-coated beads. For example, a serum containing an IgG kappa M-protein will have the spike removed by beads coated with anti-kappa, but not by those coated with anti-lambda. Similarly, the spike will be removed by beads coated with anti-IgG (or Staph protein A), but not by beads coated with anti-IgM or IgA. Although the process sounds laborious, commercially available coated beads are packaged for complete performance of IgG, IgA, IgM, kappa and lambda identification on the automated CE instrument. Unfortunately, anti-IgD and anti-IgE beads were not readily available at the time of this writing. The other drawback of ISUB is its inherent insensitivity. With IFE, there is an enhancement of the sensitivity of electrophoresis that allows detection of small M-proteins that are not seen on even high-quality systems. With ISUB, if you do not see the spike on the electrophoresis step, it obviously cannot be immunosubtracted. Nonetheless, this is a practical technique for quickly identifying most M-protein spikes (26).

REFERENCES

1. Keren DF, Alexanian R, Goeken FA, et al. Guidelines for clinical and laboratory evaluation of patients with monoclonal gammopathies. *Arch Pathol Lab Med* 1999;123:106–107.

2. Alexanian R, Weber D, Liu F. Differential diagnosis of monoclonal gammopathies. *Arch Pathol Lab Med* 1999;123:108–113.

3. Kyle RA. Sequence of testing for monoclonal gammopathies. *Arch Pathol Lab Med* 1999;123:114–118.

4. Kyle RA. Monoclonal gammopathy of undetermined significance and solitary plasmacytoma. Hematology/Oncology Clin N Amer 1997;11:71–87.

5. Keren DF. High-resolution electrophoresis and immunofixation. Butterworth Heinneman, Boston and London, second edition, 1994.

6. Keren DF. Procedures for the evaluation of monoclonal immunoglobulins. Arch Pathol Lab Med 1999;123:126–132.

7. Keren DF. Detection and characterization of monoclonal components in serum and urine. Editorial. *Clin Chem* 1998;44:1143–1145.

8. College of American Pathologists SURVEY Report EC-07, 1991.

9. Alexanian R, Barlogie B, Dixon D. Prognosis of asymptomatic multiple myeloma. *Arch Intern Med* 1988;148:1963–1965.

10. Clamp JR. Some aspects of the first recorded case of multiple myeloma. *Lancet* 1967;2:1354–1356.

11. Shustik C, Michei R, Karsh J. Nonsecretory myeloma: a study on hypoimmunoglobulinemia. *Acta Haematol* 1988;80:153–158.

12. Brigden ML, Neal ED, McNeely MDD, et al. The optimum urine collections for the detection and monitoring of Bence Jones proteinuria. *AJCP* 1990;93:689–693.

13. Diamond J, McLaughin M. Urinary parameters to assess renal function. *Clin Lab Med* 1988;8:495–505.

14. Graziani MS, Lippi U. Multiple myeloma with serum IgM kappa and Bence Jones lambda biclonal gammopathy. *Clin Chem* 1986;32:221–222.

15. Graziani MS. Immunoblotting for detecting Bence Jones proteinuria. *Clin Chem* 1987;33:1079–1080.

16. Kim M. Proteinuria. *Clin Lab Med* 1988;8:527–540.

17. Kosaka M, Iishi Y, Okagawa K, et al. Tetramer Bence Jones protein in the immunoproliferative diseases. Angioimmunoblastic lymphadenopathy, primary amyloidosis and multiple myeloma. *Am J Clin Pathol* 1989;91:639–646.

18. Norden AGW, Flynn FV, Filcher LM, et al. Renal impairment in myeloma: negative association with isoelectric point of excreted Bence Jones protein. *J Clin Pathol* 1989;42:59–62.

19. Pezzoli A, Fascali E. The clinical significance of pure Bence Jones proteinuria at low concentration. *Am J Clin Pathol* 1989;91:470–475.

20. Smith DB, Harris M, Cowland E, et al. Non-secretory multiple myeloma: a report of 13 cases with a review of the literature. *Hematol Oncol* 1986;4:307–313.

21. Bolling JP, Brazis PW. Optic disk swelling with peripheral neuropathy, organomegaly, endocrinopathy, monoclonal gammopathy, and skin changes (FOEMS) syndrome. *Am J Ophthalmol* 1990;109:503–510.

22. Troussard X, de Ligny BH, Gallet B, et al. Massive systemic amyloidosis associated with light chain deposition disease. *Nephron* 1989;52:139–143.

23. Glenner GG. Amyloid deposits and amyloidosis. *N Engl J Med* 1980;302:1283–1292.

24. Keren DF, Warren JS, Lowe JB. Strategy to diagnose monoclonal gammopathies in serum: high-resolution electrophoresis, immunofixation and K/L quantification. *Clin Chem* 1988;34:2196–2201.

25. Keren, DF. Clinical Indications for Electrophoresis and Immunofixation. In: Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of Clinical Immunology*, 5th ed. Washington, DC: ASM Press, 1997:65–74.

26. Sun T, Peng S, Narurkar L. Modified immunoselection technique for definitive diagnosis of heavy chain disease. *Clin Chem* 1994;40:664.

27. Danon F, Seligmann M. Transient human monoclonal immunoglobulins. *Scand J Immunol* 1972;1:323–328.

28. Malacrida V, De Francesco D, Banfi G, et al. Laboratory investigation of monoclonal gammopathy during 10 years of screening in a general hospital. *J Clin Pathol* 1987;40:793–797.

29. Keren DF, Morrison N, Gulbranson R. Evolution of a monoclonal gammopathy documented by high-resolution electrophoresis and immunofixation. *Lab Med* 1994;25:313–317.

30. Carter PW, Cohen HJ, Crawford J. Hyperviscosity syndrome in association with kappa light chain myeloma. *Am J Med* 1989;86:591–595.

31. Kahn SN, Bina M. Sensitivity of immunofixation electrophoresis for detecting IgM paraproteins in serum. *Clin Chem* 1988;34:1633–1659.

32. Roberts RT. Usefulness of immunofixation electrophoresis in the clinical laboratory. *Clin Lab Med Sci* 1986;6:601–609.

33. Crawford J, Eye MK, Cohen HJ. Evaluation of monoclonal gammopathies in the “well” elderly. *Am J Med* 1987;82:39–45.

34. Young VH. Transient paraproteins. *Proc R Soc Med* 1969;62:778–780.

35. Braunstein AH, Keren DF. Monoclonal gammopathy (IgMK) in a patient with Burkitt's lymphoma. *Arch Pathol Lab Med* 1983;107:235–238.

36. Nathwani BN, Winberg CD, Berman RM. Angioimmunoblastic lymphadenopathy with dysproteinemia and its progression to malignant lymphoma. In: Jaffe ES, ed. *Surgical Pathology of Lymph Nodes and Related Organs*. Philadelphia: WE Saunders, 1985:57–85.

37. Laurell CB. Electrophoresis, specific protein assays or both in measurement of plasma proteins. *Clin Chem* 1973;19:99–102.

38. Bush D, Keren DF. Over and underestimation of monoclonal gammopathies by quantification of serum kappa and lambda containing immunoglobulins. *Clin Chem* 1992;38:315–316.

39. Davenport RD, Keren DF. Oligoclonal bands in cerebrospinal fluids: significance of corresponding bands in serum for diagnosis of multiple sclerosis. *Clin Chem* 1988;34:764.

40. Dyck PJ. Intravenous immunoglobulin in chronic inflammatory demyelinating polyradiculoneuropathy and in neuropathy associated with IgM monoclonal gammopathy of unknown significance. *Neurology* 1990;40:327–328.

41. Fifield R, Keller I. An immunochemical evaluation for the identification and typing of monoclonal proteins. *Ann Clin Biochem* 1990;27:327–334.

42. Howerton DA, Check IJ, Hunter RL. Densitometric quantitation of high resolution electrophoresis agarose gel protein electrophoresis. *Am J Clin Pathol* 1986;85:213–218.

43. Keren DF, Di Sante AC, Bordine SC. Densitometric scanning of high-resolution electrophoresis of serum: methodology and clinical application. Am J Clin Pathol 1986;85:348–353.

44. Liu FJ, Fritsche HA, Trujillo JM. Underestimation of monodonal proteins by serum protein electrophoresis on cellulose acetate. *Clin Chem* 1987;33:182–184.

45. Williamson AR, Salaman MR, Kreth HW. Microheterogeneity and allomorphism of proteins. *Ann NY Acad Sci* 1973;209:210–224.

46. Weiss LM, Strickler JG, Dorfman RF. Clonal T-cell population in angioimmunoblastic lymphadenopathy and angioimmunoblastic lymphadenopathy-like lymphoma. *Am J Pathol* 1986;122:392–397.

47. Themer PS, Bedard YC, Femandes BJ. I-light chain nephropathy with Fanconi's syndrome. *Arch Pathol Lab Med* 1983;107:654–657.

48. Stemerman D, Papadea C, Martino-Saltzman D, et al. Precision and reliability of paraprotein determinations by high-resolution agarose gel electrophoresis. *Am J Clin Pathol* 1989;91:435–440.

49. Sanders PW, Herrera GA, Chen A, et al. Differential nephrotoxicity of low molecular weight proteins including Bence Jones proteins in the perfused rat nephron in vivo. *J Clin Invest* 1988;82:3086–3096.

50. Rosenfeld L. Henry Bence Jones (1si3-1873): the best chemical doctor in London. *Clin Chem* 1987;33:1687–1692.

51. Keren DF. Capillary zone electrophoresis in the evaluation of serum protein abnormalities. *Am J Clin Pathol* 1998;110;248–252.

65
Primary Immunodeficiency Diseases

Jeffrey S. Warren

Our understanding of primary immunodeficiency diseases has advanced rapidly. There is perhaps no better example than that of X-linked agammaglobulinemia (XLA). In 1952, a pediatrician named Colonel Ogden Bruton described an 8-year-old boy who, beginning at approximately 2 years of age, had begun to suffer from recurrent bacterial pneumonia and septicemia. Recognition of the association between recurrent infections with high-grade bacterial pathogens and very low concentrations of serum gammaglobulins had been made possible by the recent development of serum protein electrophoresis. The cause-and-effect nature of this association was confirmed when the infections ceased following high-dose intramuscular gammaglobulin injections. Examination of similar patients, uniformly boys, coupled with the observations that infections were caused by encapsulated bacteria such as *Streptococcus pneumoniae*, *Hemophilus influenzae*, *Streptococcus pyogenes*, etc., but not by intracellular pathogens such as mycobacteria, viruses, and fungi, and that lymph nodes from XLA patients were devoid of germinal centers but not lymphocytes in the deep paracortical regions, led to the conclusion that Bruton's agammaglobulinemia is an X-linked defect in B lymphocyte maturation. In 1993 it was determined that XLA is caused by a variety of mutations in the X-linked gene, *BTK* (B-cell–specific or Bruton's tyrosine kinase).

From a clinical, immunologic, and “pathogenesis of disease” standpoint, it remains useful to group immunodeficiency disorders into broad categories based upon whether the defect is congenital or acquired and according to which system (or systems) of host defense is defective. These host defense systems have traditionally included the complement system, the phagocytes, the humoral immune system (immunoglobulins), the cellular immune system, and combined systems. The latter grouping reflects disorders that result from defects in more than one system (e.g., severe combined immunodeficiency). It must be emphasized that even in the case of nosologically distinctive entities there are features that extend beyond a single “simple” system. For example, optimal antibody production by B lymphocytes (humoral immunity) requires intact T helper lymphocytes. Likewise, efficient phagocytosis of some microorganisms requires opsonization of the foreign particles with C3b and/or iC3b (complement system) and immunoglobulin (humoral immune system). In many diseases, seemingly unrelated immune defects are observed in the same patient. It was stated in the first edition of this textbook that “causative defects have been identified in only a few immunodeficiency disorders.” In the short time since publication of the first edition of *Clinical Laboratory Medicine*, this statement has become outdated. Literally dozens of primary immunodeficiency diseases now have been defined at the genetic level. For each disease, many questions remain. These entail such issues as: why are there variations in disease expression? How are genetic defects translated into functional defects? How can knowledge of genetic defects be exploited diagnostically and therapeutically?

Primary immunodeficiency diseases frequently present as failure to thrive, chronic or recurrent infections, unusual infections (e.g., uncommon microorganisms or atypical manifestations), and allergic disorders. Immunodeficiency patients may exhibit nonspecific manifestations such as skin rashes or eczema, chronic diarrhea, evidence of exaggerated autoimmunity, lymphadenopathy, hepatosplenomegaly, and growth failure. Some primary immunodeficiency disorders are associated with features that appear to be unrelated to the host defense system. For example, patients with the Chedick-Higashi syndrome typically express partial oculocutaneous albinism, patients with the DiGeorge syndrome may exhibit structural defects related to the development of the third and fourth branchial pouches (e.g., structural abnormalities of the great vessels and absent parathyroid glands), and patients with ataxia-telangiectasia exhibit characteristic neurological and cutaneous anomalies. In general, the frequency and severity of infections observed in a patient with an immunodeficiency disease parallel the magnitude of the immunologic defect. For example, in the absence of definitive therapy, patients with severe combined immunodeficiency (SCID) typically survive less than 1 year and suffer from frequent, often overwhelming bacterial, viral, and fungal infections. In contrast, patients who harbor IgA and/or isolated IgG subclass deficiencies may be asymptomatic, suffer from only a slightly increased frequency of upper respiratory infections, or may suffer from moderately severe lower-respiratory-tract infections that lead to complications such as bronchiectasis. The types of infection that occur may provide insight into the type of immunodeficiency. Recurrent pyogenic infections such as bacterial pneumonia and otitis media are typical of hypogammaglobulinemic patients while those with cellular immune defects commonly suffer from viral, fungal, mycobacterial, and protozoal infections. Patients who harbor deficiencies of the distal complement system (C5-C8) may suffer from recurrent neisserial sepsis.

Many primary immunodeficiency disorders can be classified based on clinical findings and the results of widely available laboratory

tests. Specialized evaluations of the immune system are often labor intensive and expensive. Clinical presentation, a detailed family history, and course of illness, provide important diagnostic clues and must be carefully considered in order to carry out an efficient laboratory evaluation. Leukopenias are relatively common given to prevalence of immunosuppressed oncology and transplantation patients. In most cases of iatrogenic immunodeficiency the cause is readily apparent. Disorders that are not primary to the “classic” immune system, per se, also may be accompanied by increased susceptibility to infections. For example, patients with cystic fibrosis often suffer from persistent, recalcitrant lower-respiratory infections secondary to *Pseudomonas* and *Staphylococcus aureus*. Among cystic fibrosis patients mucous secretions may be poorly cleared because of their high viscosity. Sickle cell anemia often leads to splenic microinfarctions and autosplenectomy, which in turn is associated with an increased risk of bacterial sepsis. There are a variety of dysfunctional cilia syndromes in which impaired ciliary clearance of respiratory secretions results in an increase in susceptibility to bronchopulmonary infections. Finally, the importance of the acquired immunodeficiency syndrome (AIDS) cannot be overemphasized. While these disorders will not be discussed further here, they serve to emphasize that a thorough general medical evaluation is required to arrive at a correct diagnosis in a patient in whom an immunodeficiency disorder is a consideration.

This chapter will provide a necessarily brief overview of the most common and/or prototypic primary immunodeficiency diseases. Brief discussions of laboratory studies that have relevance to the diagnosis and management of immunodeficiency disorders will be incorporated into each corresponding section. Excellent, comprehensive texts that review the clinical, epidemiologic, genetic, and diagnostic features of primary immunodeficiency diseases are available.

- DISEASE STATES

DISEASE STATES

Part of "65 - Primary Immunodeficiency Diseases"

Disorders of the Complement System

Since the initial characterization of a patient with a genetically determined complement deficiency in 1960, defects and/or deficiencies of nearly all of the complement components and their soluble and cell surface-linked regulatory proteins have been described (Table 65.1). Despite the existence of approximately 24 such proteins, primary complement component deficiencies account for fewer than 1% of all patients with recognized immunodeficiency disorders. Most of the inherited disorders of complement result in either recurrent infections, systemic lupus erythematosus (SLE), lupuslike disorders, other autoimmune disorders, or glomerulonephritis. Characteristically atypical clinical features may provide clues that suggest an inherited disorder of the complement system. The biology of the complement system is detailed elsewhere, but a number of less familiar or only recently recognized biological consequences of complement activation provide insight into how complement system defects lead to disease.

TABLE 65.1. COMPLEMENT COMPONENT DEFICIENCIES[a]

Complement	Gene(s)	Inheritance[b]
C1q	*C1QA, C1QB, C1QG*	AR
C1r/C1s	*C1R/C1S*	AR
C4	*C4A/C4B*	AR
C2	*C2*	AR
C3	*C3*	AR
C5	*C5*	AR
C6	*C6*	AR
C7	*C7*	AR
C8α/β	*C8A/C8B*	AR
C8γ	*C8G*	AR
C9	*C9*	AR
Factor I	*IF*	AR
Factor B	*BF*	AR
Factor D	*?*	?
Factor H	*HF1*	AR
Properdin	*PFC*	X-linked R
Mannose binding lectin	*MBL*	AR and AD
C1 esterase inhibitor	*C1NH*	AD
C4 binding protein	*C4BPA, C4BPB*	?
Decay accelerating factor (CD55)	*DAF*	AR
CD59 (Homologous restriction factor 20; membrane inhibitor of reactive lysis)	*CD59*	AR

[a] Table adapted from Ross SC, Denson P. Complement deficiency states and infection: epidemiology pathogenesis and consequences of neisserial and other infections in an immune deficiency. *Medicine* 1984; 63: 243–273 and Sullivan KE, Winklestein JA. In: HD Ochs, CI Edvard Smith, JM Tuck, eds. *Primary Immunodeficiency Diseases*, 1st ed. New York and Oxford: Oxford University Press, 1999: 397–416.

Activation of the classical complement pathway usually is triggered by fixation of C1q,r,s by stoichiometrically sufficient densities of surface-bound IgM or IgG (subclasses 1, 2, 3). While C1q binding to appropriately spaced Fc domains of either IgM or IgG results in sequential activation of C1r and C1s and C4 cleavage, C4 also can be cleaved via the lectin activation pathway in which mannose binding lectin (MBL) is bound to microbial oligosaccharide moieties. In addition to its well-recognized role as an opsonin (along with iC3b), C3b appears to play an important role in immune complex solubilization and clearance. Primate red blood cells express CR1 receptors that bind to circulating immune complexes that are coated with covalently bound C3b. CR1-bearing red cells function as a “buffering system” to clear C3b-opsonized immune complexes. The immune complex-coated red blood cells pass through the liver and spleen where mononuclear phagocytes remove the complexes from red cells, which then return to the circulation. Finally, there is accumulating evidence that peptide fragments of early complement components (e.g., C3b, C3d), via CR1 and CR2 (CD21) on B lymphocytes (and perhaps often cell types), amplify the humoral immune response and facilitate follicular dendritic cell antigen trapping. In addition to causing clinically important, albeit rare, predispositions to infections and autoimmune disease, characterization of the various defects in the complement system has led to a more detailed understanding of the biology of complement.

Complement Cascade Protein Defects

The clinical consequences of genetically determined deficiencies of individual complement components are varied. The most common

manifestations include increased susceptibility to infection and rheumatologic or autoimmune disorders. Some clinical associations are distinctive [e.g., neisserial infections among patients with late complement component (C5, C6, C7, or C8) deficiencies] or unusual (angioedema in patients with C1 esterase inhibitor deficiency). Affected individuals are frequently asymptomatic. With the debated exception that heterozygous carriers of C2, C3, or C4 deficiency may suffer from an increase in prevalence of autoimmune disorders (e.g., SLE), only patients with homozygous complement component deficiencies suffer from overt clinical illness.

C1q consists of six sets of three polypeptides, C1qA, C1qB, and C1qC. These polypeptides are encoded by three corresponding highly homologous genes (*C1QA*, *C1QB*, *C1QG*) located on chromosome 1. A variety of nonsense and missense mutations in each of the three genes has been described. Some patients produce no detectable C1q while others produce dysfunctional C1q. It appears that C1qA is required for assembly of intact C1q. C1q-deficient patients from a variety of ethnic backgrounds have been described. As would be predicted, these patients exhibit very low CH_{50} values (<5% of normal) and C1 functional activity is typically <1% of normal. Almost all C1q-deficient patients have SLE and many suffer from recurrent pyogenic infections.

C1r and C1s are homologous serine esterases. C1r and/or C1s deficiency are exceedingly rare and detailed knowledge of the specific mutations is lacking. These patients also exhibit low CH_{50} values and suffer from either SLE, glomerulonephritis, and/or bacterial infections.

C4 is synthesized as a single polypeptide which is then cleaved to form the 3-chain disulfide-linked mature C4. C4 exists as two isotypes, C4A and C4B, which differ with respect to their efficiency in transacylation reactions. Complete C4 deficiency, marked by a CH_{50} of <1% of normal, is very rare. A variety of isoform and heterozygous C4 deficiencies have been described. These patients are relatively common and typically suffer from SLE, glomerulonephritis, and/or recurrent infections. C4-deficient SLE patients, like C2-deficient SLE patients, may have prominent cutaneous features and are more frequently ANA and anti-dsDNA negative than their complement-sufficient counterparts.

C2 is closely linked to factor B. C2 deficiency is the most common complete complement component deficiency. More than 90% of completely C2-deficient individuals are homozygous for a 28-base pair deletion, which leads to a splicing abnormality and transcription defect. These patients exhibit very low CH_{50} values and nearly half develop SLE, cutaneous lupus, and/or recurrent infections.

C3 is encoded by a gene on chromosome 19 and is synthesized as a single polypeptide precursor that is processed into two disulfide-linked chains. An array of mutations that lead to C3 deficiency have been described. C3 deficiency, either through a primary genetic defect or through absence of the regulatory protein, factor I (C3b inactivator), leads to reduced serum opsonizing capacity and an increased incidence of severe infections attributable to encapsulated pyogenic bacteria, such as *S. pneumoniae*, *Haemophilus influenzae*, and nonpneumococcal streptococci, as well as others, including Gram-negative enterobacteriaceae and *S. aureus*. Patients with C3 deficiency also suffer from an increased incidence of SLE, glomerulonephritis, and vasculitis. Complete absence of C3 is inherited in an autosomal-recessive pattern while clinically normal heterozygotes possess serum C3 concentrations that approximate one half of normal levels. Because C3b is important in both the propagation of the complement activation sequence and as an opsonin, it is difficult to know how these functional defects affect the risk of pyogenic infections. Because C3-deficient patients tend to suffer from a spectrum of infections similar to that seen in a splenectomized person, C3b may be particularly important as an opsonin in the clearance of bacteria from the circulation. Patients with C3 deficiency exhibit decreased CH_{50} and C3 values.

Deficiencies of components in the lytic pathway result in impaired bacteriolysis without impairment of opsonization. Patients with C5, C6, C7, or C8 deficiencies suffer from frequent disseminated neisserial infections (*Neisseria meningitides* and *Neisseria gonorrhoeae*), but interestingly, the clinical illness in such patients is less severe and epidemiologically distinct from that observed in disseminated neisserial infections seen in complement-sufficient individuals. C5-deficient patients exhibit decreased total hemolytic complement activity. There is a diversity of C5 gene mutations that lead to the clinical disorder. The C6, C7, and C9 genes (see below) are linked on chromosome 5. Again, a variety of clinically significant mutations have been identified. C6 deficiency has been subclassified into 3 forms: the most common marked by undetectable C6 and at least two other forms of so-called subtotal C6 deficiency (C6SD). One form of C6SD is characterized by a truncated, but partially functional, C6 molecule. These individuals do not suffer from recurrent infections. Several types of C7 deficiency have been characterized. C8 is composed of three polypeptide chains (α, ß, and γ), which are encoded by three genes, *C8A*, *C8B*, and *C8G*. C8-deficient patients are rare. Defects that lead to deficiencies of the C8 α-γ chain and C8ß chain have been described. It appears that C8ß deficiency occurs primarily in whites while C8 α-γ deficiency is more common in blacks. Patients with deficiencies in the distal complement system exhibit markedly reduced CH_{50} valves.

A few C9-deficient patients have been described, primarily in Japan. C9-deficient serum contains partially reduced hemolytic complement activity. Epidemiologic studies have revealed an increase in incidence of systemic meningococcal infections but this complication is less common in C9-deficient individuals than in individuals deficient of either C5, C6, C7, or C8.

Complement Regulatory Protein Defects

C1 esterase inhibitor (C1-INH) is an autosomal codominantly expressed member of the serine proteinase inhibitor (SERPIN) family. Patients with C1-INH deficiency (because of either a silent C1-INH allele or an allele encoding a dysfunctional molecule) possess less than 50% of normal levels of functional protein. While some patients manifest SLE or lupuslike syndromes, virtually all C1-INH deficient patients manifest hereditary angioedema (HAE). Hereditary angioedema is an autosomal-dominant disorder characterized by episodic, localized, acute attacks of painless edema typically involving the extremities, face, or upper respiratory tract. Some patients manifest with abdominal pain from localized involvement of gastrointestinal mucosa. It is often unclear what triggers an angioedema attack. Type I HAE (85% of cases) is characterized by low plasma levels of antigenic

and functional C1-INH, normal or depressed C1 levels, depressed C4 levels, and normal C3 levels. Because intact C1-INH is catabolized more rapidly than normal, levels of intact C1-INH are often only 5% to 30% of reference concentrations. Type II HAE is less common. Type II HAE is associated with normal or elevated levels of dysfunctional mutant protein together with low levels of normal C1-INH. These patients typically have normal to elevated antigenic C1-INH levels and reduced functional levels.

There are several acquired forms of C1-INH deficiency. These disorders usually are associated with increased C1-INH consumption related to B-cell lymphoproliferative disorders or autoimmune disorders accompanied by autoantibodies directed against C1-INH. Acquired C1-INH deficiency patients typically exhibit depressed C1 and C4 levels, normal or increased antigenic C1-INH, and decreased functional C1-INH levels. Patients with acquired C1-INH deficiency almost always exhibit profoundly depressed C1q levels, while patients with HAE (type I or type II) may have depressed C1q levels, but often do not. Attenuated androgens reduce the frequency and severity of “edema attacks” in the hereditary forms and acquired forms of angioedema, but higher doses are typically required in the latter group.

Defects in nearly all of the other complement regulatory patients have been described but are exceedingly rare (Table 65.1). Special comment is in order regarding genetic deficiencies of decay accelerating factor (DAF) and CD59 and low serum concentrations of mannose binding lectin (MBL).

Genetic defects in DAF and *CD59* are isolated and result in clinical features that are distinct from the more generalized GPI linkage defect seen in paroxysmal nocturnal hemoglobinuria, an acquired clonal disorder. Decay accelerating factor is GPI-linked and plays a role in disassembly of the classical and alternative pathway C3 convertases. Isolated genetic DAF deficiency is associated with the Inab phenotype (absence of the Cromer complex of blood group antigens). DAF deficiency is not associated with clinical hemolysis even though red cells are somewhat more sensitive to *in vitro* complement-mediated lysis. A few DAF-deficient patients have had protein-losing enteropathy but it is unknown what relationship this has with the genetic lesion. CD59 is a GPI-linked protein that inhibits C8 and C9-mediated cell lysis. The only described isolated (primary genetic) CD59-deficient patient identified to date had hemolytic anemia and recurrent CNS infarctions.

Mannose-binding lectin functions somewhat like C1q but can activate the classical and alternative complement pathways. Serum levels are influenced by several polymorphic amino acids. MBL deficiency is believed to increase the risk of both recurrent infections and SLE.

Phagocytic Function Disorders

The importance of phagocyte function in host defense is perhaps best exemplified by the frequent occurrence of severe life-threatening infections in patients rendered granulocytopenic by chemotherapy. Severely granulocytopenic patients (e.g., absolute neutrophil count less than 0.5×10^9/L) frequently suffer from episodes of bacterial sepsis. In many cases, organisms of low intrinsic pathogenicity are responsible. Understanding the mechanisms involved in the individual steps that characterize neutrophil function in the acute inflammatory response (adhesion, chemotaxis, phagocytosis, respiratory burst, and degranulation) is helpful when classifying phagocytic cell defects. With the development of more facile clinical techniques (e.g., flow cytometry and molecular biology), it has become easier to routinely identify specific phagocytic cell defects.

Chronic Granulomatous Disease (and Related Disorders)

Chronic granulomatous disease (CGD) encompasses a genetically heterogenous group of lesions that result in a complete or near-complete absence of the respiratory burst that normally occurs upon phagocyte activation. Activation of the cytoplasmic membrane-associated NADPH oxidase in phagocytes results in the generation of O_2^- (superoxide anion), which in turn gives rise to a variety of reactive oxygen intermediates. Because of the importance of oxidants in phagocyte-mediated microbial killing, individuals who cannot generate NADPH oxidase-derived O_2^- are susceptible to recurrent bacterial and fungal infections.

The clinical, genetic, and molecular heterogeneity of CGD is well understood in terms of NADPH oxidase (Table 65.2). The NADPH oxidase complex consists of four subunits: $gp91^{phox}$, $p22^{phox}$, $p47^{phox}$, and $p67^{phox}$, and a number of associated regulatory proteins ($p40^{phox}$, rap1A, rac1, rac2). Mutations within each of the genes that encode the four major subunits have been described. Mutations in *CYBB* on chromosome Xp21.1 gives rise to the X-linked form of CGD (approximately 65% of all cases). Mutations in the genes which encode $p22^{phox}$ (*CYBA*), $p47^{phox}$ (NCF1), and $p67^{phox}$ (*NCF2*) are located on different autosomes and account for the remaining 35% of CGD cases. $Gp91^{phox}$ and $p22^{phox}$ are membrane-associated in resting phagocytes and contain a cytochrome (b_{-245} or b_{558}) while $p47^{phox}$ and $p67^{phox}$ are located in the cytosol of quiescent cells. Phagocyte activation results in the translocation of the cytosolic component to the membrane component. Patients are classified according to the component of the NADPH oxidase which is affected: $gp91^{phox}$, $p22^{phox}$, $p47^{phox}$, and $p67^{phox}$. Other than a small number of patients in whom mutations in the promotor region of CYBB has been identified, nearly all CGD mutations involve one of the genes which encode these four subunits.

TABLE 65.2. CHRONIC GRANULOMATOUS DISEASE

NADPH Component	Gene Locus/Gene	Inheritance	% of Cases
$Gp91^{phox}$	Xp21.1/*CYBB*	X	65%
$P22^{phox}$	16p24/*CYBA*	AR	6%
$P47^{phox}$	7q11.23/*NFC1*	AR	23%
$P67^{phox}$	1q25/*NFC2*	AR	6%

AR, autosomal recessive; X, X-linked

CGD patients suffer from a variety of recurrent bacterial and fungal infections. These include pneumonia, lymphadenitis, skin infections, hepatic and perirectal abscesses, osteomyelitis, otitis media, conjunctivitis, gastrointestinal infections, and septicemia. CGD patients also may exhibit failure to thrive, anemia of chronic disease, chronic diarrhea, and hepatosplenomegaly. The bacteria and fungi that infect CGD patients have in common

the ability to produce catalase. It is believed that microbial catalase degrades H_2O_2 produced by the infecting microorganisms thus allowing for the elimination of nearly all oxidant accumulation within phagolysosomes. Common microbial isolates include *S. aureus*, Gram-negative enterics, *Aspergillus* sp., *Candida albicans*, *Burkholderia cepacia* (formerly *Pseudomonas*), *Serratia marcescens*, *Streptococcus* sp., and *Nocardia* sp.

Diagnosis requires recognition of both typical and atypical clinical features and demonstration of a defective phagocyte respiratory burst. Several assays that detect the ability of neutrophils to produce O_2- via NADPH oxidase have been developed. CGD subgroup classification is carried out commonly by analysis of immunoblot assays (all four NADPH oxidase subunits can be identified), cell-free NADPH oxidase functional assays, and by molecular assays (PCR, RFLP, and DNA sequencing).

Several hundred of polymorphic variants of glucose-6-phosphate dehydrogenase (G6PD) have been described. G6PD deficiency is the most common cause of hemolytic anemia caused by a red blood cell enzyme deficiency. Hemolytic episodes often are triggered by exposure to a drug that produces oxidative stress (e.g., antimalarials, sulfa drugs, etc.), ingestion of fava beans, or an infection. Severe G6PD deficiency in which patients present with a CGD-like syndrome have been described. G6PD, part of the hexose monophosphate shunt, is an X-linked heterodimer that catalyzes the generation of NADPH from glucose-6-phosphate and NADP+. The G6PD expressed in leukocytes is encoded by the same gene as the enzyme expressed in red cells, but is normally expressed at high enough concentrations that leukocyte dysfunction does not occur. Patients who suffer from recurrent infections possess very low concentrations of leukocyte NADPH and exhibit leukocyte enzyme activity levels that are usually less than 5% of normal. Male patients have been reported more often than females.

As in patients with “typical” CGD, patients with severe leukocyte G6PD deficiency suffer from recurrent coagulase-positive *S. aureus* infections, infections with *Escherichia coli*, and infections with *Serratia marcescens*. Sites of infection commonly include the lungs, skin, mucous membranes, and lymph nodes. These patients also suffer from hemolytic anemia attributable to red cell G6PD deficiency. As in the case of CGD, leukocytes from profoundly G6PD-deficient patients fail to generate O_2- or its metabolites.

Leukocyte Adhesion Deficiency

During the mid-1980s, several groups of investigators identified a small number of patients who suffered from recurrent bacterial infections and in whom neutrophils exhibited reduced spreading on surfaces, impaired aggregation, impaired iC3b-mediated cell function, and reduced antibody-dependent cellular cytotoxicity. This disorder, which can manifest as the so-called “moderate” and “severe” phenotypes initially, was named leukocyte adhesion deficiency. Following the later elucidation of another type of LAD (see below), the original disorder was renamed LAD1. It was soon recognized that leukocytes from LAD patients lack or are partially deficient of three surface molecules that are now known to be members of the ß2 or leukocyte integrin family. The integrins represent a superfamily of adhesion molecules each consisting of noncovalently associated α and ß chains. The leukocyte integrins possess distinct α chains and a common ß chain. This family includes CD11a/CD18, also known as lymphocyte function-associated antigen (LFA-1), CD11b/CD18 (Mac-1), and CD11c/CD18 (p150,95). CD11b/CD18 now is known to be complement receptor 3 (CR3). CD11b/CD18 and CD11c/CD18 are expressed on the lining surfaces of specific and tertiary neutrophil granules and are thus detected as an increase in cell surface density when these lysosomes fuse with the cytoplasmic membrane following cell activation. In contrast, CD11a/CD18 is expressed only on the surface of neutrophils and is therefore not “up-regulated” upon degranulation. CD11a/CD18 (LFA-1) is found in highest density on the surfaces of T, B, and NK cells, while CD11b/CD18 (Mac-1) is found in highest density on neutrophils, eosinophils, monocytes, macrophages, and NK cells. Patients with LAD1 possess mutations in the ß chain gene resulting in variable decreases in surface $ß_2$ integrin expression. Several different ß chain mutations have been described. LAD1 is inherited in an autosomal-recessive manner. The clinical result is a heterogeneous group of patients who exhibit different degrees of adhesion defect and variable severity of problems.

Patients with LAD typically suffer from recurrent bacterial infections without much pus, and with periodontitis, delayed wound healing, elevated circulating leukocyte counts, and a history of delayed shedding of the umbilical cord stump. Peripheral blood neutrophil counts typically range between 15×10^9/L and 60×10^9/L between infections and can reach 100×10^9/L or more during infections. Soft-tissue infections are particularly common and are often quite severe. *S. aureus*, *Pseudomonas* sp., and members of the enterobacteriaceae are the most common isolates. Diagnosis is made based on clinical data, demonstration of impaired adhesiveness and aggregation, and flow cytometric analysis.

LAD2 (Rambon-Hasharon syndrome) is an exceedingly rare cellular adhesion defect that results from abnormal fucosylation reactions.

Chédiak-Higashi Syndrome

Chédiak-Higashi syndrome is an autosomal-recessive disorder characterized by the presence of giant cytoplasmic granules in neutrophils, monocytes, and lymphocytes, as well as in some nonhematologic cell types. The genetic defect has not been completely characterized but appears to result from mutations in the recently discovered *CHS1*gene. *CHS1* encodes a protein that may be involved in organellar protein trafficking. The basic defect in this disorder is unknown, but involves multiple organ systems as evidenced by the occurrence of oculocutaneous albinism, defective platelets, and in some patients either sensory or motor neuropathies. Patients with Chédiak-Higashi syndrome typically have very light skin, photophobia, and characteristic silver-colored hair. The defects in pigmentation are the result of abnormal melanosome aggregation. Giant neutrophil granules can be seen in myeloid precursor cells. Because many myeloid series cells die before completion of maturation, Chédiak-Higashi patients typically are leukopenic ($2–3 \times 10^9$WBC/L). Because of the neutropenia and a host of neutrophil, monocyte, and lymphocyte functional defects, patients suffer from recurrent bacterial and fungal infections. The most common infections are of the skin, mucous membranes, and respiratory tract. The platelet defect is associated with decreased storage pools of ADP and serotonin, resulting in prolonged

bleeding times. Patients with Chédiak-Higashi syndrome can enter a so-called “accelerated phase” which is marked by local lymphocytic proliferations within the liver, spleen and bone marrow as well as a progressive pancytopenia. The pancytopenia of accelerated Chédiak-Higashi syndrome is accompanied by hemorrhage and infection. Among a variety of defects described in leukocytes from Chédiak-Higashi patients are abnormally fluid cell membranes, high resting rates of cellular respiration, impaired adherence to surfaces, and reduced chemotactic responsiveness. It is thought that altered membrane fusion relates to the formation of giant cytoplasmic granules, which in the case of neutrophils, contain reduced quantities of dilute lytic enzymes.

Other Phagocyte Function Disorders

The molecular basis of specific granule deficiency appears to be the result of a regulatory defect in hematopoietic gene expression. Neutrophils are either completely or nearly completely devoid of specific granules and their contents. Specific granules contain a variety of host defense proteins. These include lysozyme, lactoferrin, and Mac-1. Deficiency of Mac-1 results in defective adhesion to endothelium. In addition to its role in leukocyte-endothelial adhesion, Mac-1 also functions as the iC3b complement receptor. Neutrophils deficient of specific granules exhibit defective chemotaxis, adhesion, and microbicidal activities. Patients with specific granule deficiency suffer from recurrent and severe bacterial (*S. aureus*, *Pseudomonas aeruginosa*, Gram-negative enterics) and fungal (*Candida* sp.) infections of the mucous membranes, skin, and lower respiratory tract. The diagnosis, suggested by the history, patient examination, and abnormal neutrophil morphology, is confirmed by data from leukocyte functional assays (e.g., specific granule release assay) and flow cytometry [reduced or absent CD11b/CD18 (Mac-1) expression].

Myeloperoxidase (MPO) deficiency is a relatively common disorder (1 in 2,500 to 4,000 people) in which granulocytes produce O_2- and H_2O_2 but cannot generate normal levels of MPO-dependent hypohalite (HOC1) or HOC1-derived chloramines. Patients with complete MPO deficiency, with partial MPO deficiency, and with acquired MPO deficiency have been described. A variety of mutations that result in defective posttranslational processing of MPO have been described. Acquired partial MPO deficiency occurs in patients with acute myelogenous leukemia (M2, M3, and M4), in patients with chronic myelogenous leukemia, and in patients with myelodysplastic syndromes. *In vitro* leukocyte function assays reveal that affected cells exhibit delayed neutrophil-mediated microbial killing. Clinically affected patients, usually diabetics, suffer primarily from recurrent candidal infections. This disorder appears to involve other undefined factors because most MPO-deficient individuals are asymptomatic. Flow cytometric analysis may reveal an augmented respiratory burst, but diagnosis usually is made by cytochemical staining procedures.

Finally, several other exceedingly rare phagocyte function defects have been described. These include cyclic neutropenia, congenital neutropenias that include the Kostmann and Shwachman syndromes, and several less well-characterized defects.

Disorders of “Humoral” Immunity (B-Lymphocyte Defects)

As mentioned above, the first well-documented description of an immunodeficiency disorder was made in reference to a boy with agammaglobulinemia. Disorders of the “humoral” immune system are relatively common, considering that this group of disorders includes patients with common variable immunodeficiency (CVID) and selective IgA deficiency. This classification scheme is imperfect when one considers that there are pivotal interactions between antibody-producing cells (B lymphocytes and plasma cells) and T lymphocytes, as well as between the effector molecules of the humoral system (antibodies) and phagocytes (e.g., antibody-dependent cellular cytotoxicity) or complement components (e.g., C1q binding to IgG or IgM molecules). It should be emphasized that most deficiencies of γ-globulins are not primary and that they occur in the context of many disease states (e.g., protein-losing enteropathy, nephrotic syndrome, multiple myeloma, light chain disease, and secondary to immunosuppressive therapy). While the basic defect remains unknown in some humoral immunodeficiency disorders, specific genetic defects have been defined in an increasing proportion of these entities (Table 65.3). While many patients with immunoglobulin deficiencies are recognized because they suffer

from recurrent infections, individuals with some disorders experience infections no more frequently than normal (e.g., some patients with selective IgA deficiency).

TABLE 65.3. DEFICIENCIES THAT AFFECT ANTIBODY PRODUCTION

Disease	Mode of Inheritance	Locus/Gene
Agammaglobulinemia		
X-linked	XL	Xq21.3/*BTK*
Selective antibody class/subclass deficiencies		
γ1 isotype	AR	14q32.33
γ2 isotype	AR	14q32.33
Partial γ3 isotype	AR	14q32.33
γ4 isotype	AR	14q32.33
IgG subclass ± IgA deficiency	?	-
α1 isotype	AR	14q32.33
α2 isotype	AR	14q32.33
ε isotype	AR	14q32.33
IgA deficiency	Varied	-
Common variable immunodeficiency	Varied	-
Other	See text	-

Agammaglobulinemia

Bruton's or X-linked agammaglobulinemia is the first human primary immunodeficiency disease in which the genetic defect was definitively identified. In 1993 it was determined that mutations in a cytoplasmic tyrosine kinase, *Bruton's tyrosine kinase (BTK)*, lead to this clinical entity. *BTK* mutations result in a maturation arrest in B lymphocyte development. It should be noted that different *BTK* mutations give rise to disease phenotypes of varying clinical severity. A few male patients with agammaglobulinemia plus growth hormone deficiency have been described. Some of these patients have had normal *BTK* while others have possessed *BTK* mutations. Finally, a few females with the XLA phenotype have been described, but no *BTK* mutations have been defined in these patients.

Males with X-linked agammaglobulinemia are usually well until 6 to 9 months of age, the time by which maternal antibodies acquired *in utero* have been largely catabolized. Patients typically suffer from recurrent respiratory tract, joint, CNS, and systemic infections with high-grade encapsulated pyogenic bacteria. The spectrum of infections is broad, including pneumonia, sinusitis, otitis, meningitis, and sepsis. The most common bacterial isolates include *S. pneumoniae*, *H. influenzae*, and *Streptococcus* sp. Patients are particularly susceptible to viral hepatitis, and a number of patients have developed paralysis after exposure to live attenuated polio vaccine. A significant number of XLA patients also have died as the result of chronic echovirus infections of the central nervous system. This disorder typically is accompanied by a dermatomyositislike syndrome. A small percentage of patients with XLA develop neutropenia, which may be transient, cyclic, or persistent. *Pneumocystis carinii* pneumonia has been reported in neutropenic agammaglobulinemic patients. In contrast to patients with defective cellular immunity, growth retardation is uncommon as are fungal or other opportunistic infections.

Patients with XLA typically possess very low serum concentrations of γ-globulin (less than 100 mg/dL) and very low to undetectable concentrations of immunoglobulin in secretions. Analysis of peripheral blood typically reveals normal concentrations of total lymphocytes but very low numbers of B lymphocytes. While pre-B lymphoid series cells can be identified in the bone marrow of these patients, biopsies of lymphoid tissues reveal hypoplasia, the absence of germinal centers, and markedly diminished numbers of plasma cells. Patients with XLA are supported by assiduous attention to signs of infection, antibiotics, and immunoglobulin replacement therapy. The major long-term problem in patients who avoid complications like chronic viral CNS infections is the development of chronic and recurrent respiratory tract infections, which can lead to the development of bronchiectasis and pulmonary insufficiency.

Common Variable Immunodeficiency

Common variable immunodeficiency (CVID) is a heterogeneous group of diseases characterized by a deficiency of γ-globulins. Nearly all patients exhibit marked reductions in serum IgG and IgA levels and approximately 50% of patients also exhibit a marked reduction in IgM. This group of disorders manifests in a manner that is somewhat similar clinically to X-linked agammaglobulinemia except that it occurs with equal frequency in both males and females and it can present at different ages. While many patients present in infancy or early childhood, most common variable immunodeficiency patients present during later childhood or adulthood. In general the γ-globulin deficiency is less severe than that seen in X-linked agammaglobulinemia, typically on the order of 200 to 300 mg/dL of total immunoglobulins. In addition to pyogenic infections, CVID has been associated with a variety of autoimmune and neoplastic disorders. Patients with CVID suffer from a markedly increased incidence of giardiasis and bronchiectasis. In addition to autoimmune and infectious complications, these patients suffer from an increased incidence of gastric carcinoma, lymphomas, and amyloidosis.

The defects responsible for common variability immunodeficiency appear to be heterogeneous. Patients with CVID typically have normal numbers of circulating B cells but they don't differentiate normally *in vivo* or become immunoglobulin-producing plasma cells *in vitro*. Indeed, there are defects in terminal B lymphocyte maturation, a process that is now better understood in the context of interleukins. Interestingly, first-degree relatives of CVID patients have an increased incidence of selective IgA deficiency, and vice versa, suggesting the possibility that these disorders have a related pathogenetic origin. In fact, some patients with selective IgA deficiency eventually develop CVID.

Patients with common variable immunodeficiency are treated with antibiotics and immunoglobulin replacement therapy.

IgA Deficiency

Selective IgA deficiency is marked by the presence of very low serum concentrations of IgA accompanied by diminished IgA concentrations in mucosal secretions. Selective IgA deficiency is the most common primary immunodeficiency, with a prevalence of between 1:350 and 1:3000. IgA deficiency generally is associated with an increased incidence and severity of respiratory, genitourinary, and gastrointestinal tract infections as well as an increased likelihood of autoimmune and lupuslike syndromes. Patients with this immunodeficiency may also develop a spruelike malabsorption syndrome, which often responds favorably to a gluten-free diet. Interestingly, not all IgA-deficient individuals manifest such problems. It appears that many, but not all, clinically ill IgA-deficient patients suffer from concomitant deficiencies of IgG subclasses, most commonly subclasses IgG2 and IgG4. The most common offending organisms are pyogenic bacteria. Some IgA-deficient patients go on to develop CVID, and some IgA deficient patients have spontaneously remitted.

The basic defect responsible for selective IgA deficiency is unknown but appears to be a regulatory defect since α1 and/or α2 heavy-chain gene deletions have been identified in only a few patients. A number of drugs also have been implicated in the induction of selective IgA deficiency. Cell surface analysis of circulating lymphocytes reveals that IgA-bearing B lymphocytes also

express IgM and IgD, an abnormal finding that suggests that such cells are maturation-arrested. Treatment of IgA-deficient patients is based on the severity of clinical illness. Antibiotics are the mainstay of therapy. Great care to avoid sensitization of IgA deficient patients with IgA-bearing blood products should be practiced because such patients can experience anaphylactic reactions following exposure to IgA-containing blood products. Unfortunately, the presence of IgE antibodies directed against IgA does not absolutely predict a predisposition to an untoward reaction following receipt of blood products. If blood products are absolutely required, IgA-poor preparations should be used. Each patient must be considered at risk and observed closely when receiving blood products.

Selective Antibody Class and Subclass Deficiencies

Many individuals harbor IgG subclass (IgG1, IgG2, IgG3, IgG4) deficiencies. The total serum IgG concentration may be normal or abnormal depending on which subclass is deficient and on the severity of the defect. For example, because nearly 70% of the total serum IgG is IgG1, a marked IgG1 subclass deficiency would be reflected in a decreased total IgG concentration. In contrast, because IgG4 accounts for only a small percentage (less than 1% to 2%) of the total IgG, a complete absence of IgG4 would not be reflected by measurements of total IgG concentration. Serum IgA and IgM leads usually are normal, although perhaps one fifth of IgA-deficient patients have a concomitant IgG subclass deficiency. Some patients mount normal antibody responses following immunization while others do not. Clinically, IgG subclass-deficient patients fall within a spectrum ranging from no symptoms to recurrent pyogenic infections, typically involving the paranasal sinuses or the lower respiratory tract. The most common isolates include *S. pneumoniae*, *H. influenzae*, and *S. aureus*. The diagnosis for IgG subclass deficiency sometimes is not made for years. The primary defects(s) are not well understood. Therapy includes appropriate antibiotics and γ-globulin replacement therapy.

Other Deficiencies of Antibody Production

Specific antibody deficiency, also known as “antibody deficiency with near-normal immunoglobulins,” is rare. These patients have defective antibody responses to certain antigens despite normal serum immunoglobulin concentrations and normal T lymphocyte function in available laboratory assays. Patients exhibit selectively subnormal antibody responses to various antigens, including pneumococcal polysaccharide vaccines, bacteriophage φX174, tetanus toxoid, and diphtheria toxoid. Most patients also are devoid of blood group antibody titers. The primary defect in selective antibody deficiency is unknown. Current therapy includes γ-globulin replacement and antibiotics.

Only a few cases of selective IgM deficiency have been reported. A variety of serious infectious, including meningococcemia, disseminated enterobacteriaceae infections, pneumococcal meningitis, staphylococcal soft-tissue infections, and respiratory tract infections have been associated with IgM deficiency. The specific defect responsible for selective IgM deficiency is unknown. Treatment includes antibiotic therapy.

Transient hypogammaglobulinemia of infancy is relatively rare and, as its name suggests, manifests with recurrent infections and hypogammaglobulinemia. In contrasts to patients with common variable immunodeficiency or X-linked agammaglobulinemia, infants with this disorder possess isohemagglutinins and can respond to administered antigens such as diphtheria and tetanus toxoid. Such specific antibody responses usually can be detected before serum immunoglobulin concentrations reach normal levels. Because of the transient nature of this disorder and the potential for suppression of endogenous antibody production, replacement γ-globulin is not recommended.

Combined Humoral and Cellular Immune Dysfunction

The combined humoral and cellular immunodeficiency disorders include syndromes that vary greatly in severity and in the spectrum of defects that give rise to individual diseases (Table 65.4). For instance, patients with adenosine deaminase (ADA) deficiency suffer from a profound humoral and cellular immunodeficiency as well as skeletal abnormalities while patients with X-linked hyper-IgM syndrome suffer from recurrent pyogenic bacterial infections and exhibit pronounced lymphadenopathy. As noted for humoral immunodeficiency disorders, classification of combined disorders can be problematic. Leukocyte adhesion deficiency (see earlier discussion) affects both lymphocyte and neutrophil-mediated host defense processes and could be considered either a disorder of phagocyte function or a combined cellular and humoral defect. In an increasing number of disorders, the specific defect(s) responsible for combined humoral and cellular disorders has been defined (Table 65.4). Clinically, patients with combined humoral and cellular defects may present with bacterial infections reminiscent of pure humoral deficiencies, opportunistic infections typically associated with cellular immune dysfunction, or features unique to given disorders. Examples of the latter include skeletal (costochondral) abnormalities in ADA deficiency, the characteristically extensive erythematous scaly rash in Omenn syndrome, and dwarfism in cartilage-hair hypoplasia.

TABLE 65.4. COMBINED HUMORAL AND CELLULAR IMMUNODEFICIENCIES

Disease	Locus/Gene	Inheritance[a]
Severe combined immunodeficiency (SCID)		
T^-B^- SCID		
Omenn syndrome	11p13/*RAG1, RAG2*	AR
RAG1 deficiency	11p13/*RAG1*	AR
RAG2 deficiency	11p13/*RAG2*	AR
Reticular dysgenesis	-	AR
T^-B^+ SCID		
Jak3 deficiency	19p13.1/*JAK3*	AR
X-linked γc-chain	Xq13.1-q13.3	X
Abnormal purine metabolism		
Adenosine deaminase (ADA) deficiency	20q13.2-q13.11	AR
Purine nucleoside phosphorylase (PNP) deficiency	14q13.1	AR
Hyper-IgM syndrome		
X-linked (CD40L deficiency)	Xq26.3-q27.1	X
Non-X-linked	-	-
Major histocompatibility complex deficiencies		
MHC class I deficiency	6q21.3/*TAP2*	AR
MHC class II deficiencies	Multiple	AR
Other combined immunodeficiencies		
CD3 deficiencies	11q23/*CD3E, CD3G*	AR
IL-2 receptor α-chain deficiency	10p14-p15/*IL2RA*	AR
ZAP-70 deficiency	2q12/*ZAP70*	AR

AR, autosomal recessive; X, X-linked

Severe Combined Immunodeficiency

While SCID vary in agenetic etiology, mode of inheritance, lymphoreticular histopathology, mechanism of cellular dysfunction, and to some extent, clinical characteristics, they have in common the near complete absence of cellular and humoral immune function. As a result, affected patients are profoundly ill and die within a short period of time, unless they can be isolated from the environment, their immune systems can be replaced by bone marrow transplantation, or in appropriate instances, specific enzyme replacement therapy can be instituted. The SCID disorders present within the first months of life and are marked by recurrent episodes of upper respiratory, lower respiratory, and cutaneous bacterial infections. Persistent opportunistic infections caused by *Candida*, measles, varicella, cytomegalovirus, and *Pneumocystis* also are prevalent. Patients with SCID suffer extreme wasting and are at risk of developing graft-versus-host (GVH) disease.

Clinical and laboratory evaluation of immune function reveal profound abnormalities. There is a virtual absence of cellular immune function manifested by lymphopenia, cutaneous anergy, and the absence of *in vitro* lymphocyte proliferative responses to antigens, mitogens, or allogeneic cells. Serum immunoglobulin concentrations usually are markedly depressed, and antibody responses to immunization are difficult to detect. A small number of autosomal recessive SCID patients in whom nearly all circulating lymphocytes appear to be natural killer (NK) cells have been reported.

As shown in Table 65.4, SCID can be subclassified into T^-B^- and T^-B^+ categories based on immunophenotypic data. While many fundamental questions remain, a great deal has been learned about the regulation of T lymphocytes and B lymphocytes antigen receptor gene recombination and expression [including V(d)J recombination]. It has become clear that defects in the recombination mechanism can lead to SCID. Defects in *RAG1* and *RAG2*, which encode "recombination activating genes 1 and 2", respectively, have been specifically linked to SCID. Patients with RAG1 and RAG2 SCID do not possess their own circulating T or B lymphocytes. Omenn syndrome is a T^-B^- severe combined immunodeficiency that, in addition to the combined system immunodeficiency is characterized by generalized erythroderma, eosinophilia, elevated serum IgE, and hepatosplenomegaly. These patients also harbor *RAG* mutations. Reticular dysgenesis (RD) is not well understood but reflects failure of bone marrow stem cell production. Patients with RD are profoundly lymphopenic and granulocytopenic.

Autosomal-recessive and X-linked recessive SCID patients present with recurrent or persistent bacterial, fungal, protozoal, and viral infections accompanied by wasting. It has become clear in recent years that X-linked SCID is caused by mutations in the gene that encodes the common γ chain of the IL-2, IL-4, IL-7, IL-9, and IL-15 receptors. Autosomal-recessive SCID patients are indistinguishable clinically and immunophenotypically from X-linked SCID patients. Autosomal-recessive T^-B^+ SCID patients harbor mutations in the gene that encodes JAK3, a tyrosine kinase that transduces the interleukin γ chain receptors noted above. Care must be taken not to subject these patients to GVH disease. GVH disease may develop following receipt of blood products contaminated with variable lymphocytes and can even occur as the result of maternal-fetal hemorrhage *in utero*. Replacement γ-globulin has not been effective in preventing the demise of SCID patients. Some patients have been immunologically reconstituted with HLA-identical or D locus-compatible bone marrow transplantation. Recent progress also has been achieved by reconstituting SCID patients with haploidentical allogeneic bone marrow purged of postthymic T lymphocytes.

Immunodeficiencies Associated with Abnormal Purine Metabolism

Adenosine deaminase (ADA) and purine nucleoside phosphorylase (PNP) are sequential enzymes in the purine salvage pathway. Dysfunction of in either of these enzymes can lead to accumulation of toxic intermediates that cause the loss of lymphocytes and can result in SCID. ADA deficiency accounts for approximately 40% of all autosomal recessive SCID cases and 20% of all SCID cases. PNP deficiency is less common. ADA-deficient SCID patients suffer from the same infections as other SCID patients but also may exhibit skeletal abnormalities. Deficiency of ADA results in the accumulation of adenosine, 2′-deoxyadenosine, and 2′-O-methyladenosine, which are toxic to lymphocytes. It appears the adenosine and deoxyadenosine inactivate the enzyme,

S-adenosylhomocysteine (SAH) hydrolase, which in turn is an inhibitor of most intracellular methylation reactions. Reduction of SAH hydrolase activity leads to excessive methylation and lymphocytotoxicity. Patients with ADA-deficient SCID typically become lymphopenic within the first year of life. ADA-deficient SCID patients have been treated successfully with bone marrow from HLA-matched siblings and allogeneic haploidentical T-lymphocyte-purged bone marrow. An approach that has also been successful in some patients is ADA enzyme replacement either through infusion of irradiated adenosine deaminase-containing red blood cells or infusion of polyethylene glycol-linked bovine ADA. Autosomal-recessive SCID with ADA deficiency also is a prime candidate for gene-replacement therapy. PNP deficiency results in a profound T lymphocyte defect and a variable B lymphocyte defect. These SCID patients also suffer frequently from neurologic abnormalities and autoimmune disorders.

Hyper-IgM Syndrome

Immunodeficiency with hyper-IgM is a heterogeneous group of disorders in which patients suffer from recurrent bacterial and opportunistic infections and exhibit low serum concentrations of IgG, IgA, and IgE. Lymph nodes from hyper-IgM patients display abnormal architecture that includes a lack of germinal centers. In most cases hyper-IgM is X-linked, but autosomal recessive, autosomal dominant, and acquired forms have been described. During the 1990s it has become clear that most hyper-IgM cases are raised by a mutation in the CD40L gene which encodes CD40 ligand, a transiently expressed T lymphocyte surface molecule that, in the presence of appropriate cytokines, induces immunoglobulin production in B lymphocytes via cell surface CD40. Defective signaling via B lymphocyte CD40 also can result in a hyper-IgM syndrome.

Patients with this disorder typically present in infancy with recurrent pyogenic infections including pneumonia, sinusitis, otitis media, and tonsillitis. In contrast to patients with X-linked agammaglobulinemia, these patients have pronounced lymphadenopathy. Serum IgA and IgG concentrations are markedly depressed, and there is typically a polyclonal increase in serum IgM. Patients with immunodeficiency with elevated IgM occasionally have concomitant neutropenia and may manifest autoimmune hemolytic anemia and/or thrombocytopenia.

Major Histocompatibility Complex Deficiencies

Only a few cases of HLA class I deficiency (type 1 bare lymphocyte syndrome) have been described. In these cases, very low levels of HLA class I molecules can be detected. (Complete HLA class I deficiency has not been described and may be presumed to be lethal.) HLA class I molecule deficiency is not associated uniformly with an increase in susceptibility to infections. Clinically affected patients have typically suffered from chronic bacterial infections of the respiratory tract and in a couple of cases, vasculitis. Several HLA class I-deficient individuals without disease have been reported. Several lesions can give rise to HLA class I molecule deficiency. Studies to date have revealed constitutive decreases in class I HLA gene expression in lymphocytes and mutations in TAP2, a gene that encodes a subunit of TAP (a transporter associated with antigen processing), a protein involved in the transportation into the endoplasmic reticulum of peptide antigen fragments that have been endogenously degraded.

Major histocompatibility complex class II (MHC II) deficiency results in a profound combined immunodeficiency fungal and protozoal infection, which primarily involves the respiratory and gastrointestinal tracts. As in the case of HLA class I deficiency, this syndrome also is referred to as the “bare lymphocyte syndrome” (BLS). One classification scheme utilizes the designations BLS type I, BLS type II, and BLS type III (for classes I and II). MHCII deficiency is an autosomal-recessive disease that is caused by mutations in at least three different transcription factors (*RFX5, RFXAP*, and *MHC2TA*), which regulate HLA class II molecule expression.

Other Combined B- And T-Cell Immunodeficiencies

Additional combined immunodeficiencies include mutations in *CD3E* and *CD3G*, genes that encode CD3ε and CD3γ, respectively. CD3ε and CD3γ are proteins that are required for T lymphocyte receptor (TCR) function. Some CD3-deficient individuals have been healthy while others have suffered from recurrent lower respiratory tract infections, intractable diarrhea, and failure to thrive.

IL-2 receptor α-chain deficiency and ZAP-70 deficiency, the latter the result of a mutation in the T cell receptor zeta-chain associated protein kinase gene (*ZAP70*), are exceedingly rare immunodeficiency syndromes.

Miscellaneous Well-Defined Immunodeficiency Disorders

Wiscott-Aldrich Syndrome (and X-Linked Thrombocytopenia)

Wiskott-Aldrich syndrome (WAS), also known as immunodeficiency with thrombocytopenia and eczema, is an X-linked disorder marked by cellular and humoral immunodeficiency, thrombocytopenia, eczema, and an increased risk of hematopoietic malignancies. Mutations responsible for both WAS and X-linked thrombocytopenia (XLT) (see below) have been mapped to Xp11.22. The *WAS* gene encodes the WAS protein (WASP), which is expressed in all hematopoietic stem-cell–derived lines. It is believed to be important in signal transduction and the regulation of cytoskeletal function. Boys with WAS sometimes present with a thrombocytopenic bleeding diathesis heralded by prolonged bleeding following circumcision. Skin lesions and recurrent infections typically occur during the first year. Early in the course of the disease patients exhibit impaired humoral responses to encapsulated high-grade bacterial pathogens resulting in otitis media, pneumonia, meningitis, and sepsis. Later, patients suffer from herpes infections and *Pneumocystis* infections. WAS carries a grave prognosis and most patients die during childhood. There is an increased incidence of lymphoreticular malignancy.

Thrombocytopenia is accompanied by normal numbers of

megakaryocytes in the bone marrow. Immunologic abnormalities reflect the clinical course of the disorder. Patients typically have borderline or low immunoglobulin levels with depressed serum IgM levels, low or low-normal IgG levels, and elevated serum IgA and IgE levels. Defective humoral immune responses to carbohydrate antigens and depressed isohemagglutinin levels are apparent early. Later, poor humoral responses to protein antigens become apparent. Cutaneous anergy and decreased concentrations of peripheral blood T lymphocytes are also present.

Treatment, including platelet transfusions, splenectomy, and γ-globulin therapy, generally has been supportive. Some recent successes have been achieved with stem cell transplantation.

Hereditary XLT, a less severe phenotypic variant of WAS, is characterized by thrombocytopenia, mild or no eczema, and moderate or no immunologic abnormalities. It appears that XLT is caused by different mutations in the WAS gene.

DiGeorge Syndrome

The DiGeorge syndrome is the result of defective embryogenesis involving the 3rd and 4th pharyngeal pouches. The recognition of partial forms of DiGeorge syndrome and the discovery that approximately 90% of cases are caused by a deletion of chromosomal region 22q11 have led to the recognition that this disorder is relatively common (1:3000 to 1:4000 live births) among primary immunodeficiency disorders. It appears that there is a common genetic abnormality and phenotypic overlap with velocardiofacial syndromes and conotruncal anomaly face syndrome. It is believed that the DiGeorge syndrome reflects a defect in neural crest cell migration into the third and fourth pharyngeal pouches. The result is hypoplasia or aplasia of the thymus and defective T-lymphocyte-dependent immunity. Patients with DiGeorge's syndrome usually suffer from defects involving other structures derived from the third and fourth branchial arches including the parathyroid glands, the great vessels, esophagus, cardiac septa, and several facial structures. The clinical results of these defects include neonatal hypocalcemic tetany secondary to agenesis of the parathyroid glands, persistent right-sided aortic arch, esophageal atresia, atrial and ventricular septal defects, and facial defects such as hypertelorism, short philtrum, low-set ears, and mandibular hypoplasia.

The degree of thymic hypoplasia is variable. Some patients manifest only limited defects in cellular immunity and thus suffer from few infectious complications. At the other end of the spectrum are patients with thymic aplasia who may present like patients with severe combined immunodeficiency. These patients suffer from recurrent life-threatening viral, fungal, and protozoal infections, and may develop GVH disease after receipt of blood products.

Patients with DiGeorge's typically have decreased numbers of circulating T lymphocytes with normal immunoglobulin concentrations. Some patients have decreased serum IgA and may have increased serum IgE concentrations. *In vitro*, the response of peripheral blood lymphocytes to T-lymphocyte mitogens usually is depressed. Thymic tissue is generally atrophic, although in some cases small tissue rests containing Hassall's corpuscles and thymocytes can be identified. Lymph nodes and splenic tissue typically show depleted T-cell-dependent areas. Therapy is supportive. Some groups have reported benefit from fetal and more recently, postfetal, thymus transplantation.

HYPER-IgE SYNDROME (JOB'S SYNDROME)

Hyper-IgE syndrome is an autosomal-dominant disorder of unknown cause. Patients typically suffer from recurrent staphylococcal abscesses involving the skin, lungs, joints, and soft tissues. Some patients with hyperimmunoglobulinemia E syndrome also suffer from a generalized dermatitis. The so-called Job's syndrome is considered to be related. The Job syndrome is characterized by recurrent skin abscesses in association with a poorly characterized neutrophil motility defect.

Patients with hyperimmunoglobulinemia E typically exhibit normal serum immunoglobulin concentrations except for IgE, which is typically markedly increased to concentrations in excess of 2000 IU/mL and as high as 20,000 to 50,000 IU/mL. In addition, peripheral blood eosinophilia is characteristic. Some patients have neutrophil and/or monocyte chemotactic defects. The mainstay of therapy is antibiotic treatment.

Chronic Mucocutaneous Candidiasis

Chronic mucocutaneous candidiasis (CMC) is a syndrome marked by chronic candidal infections of the skin and mucous membranes. In contrast to the other cellular immune disorders, CMC is rarely fatal. The cause of this disorder is unknown. In some cases, no immunodeficiency cause can be identified. Some patients have accompanying endocrinopathies involving the parathyroid, thyroid, adrenal glands or pancreas. Ketoconazole is the most efficacious therapy reported to date.

X-Linked Lymphoproliferative Syndrome (Duncan Disease)

X-linked lymphoproliferative disease (XLP) was initially described in a group of patients from a single kindred, the Duncan family. Recent studies have linked the syndrome to defects in the XLP (*LYP*) gene located at Xq25. Manifestations (see below) have been reported in individuals who harbor an abnormal XLP (*LYP*) gene but have never been exposed to the Epstein-Barr virus (EBV). This observation indicates that the disease is not a specific sequel to EBV infection (except in cases of fatal EBV-induced infectious mononucleosis). The syndrome is varied in its clinical manifestations but has in common an abnormal response to EBV infections. Patients with X-linked lymphoproliferative disease are well until they first suffer from an EBV infection. Nearly 70% of patients die as the result of an intense lymphocyte proliferation that occurs during mononucleosis. Most survivors develop hypogammaglobulinemia, B-cell lymphoma or a combination of both. Most patients who survive mononucleosis do not mount a normal antibody response to EBV nuclear antigen, whereas titers to the viral capsid antigen may range from zero to markedly elevated. There is often an accompanying depression in T-cell immunity to EBV. Nonspecific immunologic defects also have been associated with Duncan's disease. Such patients may have normal numbers of circulating B and T lymphocytes

with an elevated percentage of CD8+ (T-suppressor) cells. Finally, *in vitro* lymphocyte immunoglobulin synthesis to B-cell mitogens is also often blunted. There is currently no specific effective therapy for this disorder.

Cartilage — Hair Hypoplasia

Cartilage-hair hypoplasia is an autosomal-recessive disorder with variable penetrance. It has been linked to mutations in the *CHH* located at chromosome 9p13. This disorder is prevalent particularly among the Amish and is characterized by short-limbed dwarfism accompanied by frequent, often life-threatening infections. Patients with this disorder possess fine light hair and eyebrows, hyperextensible joints of the hands and feet, and a characteristic short-limbed dwarfism that is particularly pronounced in the hands. While defective antibody mediated immunity is demonstrable in most patients, there also is defective cellular immunity and a markedly increased propensity for life-threatening infections. Progressive vaccinia and vaccine-associated poliomyelitis have been observed in some patients.

DNA Breakage-Associated Syndromes

Ataxia-telangiectasia (A-T) is an autosomal-recessive disorder that was first recognized in 1926. This entity is characterized by progressive neurological abnormalities (most commonly cerebellar ataxia, choreoathetosis, and oculomotor abnormalities), ocular and cutaneous telangiectasias, immunodeficiency, growth retardation, and sensitivity to ionizing irradiation. A-T was mapped to chromosome 11q22-23 in 1988 and the ATM (ataxia-telangiectasia mutated) gene identified and cloned in 1995. *ATM* is a phosphatidylinositol -3-kinase which is involved in cell cycle control, cellular response to DNA damage, and intracellular protein transport.

Typical patients manifest progressive cerebellar ataxia within the first 12 to 18 months of life, and most patients suffer from recurrent bacterial sinopulmonary infections. Telangiectasias typically become apparent on the conjunctiva and skin early in childhood. In addition to disordered immune function, ataxia, and telangiectasias, patients with ataxia-telangiectasia are at risk of developing malignancies involving the hematopoietic system.

Both patients and heterozygotic carriers of this disorder exhibit defective DNA repair accompanied by inordinate sensitivity to ionizing irradiation. *In vitro*, B lymphocytes often exhibit an intrinsic inability to synthesize IgA.

The immunologic profile in ataxia-telangiectasia patients is variable. Most patients exhibit decreased serum IgA and IgE concentrations accompanied by variable decreases in serum IgG. Cellular immune function usually is depressed but not to the degree observed in severe combined immunodeficiency disorders or complete cases of the DiGeorge syndrome.

Chromosomal instability syndromes other than ataxia-telangiectasia include Fanconi anemia, Bloom syndrome, and ataxia-telangiectasia variants (AT-V). The latter classification was proposed in the late 1980s, for the group of patients who exhibit cytogenetic abnormalities identical to those seen in A-T patients but who display different clinical features. AT-V patients suffer from microcephaly and do not develop ataxia (which occurs in 100% of A-T patients). Some investigators have subclassified AT-V into two groups; AT-V1 (Nijmegen breakage syndrome) and AT-V2 (Berlin breakage syndrome). This rationale for this subdivision has come into recent question and is thus controversial.

Interferon-γ-Associated and Interleukin-12-Associated Immunodeficiencies

Mutations that result in interferon-γ receptor defects, an interleukin-12 (IL-12) deficiency, and an IL-12 receptor defect have been defined recently. These include mutations in the *IFNGR1*, *IFNGR2*, *IL12B*, and *IL12RB1* genes, which encode subunits of the IFNγ1-receptor, the IFNγ2-receptor, IL-12(p40 subunit), and the IL-12 receptor ß1, respectively. All of these patients exhibit a marked increase in susceptibility to disseminated mycobacteriosis. This is a particular issue in locales where live *Bacillus Calmette-Guerin* is administered as a vaccine.

The great advances that have occurred in understanding of the genetic and molecular bases for primary immunodeficiency disorders have important implications in the classification, diagnosis, and management of such patients. From a practical standpoint it remains important to recognize the typical as well as unusual clinical features of immunodeficiency diseases and to judiciously apply appropriate testing approaches to their diagnosis. The impact that these advances have had in deepening our understanding of the biology of the immune system cannot be overstated.

SUGGESTED READING

1. Ochs HD, Smith CIE, Puck JM, eds. *Primary Immunodeficiency Diseases. A Molecular and Genetic Approach*, 1st ed. New York: Oxford University.

2. Gelehrter TD, Collins FS, Ginsburg D. *Principles of Medical Genetics*, 2nd ed. Baltimore: Williams & Wilkins, 1998.

3. Conley ME, Stiehm ER. Immunodeficiency disorders: general considerations. In: Stiehm ER, ed. *Immunologic Disorders in Infants and Children*. Philadelphia: W.B. Saunders, 1996:201–252.

4. Rose NR, de Macario EC, Folds JD, et al., eds. *Manual of Clinical Laboratory Immunology*. Washington, D.C.: ASM Press, 1997.

5. Winkelstein JA, Sullivan KE, Colten HR. Genetically determined deficiencies of complement. In: Scriver CR, Beaudet AL, Slu WS, et al., eds. *Metabolic Basis of Inherited Disease*, vol. 2. New York: McGraw-Hill, 1995.

6. Slingsby JH, Norsworthy P, Pearce G, et al. Homozygous hereditary C1q deficiency and systemic lupus erythematosus. *Arthritis Rheum* 1996;39:663–670.

7. Lee SL, Wallace SL, Barone R, et al. Familial deficiency of two subunits of the first component of complement: C1r and C1s associated with a lupus erythematosus-like disease. *Arthritis Rheum* 1978;21:959–967.

8. Truedsson L, Sturfelt G, Nived O, et al. Prevalence of the type I complement C2 deficiency gene in Swedish systemic lupus erythematosus patients. *Lupus* 1993;2:325–327.

9. Sim RB, Kolble K, McAleer MA, et al. Genetics and deficiencies of the soluble regulatory proteins of the complement system. *Int Rev Immunol* 1993;10:65–86.

10. Oltvai ZN, Wong ECC, Atkinson JP, et al. C1 inhibitor deficiency: molecular and immunologic basis of hereditary and acquired angioedema. *Lab Invest* 1991;65:381–388.

11. Lublin DM, Mallinson G, Poole J, et al. Molecular basis of reduced or absent expression of decay-accelerating factor in Cromer blood group phenotypes. *Blood* 1994;84:1276–1282.

12. Sullivan KE, Wooten C, Goldman D, et al. Mannose binding protein genetic polymorphisms in black patients with systemic lupus erythematosus. *Arthritis Rheum* 1996;39:2046–2051.

13. Thompson RA. Immunodeficiency due to defects of polymorphonuclear leukocyte function. *Immunol Invest* 1988;17:85–92.

14. Cross AR, Curnutte JT, Rae J, et al. Hematologically important mutation: X-linked chronic granulomatous disease. *Blood Cells Mol Dis* 1996;22:90–95.

15. Cross AR, Curnutte JT, Rae J, et al. Hematologically important mutations: the autosomal recessive forms of chronic granulomatous disease. *Blood Cells Mol Dis* 1996;22:68–270.

16. Anderson DC, Kishimoto TK, Smith CW. Leukocyte adhesion deficiency and other disorders of leukocyte adherence and motility. In: Scriver CR, Beaudet AL, Sly WS, et al., eds. *The Metabolic and Molecular Basis of Inherited Diseases*, 7th ed., New York: McGraw-Hill, 1995:3955–3995.

17. Etzioni A, Harlan JM, Pollack S, et al. Leukocyte adhesion deficiency (LAD) II: a new adhesion defect due to absence of sialyl lewis x, the ligand for selectins. *Immunodeficiency* 1993;4:307–308.

18. Hernaszewski RA, Webster ADB. Primary hypogammaglobulinemia: a survey of clinical manifestations and complications. *QJ Med* 1993;86:31–42.

19. Mattson PT, Vihinen M, Smith CIE. X-linked agammaglobulinemia (XLA): a genetic tyrosine kinase (Btk) disease. *Bioessays* 1996;18:825–834.

20. Paller AS. Immunodeficiency syndromes. X-linked agammaglobulinemia, common variable immunodeficiency. Chediak-Higashi syndrome, Wiskon-Aldrich syndrome, and X-linked lymphoproliferative disorder. *Dermatol Clin* 1995;13:65–71.

21. Ashman RF, Schaffer FM, Kemp JD, et al. Genetic and immunologic analysis of a family containing five patients with common-variable immune deficiency or selective IgA deficiency. *J Clin Immunol* 1992;12:406–414.

22. Fischer A. Severe combined immunodeficiencies. *Immunodefic Rev* 1992;3:83–100.

23. Fischer A, Malissen B. Natural and engineered disorders of lymphocyte development. *Science* 1998;280:237–243.

24. Buckley RH, Schiff RI, Schiff SE, et al. Human severe combined immunodeficiency (SCID): genetic phenotypic and functional diversity in 108 infants. *J Pediatr* 1997;130:378–387.

25. Schwarz K, Gauss GH, Ludwig L, et al. RAG mutations in human B cell-negative SCID. *Science* 1996;274:97–99.

26. Macchi P, Villa A, Giliani S, et al. Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency (SCID). *Nature* 1995;377:65–68.

27. Puck JM, de Saint Basile G, Schwarz K, et al. IL2RGbase: a database of γc-chain defects causing human X-SCID. *Immunol Today* 1996;17:507–511.

28. Hershfield MS, Mitchell BS. Immunodeficiency diseases caused by adenosine deaminase deficiency and purine nucleoside phosphorylase deficiency. In: Scriver CM, Beaudet AL, Sly W, et al., ed. *Metabolic Basis of Inherited Disease*, 7th ed. New York: McGraw Hill, 1995: 1725–1768.

29. Allen RC, Armitage RJ, Conley ME, et al. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. *Science* 1993;259:990–993.

30. Elder ME, Lin D, Clever J, et al. Human severe combined immuno-deficiency due to a defect in ZAP-70, a T cell tyrosine kinase. *Science* 1994;264:1596–1599.

31. Derry JMJ, Kerns JA, Weinberg KI, et al. WASP gene mutations in Wiskott-Aldrich syndrome and X-linked thrombocytopenia. *Hum Mol Genet* 1995;4:1127–1135.

32. Driscoll DA, Budarf ML, Emanuel BS. A genetic etiology for Di-George syndrome: consistent deletions and microdeletions of 22q11. *Am J Hum Genet* 1992;50:924–933.

33. Burn J, Takao A, Wilson D, et al. Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22. *J Med Genet* 1993;30:822–824.

34. Seemayer TA, Gross TG, Egeler RM, et al. X-linked lymphoproliferative disease: twenty-five years after the discover. *Pediatr Res* 1995;38:471–478.

35. Hill HR, Augustine NH, Alexander G, et al. Familial occurrence of Job's syndrome of hyper-IgE and recurrent infections. *J Allergy Clin Immunol* 1997;995:395.

36. Meyn MS. Ataxia-telangiectasia and cellular responses to DNA damage. *Cancer Res* 1995;55:5591–6001.

37. Casanova J-L, Jouanguy E, Lamhamedi S, et al. Immunological conditions of children with disseminated BCG infection. *Lancet* 1995;346:581.

66
Allergic Conditions

Eric Brestel

Donald R. Hoffman

TYPES OF IMMUNOPATHOLOGICAL HYPERSENSITIVITY

Part of "66 - Allergic Conditions"

There are several mechanisms by which the immune response can initiate a pathological process (Table 66.1). Coombs and Gell were the first to propose this classification scheme, and it has been updated by Snyderman (1). The type I response is initiated by the reaction of allergen with IgE bound to tissue mast cells via the Fc epsilon receptor. This antigen recognition by IgE results in a transmembrane signal that initiates the release of mediators from the mast cell. These include histamine, leukotrienes, platelet-activating factor, prostaglandin D2, and a variety of intracellular enzymes, including a specific tryptase. These mediators can result in an immediate (within 5 to 30 minutes) response depending on the route of allergen exposure: a wheal-and-flare in the skin, rhinitis, asthma, vomiting and diarrhea, or systemic anaphylaxis. IgE-mediated responses also can cause systemic anaphylaxis and gastrointestinal allergic reactions manifested by vomiting, crampy abdominal pain, and diarrhea. The IgE-mediated response occurs within several minutes following exposure to an allergen and may gradually subside over 60 minutes. A secondary exacerbation of this response, however, may occur several hours after the initial mast cell degranulation and has been called the late response. This late response is felt to be a key component of chronic rhinitis and asthma and also can be seen in the skin following insect sting reactions and after the injection of allergens to which the patient is sensitive. The late response is caused by mediators released by eosinophils and lymphocytes, which are recruited to the site by upregulating vascular endothelial adhesion proteins and chemoattractants.

TABLE 66.1. TYPES OF IMMUNOPATHOLOGICAL HYPERSENSITIVITY

Type	Cell Types Involved	Mediators	Examples
Immediate (I)	Mast cells with cooperation from basophils, eosinophils, mononuclear cells, and possibly platelets for the late response	Histamine, leukotrienes, platelet activating factor	Allergic rhinitis and asthma, insect venom hypersensitivity, penicillin hypersensitivity
Cytotoxic (II)	Macrophages and neutrophils	Complement and phagocyte-derived mediators (enzymes, oxidants, eicosanoids, etc.)	Goodpasture syndrome, ABO incompatability, thrombocytopenic purpura, myasthenia gravis
Immune complex (III)	Same as type II	Same as type II	Serum sickness, glomerulonephritis
Cell mediated (IV)	Lymphocytes, macrophages, and occasionally basophils	Cytokines	Allergic contact dermatitis, the tuberculin reaction, graft-versus-host disease, chronic allograft rejection, possibly multiple sclerosis

The type II response involves antibodies, usually of the IgG or IgM class, binding with a protein that is present on a-cell membranes or other biological tissues. Examples of this form of immunopathology include Goodpasture disease, immune thrombocytopenic purpura, myasthenia gravis, and ABO and Rh incompatibilities.

The type III response involves the reaction of antibodies, usually of the IgG class, with a soluble antigen to form immune complexes. Immune complexes react with complement, deposit on subendothelial surfaces, and recruited leukocytes activate vasculitis.

In the type IV responses, lymphocytes recognize antigen via the T-cell receptor (CD3). Lymphocytes specific for the antigen release cytokines, which recruit more lymphocytes to the area and further recruit monocytes and macrophages. These cells aid in antigen processing, but they also contribute to the inflammatory response. An example of the type IV response is the intracutaneous reaction to purified protein derivative of *Mycoplasma tuberculosis*. The type IV response also is involved in the host defense against *M. tuberculosis*. Another example is allergic-contact dermatitis, which, more specifically, is cutaneous basophil hypersensitivity because of the intense basophilic infiltrate. This is the mechanism for the inflammatory response seen in poison ivy (Rhus dermatitis).

In this chapter, we will primarily concentrate on type I and type IV hypersensitivities, which are IgE- and lymphocyte-mediated, respectively.

DEFINITION OF ALLERGY

Part of "66 - Allergic Conditions"

Clinicians usually reserve the term allergy to apply to acquired hypersensitivities that have an immune basis. More specifically, allergy usually is reserved for conditions mediated by immunoglobulin E (IgE) and lymphocytes. At present, no single laboratory test or other diagnostic procedure can define a patient as being an allergic individual. Rather, the diagnosis usually is based on the accumulated evidence of the patient's history and *in vivo* and *in vitro* diagnostic procedures. Many investigators define the IgE-mediated allergic state as positivity on skin testing to one or more common aeroallergens, foods, insect venoms, some drugs, and perhaps an elevation of the total serum IgE concentration. This method of defining allergy presumes the allergic individual is capable of producing exaggerated levels of specific IgE to foreign proteins or drugs. It is not uncommon for persons to have positive skin tests to allergens without exhibiting clinical sensitivity on contact with the allergen. Studies with college students, however, have indicated that individuals with positive skin tests are very likely to develop clinical symptomatology within the next 3 years. The term allergy also includes the lymphocyte-mediated allergic contact dermatoses.

REALM OF CLINICAL ALLERGY

Part of "66 - Allergic Conditions"

IgE-mediated hypersensitivity usually results in the clinical conditions of allergic rhinitis, asthma, insect-sting hypersensitivity,

food hypersensitivity of an immediate nature, and occasionally, urticaria and angioedema. The clinical allergist also may be asked to evaluate responses to drugs and foods that do not have a known immunopathological basis, including hypersensitivity to the nonsteroidal antiinflammatory drugs, radiocontrast media, sulfite, and benzoate. There are nonallergic forms of rhinitis in asthma, and the amount of the IgE-mediated component of atopic dermatitis may vary among patients. Likewise, occupational lung diseases may have an immune component or may have an idiopathic etiology. Occasionally, the allergist is asked to evaluate a patient suspected of having a hypersensitivity pneumonitis. This is a group of conditions that often is thought to have an immunopathological basis and precipitated by encounters in the workplace. With the exception of penicillin hypersensitivity, the evaluation of adverse drug reactions is probably the most enigmatic problem facing the clinical allergist. It is rare to find a specific immunopathological mechanism for most adverse drug reactions, and clinical testing is not usually available.

Allergic contact dermatitis results when a sensitizing material is recognized as foreign by the lymphocyte. Unlike the more complex proteins that can induce IgE-mediated allergic processes, small organic compounds and inorganic ions are the usual allergens that induce allergic contact dermatitis. Some time is required for the recruitment of the cells required to mediate type IV hypersensitivity, and reactions usually are not noted for 24 to 72 hours after initial contact with the allergen. Common examples of this form of hypersensitivity include poison ivy dermatitis, dermatitis to nickel contained in jewelry and other apparel items, and the dermatitis resulting from hypersensitivity to the hair-coloring material p-phenylenediamine.

DISEASE STATES

Part of "66 - Allergic Conditions"

Rhinitis

Allergic Rhinitis

Allergic rhinitis is one of the most prevalent disease states in western civilization. Approximately 15% to 20% of the general population is allergic as defined by positive skin tests to inhaled aeroallergen(s) and clinical symptomatology that results from exposure to them. The symptomatology may be seasonal, corresponding to the pollination of plants whose pollen is airborne (grasses, trees, and weeds) and the spores of molds. It also may be perennial, as in the case of house dust mite and animal dander hypersensitivity, where the allergens are in the home or workplace. Some patients have perennial allergic rhinitis from house dust mite hypersensitivity with exacerbations during the pollination seasons. Clinical symptomatology of seasonal allergic rhinitis includes sneezing, watery nasal discharge, postnasal drainage of mucus, and nasal airway obstruction. Itching and swelling of the conjunctival surfaces of the lids of the eyes with increased tearing is a common accompaniment. Allergic rhinitis can be complicated by serous or infectious otitis media or sinusitis resulting from edema of the nasal airway with the occlusion of the Eustachian tubes or sinus ostia, respectively.

Nonallergic Rhinitis

In some subjects who present with a history of nasal congestion and increased nasal mucus production, no evidence of an IgE-mediated disease process can be determined. This condition often has been referred to as vasomotor rhinitis. It is likely that nonallergic rhinitis is actually a group of disorders with different etiologies presenting as nasal congestion with increased mucus production. The patient with vasomotor rhinitis often has the symptomatology of congestion with exacerbation of nasal symptoms on exposure to irritating vapors, bright sunlight, or the inhalation of cold, dry air. Subjects with nonallergic rhinitis with eosinophilia often have symptomatology typical of allergic rhinitis, with marked congestion, sneezing, and rhinorrhea. However, the symptoms often occur outside of the typical pollination seasons. This should alert the physician to look for eosinophils in the nasal mucus. It is presumed that the eosinophils are playing a role in this disorder, but the reason for their appearance in the nasal mucosa and secretions has not been determined.

Patients with disorders of immune function, such as immunoglobulin deficiency states and ciliary dyskinesia, also may present with symptoms of rhinitis. A clue to these conditions is a history of recurrent pyogenic infections of the sinuses, middle ear, and lungs.

Asthma

Asthma may be defined as a condition of reversible obstructive airways disease in which there is a high degree of airway hyperirritability to agents such as methacholine and histamine. While many asthmatics may have their obstruction completely reversed with bronchodilator (beta agonist) therapy, there is growing evidence that poorly controlled chronic asthma may result in a condition of poorly reversible airway obstruction. Asthma is best diagnosed by performing spirometry and determining the patient's reversibility with inhaled bronchodilators. Patients suspected of having asthma but having a normal spirometry can be challenged with methacholine, which can induce airways obstruction in asthmatics at a much lower concentration than normal subjects.

The prevalence of asthma in the general population is increasing in western societies, and death rates from asthma have been increasing in all age groups (2). The reason for these increases has not yet been determined, but it may have its basis in the nature of the life-style of western civilization.

Allergic Asthma

For many years, it was estimated that approximately 20% to 25% of all asthmatics had an allergic component to their disease. More recent data suggest that allergy may play a role in a more significant percentage of asthmatics (3). These estimates, however, are based on skin testing to common aeroallergens and elevations of total serum IgE. It remains to be determined if either of these criteria specifically indicates that a given aeroallergen, or that allergy in general, is contributing to the disease process. It generally is accepted that allergy is a common feature of the childhood asthmatic, and the severity of childhood asthma usually is proportional to the number of positive skin tests and the severity of the reaction to those skin tests (4). Allergy seems to be less frequently diagnosed in older asthmatics, especially if the onset of the asthma occurs late in life. Specific treatment of the allergic condition with allergen immunotherapy has been demonstrated to benefit allergic asthmatics and to diminish the airway's response to allergen (5).

Occupational Asthma

Asthma can result from exposure to specific agents unique to the workplace. Occupational asthma may have an allergic or nonallergic component. Allergic asthma has been noted in seafood processors, bakers sensitive to flour, animal handlers, grain workers sensitive to storage mites, egg processors, and workers exposed to molds, wood dusts, and insect products (6).

Occupational asthma also can occur in response to low-molecular-weight chemicals on an allergic or a nonallergic basis (7). Diisocyanates are used in the production of polyurethane, adhesives, and plastics. It is possible that a small percentage of workers in these industrials have an IgE-mediated disease, but the mechanism in the majority of individuals has not been determined. Isocyanate sensitivity can be disabling and can result in asthmatic symptoms recurring hours after the initial exposure. Airway hyperreactivity can persist for years after removal from isocyanate exposure. Workers exposed to a variety of wood dusts can develop asthmatic symptoms. Plicatic acid is a material derived from western cedar that can cause asthma in some individuals. It is possible that some of the responses to plicatic acid are mediated by IgE. Occupational asthma also has been described in persons exposed to acid anhydrides used in the production of epoxy resins, platinum salts, and to colophony present in solder fluxes. Many other occupational asthmas can result from exposure to organic compounds capable of eliciting IgE responses.

Asthma Exacerbated by Nonsteroidal Antiinflammatory Drugs and Food Additives

Approximately 10% of all chronic asthmatics can have an exacerbation of their asthma within 30 to 60 minutes of the ingestion of a nonsteroidal antiinflammatory drug (NSAID) (8). The mechanism of this disorder is pharmacologic, and there is no evidence of its being immune-mediated.

These drugs inhibit the enzyme cyclooxygenase, the enzyme responsible for prostaglandin synthesis. It is presumably by this mechanism that they induce the exacerbation of asthma. There have been several proposed mechanisms, and they include the suppression of the production of prostaglandin E2, which is a natural inhibitor of the 5-lipoxygenase, or the shunting of arachidonic acid through the 5-lipoxygenase pathway to produce bronchospastic leukotrienes. A heightened responsiveness to leukotriene E4 also has been proposed. Asthmatics sensitive to NSAIDs often have chronic asthma that may require oral or inhaled corticosteroids for control, and they often have nasal polyposis and chronic sinusitis. Subjects identified as being sensitive to NSAIDs must avoid them because life-threatening asthmatic reactions can occur.

Chronic asthmatics also appear potentially sensitive to sulfite, which often is used as a preservative for foods, beverages, and medications (9). The mechanism of sulfite sensitivity has not been determined, but it appears the response is triggered by the inhalation of sulfur dioxide gas that is in equilibrium with sulfite and not by the gastrointestinal absorption of sulfite. Foods and beverages that commonly contain sulfite include beer, wine, dried fruits, sauerkraut, cider, fresh red meat, and processed potatoes. The use of sulfite in restaurant salads has been outlawed, but sulfite may still be applied to grapes. The flavor enhancer monosodium glutamate also has been associated with rare cases of exacerbation of asthma.

Urticaria and Angioedema

Idiopathic Urticaria and Angioedema

The specific cause of recurrent or chronic urticaria and angioedema usually cannot be determined (10). Acute urticaria (urticaria persisting for less than 6 weeks) has been associated with a variety of infectious disorders, including streptococcal pharyngitis, infectious mononucleosis, sinusitis, cholecystitis, and during the prodrome of hepatitis B infection. It is possible that the acute phase response lowers the threshold for mast-cell degranulation. It is certainly possible that urticaria can be associated with an IgE-mediated response, as in the case of some forms

of food and insect sting hypersensitivity. Most cases of angioedema are from mast-cell degranulation occurring in deeper dermal tissues; therefore, individual urticarial wheals are not visible. Many subjects with angioedema give a history of having urticaria concomitantly or on separate occasions.

In addition to IgE-mediated and idiopathic urticaria/angioedema, a number of physical factors can precipitate the disorder. Cold urticaria can occur on exposure to cold air or contact with cold objects. Cholinergic urticaria is triggered by the initiation of sweating. Delayed pressure urticaria is delayed dermal swelling (angioedema) occurring several hours after some form of pressure to a given area. Dermatographism is triggered by scratching of the skin and can be caused by the rubbing of clothing. It presents as linear streaks of urticaria and can be readily diagnosed by stroking the skin and noting the characteristic wheal-and-flare response. Urticaria pigmentosa and systemic mastocytosis can present with generalized urticaria as well as anaphylaxis.

Hereditary Angioedema

Hereditary angioedema (HAE) is a rare condition resulting from the functional absence of the C1 esterase inhibitor molecule (11). The clinical state presents with recurrent orofacial and peripheral angioedema or with acute abdominal crises of pain and diarrhea. Trauma, including surgical procedures, can trigger an attack. The absence of this inhibitor allows for the uninhibited activation of complement by the action of the C1 esterase on C4 and C2. Plasmin, whose formation is triggered by the activation of the fibrinolytic system after thrombosis, is also inhibited by the C1 esterase inhibitor, and plasmin can activate the formation of bradykinin. It is possibly through this mechanism that trauma initiates the angioedematous process.

Acquired C1 Esterase Inhibitor Deficiency

A syndrome similar to HAE can be acquired in patients with disorders associated with large levels of circulating neoantigens (12). Disorders such as autoimmune hemolytic anemia and lymphosarcoma can be associated with the presence of antigens recognized as foreign and to which large amounts of antibodies can be produced. This large burden of immune complexes actually consumes C1 and the C1 esterase inhibitor molecule, resulting in low circulating levels. The angioedema results from the same mechanism as in patients with the hereditary form. Patients with the onset of angioedema after the second decade should have an evaluation for a possible coexisting disorder that is causing the process. Angioedema may precede by many months the other signs of the process causing the decline of the C1 esterase inhibitor levels. If angioedema occurs in the setting of low serum C4 and C1 esterase inhibitor levels in a patient in the third decade or later, a careful physical examination and further laboratory screening tests should be ordered, including a complete blood count, antinuclear antibody determination, hepatocellular enzymes, and testing for lymphoma.

Insect Venom Hypersensitivity

Hymenoptera venoms from bees, wasps, and ants are capable of stimulating the production of specific IgE. Reexposure to these venoms then can initiate an IgE-mediated response that can result in allergic symptoms including generalized urticaria, oropharyngeal angioedema, asthma, and anaphylactic shock. Persons with severe prior systemic adverse reactions are prone to potentially life-threatening reactions after future stings. One exception to this rule appears to be children who have only had dermatologic responses to venom exposure; these children do not appear to be prone to life-threatening events on subsequent stings any more so than the general population (13).

Adverse Reactions to Foods and Food Additives

The most common manifestations of IgE-mediated hypersensitivity to food proteins are nausea, vomiting, diarrhea, and urticaria and angioedema occurring within minutes after ingestion. Anaphylaxis and asthma also may occur. There are delayed reactions to some food substances, and food hypersensitivity has been implicated in the pathogenesis of some cases of atopic dermatitis. In this particular case, ingestion of an offending food allergen can cause an exacerbation of the eczematous dermatitis. Food hypersensitivity has been suggested as the etiologic basis for many other conditions, including hyperactivity, migraine headache, chronic fatigue, depression, and a variety of other somatic complaints not typically associated with mast cell-mediated events. Definitive evidence for the involvement of IgE in these conditions is lacking.

Food protein intolerance has been associated with several other disorders for which definitive evidence of an immune-mediated mechanism is lacking. Food protein-mediated gastroenteropathy occurs in infants and young children and has been associated with intolerance to a variety of foods. The process usually is transient, but it may be associated with vomiting, diarrhea, occult or gross bleeding, and growth retardation. An immunoglobulin E-mediated hypersensitivity to cow milk has been identified in a subset of these individuals. Eosinophilic gastroenteritis can present with vomiting, diarrhea, weight loss, and abdominal pain in either adults or children. Intolerance to milk protein has been documented in some cases even though an immune-mediated hypersensitivity is often lacking. Peripheral eosinophilia and an eosinophilic infiltration in the bowel wall are characteristic features of the disorder.

Gluten-sensitive enteropathy is associated with an intolerance to the wheat protein gliadin. An immune response to gliadin has been demonstrated. The disease usually begins within the first year of life, and symptoms are episodic diarrhea and abdominal pain. Growth retardation with malabsorption of a variety of nutrients may be seen. Immunoglobulin A deficiency exists in these patients at a frequency higher than that of the general population. Dermatitis herpetiformis is a papulovesicular skin disease often associated with gluten-sensitive enteropathy.

Atopic Dermatitis

Atopic dermatitis is a rather common disorder occurring in up to 4% of the population. As its name implies, it is often associated with the presence of IgE-mediated hypersensitivity and is frequently associated with other IgE-mediated disease states such as allergic rhinitis and asthma. This disorder often appears in infancy,

but it can have its onset at any age of life. It is associated with an eczematous dermatitis that can at times be generalized but tends to concentrate in the flexures of the neck and inguinal regions and in the popliteal and antecubital fossae. Xerosis often is found in areas of the skin that are not frankly eczematous, and skin biopsies of these areas show an infiltration of lymphocytes in the dermis. The reason that the skin is the particular shock organ in this atopic diathesis is unknown, but the frequent colonization of the skin with *Staphylococcus aureus* and the frequent improvement of the symptomatology with antistaphylococcal antibiotics suggests that this organism may contribute to the inflammatory response. Patients with head and neck involvement usually show colonization with and antibody responses against the dermatophyte fungus, *Malassezia furfur*. Food hypersensitivity has been associated with exacerbations of the disease; proteins from cow milk, peanut, and egg are common offenders. Double-blind, placebo-controlled food challenges have documented that such foods can exacerbate the dermatitis in patients in whom there is immediate (IgE-mediated) skin test reactivity to the food extracts (14).

Allergic Contact Dermatitis

The term contact dermatitis refers to any dermatologic condition that is triggered by contact with a substance. This does not necessarily imply an immune-mediated etiology. To refer to such a syndrome as allergic is to imply that an immune-mediated response is the cause of the inflammatory state. Allergic contact dermatitis is from a hypersensitivity triggered by a lymphocyte recognizing the offending material as foreign (a type IV response). Because only a small percentage of circulating lymphocytes have specificity for the antigen, a considerable period of time is required for the inflammatory response to become manifest. This is because the immediate reaction of the lymphocyte with the antigen results in the release of cytokines that recruit and stimulate nonspecific lymphocytes and also recruit Langerhans' cells and monocytes to the area to participate in the inflammatory response. It takes approximately 48 hours for a significant response to appear, although reactions may appear as early as 24 hours or as late as 72 hours. The urushiols, which are the contact sensitizers derived from plants of the *Rhus* genus (poison ivy, poison oak, and poison sumac) can trigger this type of response. Nickel sensitivity, as demonstrated by an intolerance to nickel-contaminated metals found in jewelry and clothing apparel, is also very common, and its sensitivity occurs in approximately 10% of the female population in western societies. Other common contact sensitizers include chromates, the hair dye p-phenylenediamine, and chemicals used in the production of rubber. The eczematous dermatitis that results from allergic-contact hypersensitivity may be localized to the site of initial contact, but the antigen may be spread by the fingers to other areas of the body; this is frequently the case in nickel and Rhus dermatitis. When allergic contact dermatitis involves the hands, a particularly severe hand eczema can result and may be recalcitrant to medical therapy.

Drug Hypersensitivity

In the case of drug hypersensitivity, the term hypersensitivity is used to refer to an adverse reaction to a pharmacologic agent that is not a manifestation of the drug's normal pharmacologic action or interaction with other drugs. Drug hypersensitivities may or may not have an immune etiology. Certainly, the most common cause of drug hypersensitivity is the IgE-mediated hypersensitivity to penicillin and its semisynthetic derivatives such as ampicillin and dicloxacillin. Penicillin hypersensitivity also is associated with occasional cross-reactivity to the cephalosporins.

Drug hypersensitivity also may be manifested by erythema multiforme, which is a skin rash with lesions resembling targets and may involve the mucosal surfaces with bulla formation, erythema, and desquamation. Extreme cases may result in the desquamation of the skin as well. There is no firm evidence for an immune-mediated cause for this disorder, and it may result in some cases from an idiopathic toxic response to the drug. Other cutaneous manifestations of drug hypersensitivity include urticaria/angioedema, allergic contact sensitivity, fixed drug eruptions, and a variety of nonspecific, papular or macular eruptions.

Drugs may fix to certain elements in the blood, such as erythrocytes, platelets, and neutrophils, and act as haptens to which antibodies are made. Immunoglobulins responding to the drug on the cellular surfaces can facilitate the removal of these cells, resulting in anemia, thrombocytopenia, and granulocytopenia. A number of drugs have resulted in lupuslike illnesses. They include quinidine, hydralazine, anticonvulsants, isoniazid, and procainamide. An interesting drug-induced syndrome can occur in patients with the acquired immunodeficiency syndrome. Nearly 50% of these individuals will have adverse reactions to sulfonamides. The underlying mechanism for these reactions is unknown.

IgE-mediated hypersensitivity is a common mechanism for adverse reactions to drugs that are proteins, such as insulin, protamine, chymopapain, and heterologous antisera. If immunoglobulin G is the primary antibody produced by the host, serum sickness can occur.

The iodinated radiocontrast media (RCM) can produce acute anaphylactoid reactions that do not appear to have an immune basis. The hyperosmolar nature of many of these preparations may contribute to their ability to initiate reactions. Persons with a history of previous RCM adverse reactions possess a greater susceptibility than the general population to react to these agents on subsequent exposure. The most common manifestations of adverse reactions to RCM are flushing, generalized urticaria, hypotension, and respiratory distress.

Hypersensitivity Pneumonitis

Hypersensitivity pneumonitis refers to a group of lung disorders having an immune etiology that are often caused by the occupational exposure to a sensitizing protein. Primary sensitization to proteins can occur through the pulmonary route, and repeated exposure to these proteins can result in an inflammatory state. There has been considerable controversy over the years as to the mechanism of the inflammatory response. Early work suggested that complement-fixing immunoglobulins were responsible for the syndrome (i.e., type III hypersensitivity), but more recent evidence suggests that antigenic stimulation of lymphocytes may be responsible for the chronic inflammatory response seen in this group of conditions. Nevertheless, the presence of precipitating (IgG) antibodies in the serum of affected individuals may assist

in the diagnosis and will be discussed later in this chapter. Further evidence that delayed (type IV) hypersensitivity is instrumental in the disease process is given by the pathological finding of noncaseating granulomata containing epithelioid and giant cells.

Farmer's lung is a disorder triggered by the inhalation of spores of *Micropolyspora faeni* or thermophilic actinomycetes, which often contaminates hay used for the feeding of livestock. Pigeon breeder's disease, or bird fancier's disease, is caused by the inhalation of proteins contained in bird droppings when hobbyists or laboratory personnel enter the animal's living facilities. Similar respiratory disorders are seen in maple bark stripper's disease, mushroom worker's lung, and many other pneumonitides resulting from the inhalation of protein antigens (15) (Table 66.2).

TABLE 66.2. EXAMPLES OF HYPERSENSITIVITY PNEUMONITIS

Disorder	Antigen Source
Farmer's lung	Thermophilic *actinomycetes* in grains and hay
Bird fancier's disease	Proteins in bird droppings
Bagassosis	Thermophilic *actinomycetes* in moldy sugar cane
Malt worker's lung	*Aspergillus* species in moldy barley
Maple bark disease	Mold in maple bark
Woodworker's lung	Wood dusts or molds in wood

Clinical symptomatology often presents as fever, malaise, myalgias, and respiratory symptoms such as cough or dyspnea, occurring several hours after exposure to the offending protein. Repeated exposures over many years may result in a chronic pulmonary condition finally ending in fibrosis and irreversible pulmonary disease. Suspecting the disease process and making the diagnosis is essential so that counseling can be instituted to reduce exposure to the offending proteins and thus to prevent the end-stage lung disease that may result.

Allergic Bronchopulmonary Mycosis

Allergic bronchopulmonary mycosis (ABPM) is a complication that can occur in allergic asthmatics. It is believed that patients with IgE-mediated hypersensitivity to *Aspergillus* and other species may inhale spores of the organism, resulting in an acute asthmatic response associated with the increased production of tracheobronchial mucus characteristic of allergic asthma. The organism then can reside in mucus creating chronic inflammation resulting in a chronic and progressive asthmatic state. Because of this heavy allergen exposure in the lung, the patient develops precipitating IgG antibodies to offending fungal antigens as well as very high titers of specific and nonspecific IgE. Like hypersensitivity pneumonitis, ABPM was originally thought to be caused by complement-fixing, (IgG) antibodies initiating the inflammatory response, but more recent evidence suggests that this condition, too, may be associated with a significant type IV hypersensitivity component. It also is possible that a chronic late-phase response (see beginning of chapter) exists in these individuals. If unrecognized, the condition may proceed to bronchiectasis and finally fibrosis. ABPM may also be a complication of cystic fibrosis.

METHODOLOGIES

Part of "66 - Allergic Conditions"

IgE-mediated Disease

Skin Testing

In patients suspected of having IgE-mediated disorders, skin testing remains the most accurate and cost-effective method of diagnosis (Table 66.3). Skin testing relies on the specific IgE-mediated release of mediators from tissue mast cells that cause a wheal-and-flare response. Skin testing should be performed by personnel trained in its application with supervision by physicians experienced in the diagnosis of IgE-mediated diseases. Extracts must be purchased from reliable suppliers, and there is a growing emphasis on the need for allergen extract standardization.

TABLE 66.3. METHODS FOR EVALUATING HYPERSENSITIVITY

Disorder	Diagnostic Test(s)
IgE-mediated diseases:	Skin testing
Allergic rhinitis	RAST and other *in vitro* methods for determining allergen-specific IgE
Allergic asthma	
Atopic dermatitis	
Insect venom allergy	
Food allergy	
Penicillin and protein drug allergy	
Hypersensitivity pneumonitis	Precipitating antibody to antigen by double immunodiffusion
	Chest radiograph
Hypersensitivity to NSAID and sulfite	Oral challenge with monitoring spirometry
Allergic bronchopulmonary aspergillosis	Skin testing to *Aspergillus*
	Total serum IgE (RIST)
	Precipitating antibody to *Aspergillus* by double immunodiffusion
	Chest radiograph
Allergic contact dermatitis	Patch testing

In persons suspected of having IgE-mediated hypersensitivity to inhalants or foods, the percutaneous or epicutaneous skin tests should be applied first. A 1:10 or 1:20 weight/volume (w/v) aqueous extract containing 50% glycerol is applied as a drop to the volar aspect of the forearm or to the back. The allergen then is introduced into the epidermis with a sharp device by passing the point of the device through the drop and into the epidermis. Reproducible application has been demonstrated with the Wyeth bifurcated needle (resembling the old vaccinia scarification device), the Pharmacia Lancet, and the Staller point (a variation of the Morrow Brown needle) (16, 17). In performing percutaneous tests, hypodermic needles, the standard Morrow Brown needle, and the Staller kit give slightly less reproducible skin tests than the devices previously mentioned. After the allergenic extract has been introduced into the epidermis, the extract is wiped from the skin, and the wheal-and-flare is measured 15 to 20 minutes later.

When epicutaneous tests are negative, intracutaneous testing may be required as a more sensitive test to identify aeroallergen hypersensitivity. The intracutaneous test is less specific than the epicutaneous test but can still be interpreted with confidence providing the proper concentrations of materials are used. A 1:1000 w/v concentration of allergen extract is injected at a very shallow angle so that material is introduced into the dermis just beneath the epidermis. Approximately 0.02 mL of material is injected. This volume is critical because wheal size is directly proportional to the volume of the material injected. The intracutaneous test result also is interpreted 15 to 20 minutes after its application.

Total Serum IgE, Specific IgE, and Other in vitro Methods (18)

Total serum IgE can be determined by a variety of methods. The most commonly used methods include solid-phase immunoassay with fluorescence detection, radiolabeled immunoprecipitation, immunoplate enzyme-linked immunosorbent assay (ELISA), microparticle-enhanced immunoassay, and solid-phase immunoassay with luminescence detection. IgE is measured in international units that are equal to 2.2 ng. The method chosen should have a minimum operating range from 5 to 1,000 IU/mL. The method should be tested to assure that it does not exhibit a large “hook effect”, giving falsely low values for specimens with elevated levels. Occasionally, specimens will be encountered with total IgE levels at least 100-fold out of range. Typical coefficients of variation are 15% for better commercial assays, except for specimens that are substantially out of operating range. A typical standard curve has 7 points, because the assays do not exhibit linear performance. Semiquantitative and screening IgE assays are of very limited value.

The radioallergosorbent test (RAST) was developed in the late 1960s and early 1970s following the discovery of IgE as the antibody causing type I hypersensitivity reactions. RAST is the classic method for determining specific IgE antibody. While originally developed as a research laboratory method, it was rapidly translated into a commercial diagnostic procedure. Allergen is usually covalently linked to a cellulose disk, but other insoluble support systems such as agarose may be used. The patient's serum is added to the allergen. The solid phase is washed and then incubated with a radiolabeled antibody specific for IgE. The solid phase is rinsed again, and the amount of bound radioactivity is quantitated to determine the amount of specific IgE directed toward the allergen in the original serum sample. This method generally has been replaced by the Immuno-CAP procedure, which uses an expanded cellulose matrix with high ligand binding capacity and enzyme based fluorescence detection. This method is significantly more sensitive and slightly less specific than most other methods in current use.

Other methods have been developed to quantitate specific IgE *in vitro*. The antibody binding to the allergen-affixed IgE can be labeled with an enzyme that uses a substrate that is converted to a colored product (ELISA) or produces a fluorescent product (the fluorescent allergosorbent test or FAST). One method utilizes luminescence generated by radiographs released in the radioactive decay of 125I-labeled anti-IgE (the multi-thread allergosorbent test or MAST). Others use chemiluminescence, colorimetry, or particle counting methods. In various methodologies, allergen is bound to cyanogen bromide activated paper disks, microtiter plates, activated fibers, microparticles or is labeled with biotin. Each of these methodologies is available through a commercial vendor. Basically, these methods are all similar in principle to the RAST and share similar advantages and disadvantages of RAST. There are significant differences in the performance of each commercial method, and for some allergens performance of many methods is poor. Manufacturers should provide information on the sources and characterization of allergens used in their systems. Fungal allergens and many fresh fruits and vegetables are particular problems in most methods. The performance characteristics of assays used for diagnosis of life-threatening conditions, such as venom anaphylaxis, and occupational conditions, such as latex allergy, need to be carefully evaluated before relying on negative results.

In vitro methods with multiple allergens coupled to the support are available and can be used for screening for atopy. These methods are slightly less sensitive than using individual allergen tests, but can be used in a cost effective manner.

Although most specific IgE tests can be reported in quantitative units, these methods are not truly quantitative because of the unknown composition of most allergen extracts. Results are not comparable between different allergens and among different methods.

Clinical States Where Skin Testing, Specific IgE Antibody and Total IgE Are Useful (19)

Perhaps the greatest utility for skin testing and specific IgE antibody lies in their ability to aid in the diagnosis of aeroallergen hypersensitivity and the resulting conditions of allergic rhinitis and asthma. Skin testing and, to a more limited extent, specific IgE antibody can be used to confirm IgE-mediated hypersensitivity to hymenoptera, including imported fire ant venoms. True IgE-mediated reactions to foods can be confirmed with skin testing to food extracts, but many adverse reactions to foods have no known immunopathological basis. Furthermore, skin tests to fruits and vegetables are unreliable and false-negative reactions are common. Epicutaneous testing by pricking a fresh fruit or vegetable and then pricking the skin may be necessary when testing with a commercial extract is negative and true IgE-mediated hypersensitivity is suspected. The avoidance of foods to which there is skin test positivity can also benefit patients with atopic dermatitis.

Skin testing for drug hypersensitivity is much more limited in usefulness. Skin testing to drugs that are proteins, such as insulin or protamine, can be informative, but testing to nonprotein drugs is greatly limited. There may be several reasons for this. First, mast-cell degranulation and the positive skin test depend on the reaction of at least two IgE molecules with an antigen. Most drugs are haptens, and without a “carrier” molecule with multiple drug-binding sites, IgE bridging cannot occur. Second, drug metabolites may be the sensitizing antigens. Finally, it may be that many adverse reactions to drugs do not have an IgE-dependent basis. Penicillin and anesthetics are exceptions. Penicillin sensitivity can be determined with a polyhaptenic synthetic

penicilloyl polylysine (PrePen). Skin test positivity to this so-called major determinant will predict an urticarial reaction to penicillin administration. Unfortunately, this skin test reagent will not predict all hypersensitivity (especially anaphylactic) reactions to penicillin, and one also should test with a fresh penicillin preparation as well as with minor determinants for which a commercial test preparation is in development. Specific IgE testing is useful in evaluation of latex sensitivity and standardized skin test reagent is in development. Total serum IgE usually is not useful in determining whether or not a condition has an IgE-mediated basis. The test is helpful in making the diagnosis of allergic bronchopulmonary aspergillosis and can be used to monitor disease activity. Total IgE also can be used to aid in the diagnosis of the hyper-IgE syndrome, a condition manifested by eczema, recurrent cutaneous and visceral staphylococcal abscesses, and mucocutaneous candidiasis.

In Vitro Activation of Basophils and Measurements of Mediator Release (18)

Circulating basophils possess a high-affinity receptor for IgE and contain histamine that is released when an allergen crosslinks IgE on the surface of the cell. The test is performed either on whole blood or on various preparations of partially purified basophils. Cells from some donors are poor releasers and are not suitable for testing. The testing may be carried out using either cells from the subject or by passive sensitization of donor cells. Methods are available to detect histamine release by either coupling to o-phthalaldehyde with fluorescence detection, labeling by enzymatic addition of radioactive methyl group, or immunoassay. Autoanalyzer-based methods and microplate assays are commercially available. Leukotriene production can be measured by immunoassay. Alternatively, degranulation can be monitored by flow cytometry. These methods are useful in pharmacologic studies of IgE mediated reactions as well as for diagnosis of IgE mediated reactions.

Tryptase, an enzyme mediator preformed in mast cells, is released upon mast-cell activation. It can be measured by immunoassay and is useful for the evaluation of anaphylactic and anaphylactoid reactions and systemic mastocytosis. Tryptase has a plasma half-life of about 90 minutes.

Recently there has been interest in the measurement of eosinophil-derived mediators in tissues and body fluids. Research studies have examined the distribution of eosinophil major basic protein and eosinophil-derived neurotoxin by immunostaining procedures. There is a commercial method available for eosinophil cationic protein measurement, which can be measured in serum, urine, nasal secretions, bronchial secretions, and specimens derived from the gastrointestinal tract. This may provide a useful adjunct method for evaluation of inflammatory aspects of the IgE mediated response. There also are studies of nitric oxide production in sputum and nasal secretions to evaluate inflammatory responses.

Experimentally, a number of other mediators have been studied including prostaglandins, thromboxanes, platelet activating factor, adenosine, chymase, eosinophil major basic protein, reactive oxygen species, kinins, and anaphylatoxins. Presently, none of these tests is utilized clinically.

Cytology

Cytologic examination of nasal secretions and lower respiratory mucus can be helpful in the diagnosis of certain pathological conditions. Nasal mucus may be obtained by scraping the inferior turbinate or by collecting nasal mucus by having the patient blow onto a waxed paper. The latter method will reveal inflammatory cells that have migrated into the mucus and will have fewer contaminating epithelial cells that are obtained when the nasal mucosa is scraped. The mucus is spread thinly on a glass slide, allowed to dry in air, fixed with methanol, and then stained with a solution that is sensitive in detecting eosinophils. Hansel's stain is popular for this purpose, but stains used for making peripheral blood smears may be adequate. The absence of eosinophils does not rule out an allergic condition because neutrophils also migrate into the nasal passages in allergic rhinitis. Furthermore, a secondary sinus infection may result in a greater percentage of neutrophils in the mucus of the allergic subject. Nasal cytology may be most helpful in confirming the diagnosis of nonallergic rhinitis with eosinophilia (NARES) and in differentiating acute viral rhinitis from allergic rhinitis. When a large percentage of the nasal inflammatory cells are eosinophils (usually greater than 50%) and all of the skin tests to common aeroallergens are negative, one should feel comfortable in making the presumptive diagnosis of NARES. It may be difficult clinically to determine the cause of acute coryza in a patient presenting during a pollination season. In such cases, nasal cytology often reveals a high percentage of eosinophils in allergic subjects, but primarily epithelial cells are seen in patients with acute viral coryza.

Lower respiratory mucus can be more difficult than nasal mucus to examine for inflammatory cells. These cells do not stain as well, and it is difficult to differentiate them. Lower respiratory mucus, however, can be examined as a wet mount preparation pressed snuggly between a cover slip and a glass slide. With a little practice, one can recognize lymphocytes, neutrophils, eosinophils, ciliated epithelial cells, erythrocytes, and alveolar macrophages at 400X magnification without the use of staining materials (Fig. 66.1). If eosinophils have been in the airway at high numbers for a fairly long period of time, Charcot-Leyden crystals (elongated, diamond-shaped bodies 15 to 50 μm in length) may be seen.

Cell-mediated Immunity (CMI)

Patch Testing

Patch testing is the preferred method for diagnosing allergic contact dermatitis that is a result of a type IV hypersensitivity response. Performance of this testing requires an experienced investigator who is knowledgeable in the limitations of patch testing, experienced in the application of the test, and capable of interpreting skin reactions and counseling the patient regarding the avoidance of materials that may contain the offending allergen.

A standard set [TRUE test kit (Glaxo Wellcome, Research Triangle Park, NC)] containing 24 common contact allergens is commercially available (20). Most of the allergens used in testing are mixed in white petrolatum, but occasionally an aqueous vehicle is preferred. Only physicians well experienced in patch testing

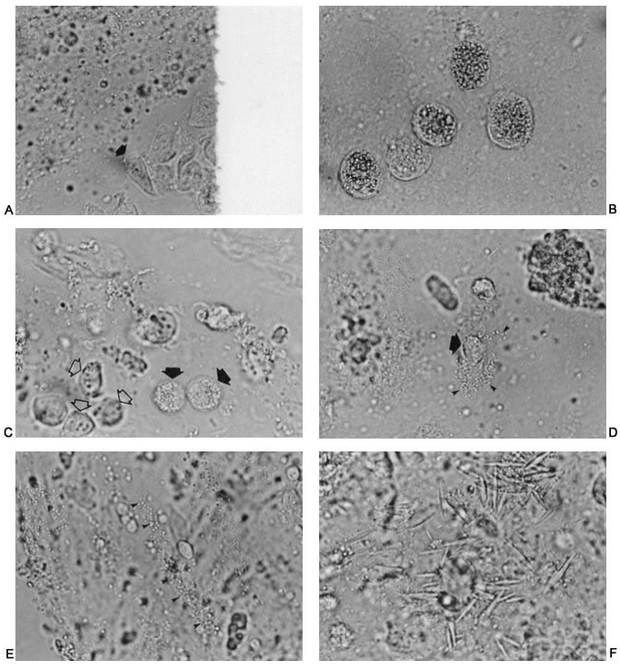


FIGURE 66.1. Photomicrographs of a sputum sample from an asthmatic patient. All of the micrographs were taken of unstained sputum samples pressed between a cover slip and a clear microscope slide. A sample was observed using the 40X objective and the 10X eyepiece. The final magnification of the micrographs is 1860X. (**A**) The broad arrow points to the brush border of a ciliated epithelial cell. Note the cell's apical nucleus and its pointed tail, which serves to fix the cell to the basement membrane. These cells often are sloughed in an inflammatory condition such as asthma or bronchitis. (**B**) A cluster of alveolar macrophages. Note the large intracytoplasmic inclusion bodies and granules. (**C**) The empty arrows point to neutrophils and the solid arrows point to eosinophils. The eosinophils are best distinguished by their fine granules, which easily stand out, especially if one gently focuses with the fine adjust knob on the microscope. The neutrophils are approximately the same size as the eosinophils, but their granules are much less distinct. (**D**) The large arrow points to an eosinophil that is degranulating, and its granules have spilled extracellularly (fine arrows). (**E**) In sputum samples that have been allowed to stand, only eosinophil granules may be seen and intact eosinophils may be difficult to locate. The fine arrows point to eosinophil granules that are somewhat linearly distributed in the sputum sample. (**F**) After sitting at room temperature for several hours, the granules themselves disrupt, allowing proteins to leak. Charcot-Leyden crystals form when lysophospholipase crystallizes out of solution. The presence of these crystals indicates that eosinophils have been present in the past, and it is often difficult to find intact eosinophils where Charcot-Leyden crystals are seen.

should prepare a test material that previously has not been investigated. Unstandardized materials raise the risk of irritant reactions, and marked allergic and toxic reactions can occur if the concentration of the allergen is too high. In addition to avoiding the use of unstandardized testing materials, placing patches on areas with active dermatitis must be avoided.

Testing material is applied from the squeeze bottle to filter paper disks spaced on an aluminized paper strip. The strip is then fixed to a hypoallergenic strip of tape and applied to the patient's back. Care must be taken to be sure that the tape is securely in place; to prevent the tape from loosening when the skin on the back is stretched, the patient should have the back bent forward while the strips are affixed. Once in place, the patient should be cautioned to keep the areas dry while bathing and to avoid situations where sweating might occur. The test should be left in place for 72 hours with the patient instructed to remove those tests where intense itching or irritation occur prior to the time when the tests are to be interpreted.

In Vitro Cytokine Production

Incubation of lymphocytes with antigen results in cytokine production with cellular proliferation. The incorporation of radiolabeled thymidine into dividing cells can be used as an index of hypersensitivity to an antigen. This method remains a laboratory tool, and reliable *in vitro* testing for cell-mediated immunity is not widely available.

Lymphocytes and other leukocytes can be incubated with mitogens, antigens, or peptides for stimulation and response studies. Various lymphokines, cytokines, and interleukins (IL) can be measured in supernatants by enzyme-linked immunoassay using commercially available kits. Interferon-gamma and migration inhibition factor are correlates of cell-mediated immunity. IL-4 production is a correlate of IgE production. Production of IL-2, interferon-gamma, and tumor necrosis factor is typical of Th1 responses and production of IL-4 and IL-13 are typical of Th2 responses. IL-5 induces maturation of and is chemotactic for eosinophils.

In Vivo Challenges

Food Hypersensitivity

Besides the history, the first step in diagnosing IgE-mediated food hypersensitivity is percutaneous skin testing. Foods identified as possible allergens can be confirmed by oral challenges. If a patient has had a life-threatening adverse reaction to a food to which he or she is sensitive, oral challenges should not be performed. Indeed, oral food challenges are rarely necessary in clinical practice but have been used as a research tool to prove that food hypersensitivity can exacerbate some disorders. Double-blind food challenges have been helpful in demonstrating that food hypersensitivity can contribute to atopic dermatitis; egg, peanut, and milk account for the majority of the adverse reactions. A major drawback to this testing is the difficulty in the preparation of gelatin capsules containing freeze-dried foods or food extracts.

NSAIDs

Up to 15% of all chronic asthmatics may have exacerbations of their asthma on ingesting an NSAID. This figure may rise to 70% of patients who have chronic asthma and paranasal sinusitis or nasal polyposis. Many of these individuals will not have a history of adverse reactions because pulmonary function values may not have dropped sufficiently to result in symptomatology. If, however, an asthmatic suspected of having NSAID hypersensitivity requires an NSAID for another condition, an oral challenge may be required under close supervision. Several precautions need to be taken because adverse reactions to these compounds can be life threatening. First, the FEV-1 (forced expiratory volume at 1 second) should be 70% of predicted or better prior to testing. It may be necessary to give a short course of corticosteroids to bring the lung functions up to an acceptable level for testing. One proposed method suggests administering a placebo at 3-hour intervals beginning at 8 a.m. for a total of three doses (21). The FEV-1 is recorded hourly during this time period and for 3 hours following the last dosage. On the second day of the challenge procedure, aspirin is given at 8 a.m. Either a 3-mg or 30-mg dose is used initially, depending on whether or not there is a prior history of a serious adverse reaction to aspirin. Again, the FEV-1 is measured hourly or if symptoms develop. A second dose (30 mg if 3 mg was the first test dose) of aspirin may be given at 11 a.m., and a final dose of 100 mg is given at 2 p.m. Again, the FEV-1 is measured hourly after each dose. If no adverse reaction occurs, the patient returns to the clinic on the third day and is then given 150 mg at 8 a.m., 325 mg at 11 a.m., and 650 mg at 2 p.m., provided that the FEV-1 has not dropped by greater than 20% at any dosage. Such a drop is an indication of a positive test and hypersensitivity to NSAID.

Sulfite

The suspicion of adverse reactions to sulfite (or metabisulfite) should occur after a person has had an adverse reaction following the ingestion of a food or beverage. The adverse reaction may be because of sulfite contained in the food or beverage, to an IgE-mediated hypersensitivity reaction to a food protein, or to some other food additive, such as monosodium glutamate (MSG). It is reasonable to do a sulfite challenge in such individuals to help determine the presence or absence of sulfite sensitivity. In aqueous solutions, sulfite is in equilibrium with sulfur dioxide. Most likely, the sulfur dioxide inhaled during the ingestion of a food or beverage that contains sulfite is the cause of the adverse reaction in the lung. Therefore, sulfite is placed in acidic solutions, such as a commercial lemonade preparation. Sulfite is prepared at 100 mg/mL in water on the day of the challenge. Spirometry is performed, and the FEV-1 must be greater than 1.5 L or 70% of predicted prior to performing the challenge. Theophylline and inhaled corticosteroids are continued. Inhaled bronchodilators are discontinued on the day of the challenge, and cromolyn is discontinued for 24 hours prior to the challenge. The patient is challenged with progressively increasing doses of sulfite (1, 5, 10, 15, 25, 50, 75, 100, 150, and 200 mg) by adding the sulfite to 20 mL of the lemonade mixture and asking the patient to swish the solution in the mouth for 15 seconds before swallowing (22). This allows the sulfur dioxide that is in equilibrium

with sulfite to effervesce and be inhaled into the lung. Increasing doses are given at 10-minute intervals with spirometry being performed just prior to the next dose. A positive test occurs when the FEV-1 drops 20% or more below the baseline value. Placebo challenges may be given periodically among the sulfite challenges to be sure that there is not a spontaneous drop of the FEV-1 occurring as the challenge progresses. Positive challenges should be repeated in a double-blind fashion.

Occupational Asthma

Asthma occurring in the workplace can be from a type I or other immune response or may have an idiopathic etiology. Several approaches can be taken to document occupational asthma (23). The easiest way to reproduce the environmental workplace is to have the patient perform spirometry there. However, this may be time-consuming and cumbersome, and the performance of a peak expiratory flow rate (PEFR) is a practical alternative. The subject should record the PEFR prior to the work shift and at regular intervals throughout the day. Careful notation should be made of what exposures occurred prior to the drop in the PEFR. The PEFR should be followed after the work shift to determine if a late response occurs. Despite positive data on a drop of the PEFR at the workplace, bronchoprovocation with specific materials may be required if compensation is being sought or if there may be several potential offending materials. Bronchoprovocation should be performed by clinicians familiar with the evaluation of occupational lung disease and who are aware of the test's limitations. Detailed guidelines have been published (24). Spirometry performed before and after exposure is necessary to make the diagnosis. The same precautions outlined in the sections above on sulfite and NSAID hypersensitivity should be observed while performing bronchial challenges. In some cases, a specific compound, such as toluene diisocyanate (TDI), may be suspected, and challenge may be performed with this material. For volatile compounds such as TDI, special challenge chambers are required that usually are not available in clinical settings. Challenges to proteins, however, can be performed by mixing the material in an aqueous solution and inhaling the material with a dosimeter delivery system. The protein concentration is increased serially 10-fold until the FEV-1 drops by 20% or more. If concentrations for challenge have not been standardized, it may be necessary to first perform a percutaneous skin test with the material to find a concentration that does not elicit a large response. In most situations the most effective evaluation can be made by monitoring the worker at intervals as he performs his normal duties. This method was particularly useful in dealing with reactivity to colophony in soldering. In all cases where bronchial challenges are performed, the physician must be aware that late responses may occur 4 to 8 hours following the initial drop of the FEV-1. The patient may be at home when this occurs, and instructions must be given on the proper procedures necessary to treat the response. If severe drops in the FEV-1 occur with a challenge, it may be necessary to administer corticosteroids (such as prednisone, 1 mg/kg) at the time of the initial challenge to help prevent the late phase response from occurring. If milder reactions occur, it may be informative to monitor the patient for late phase responses, either as an inpatient or with PEFR monitored at home.

Hypersensitivity Pneumonitis

Hypersensitivity pneumonitis has an immune basis, but a specific hypersensitivity type cannot be ascribed to any of these conditions. A type III response has been implied because of the frequency in which IgG precipitating antibodies can be detected against a specific antigen. The clinical features of the disease, however, more closely resemble a type IV immune response with granuloma formation, and an eventual progression to pulmonary fibrosis can be seen.

Patients suspected of having hypersensitivity pneumonitis should have a radiograph of the chest. During the acute phase of the illness, a rather diffuse, small nodular appearance is noted. Occasionally, infiltrates can be seen. Eventually, the patient enters a chronic fibrotic phase, and the typical radiographic features of interstitial fibrosis are prominent. Generally, pulmonary function testing reveals a restrictive defect with impaired gas exchange. Thus, spirometry should reveal a proportionate decrease in both the FEV-1 and the forced vital capacity (FVC) with an FEV-1/FVC ratio of 0.75 or greater. In severe cases, the single breath diffusion capacity for carbon monoxide (DLco) will be impaired, and there may be a fall of the PaO2 with exercise.

Heavy exposure to organic dust may result in the production of IgG antibodies to antigenic proteins in the dust. Commonly available antigens for screening for hypersensitivity pneumonitis include thermophilic actinomycetes (*Micropolyspora* and *Thermoactinomyces* species), and avian proteins. The presence of precipitating antibody in the patient's serum is determined by double immunodiffusion in agarose gels. Many other antigens have been implicated in causing hypersensitivity pneumonitis, and these antigens may not be readily available for testing for precipitating antibodies (15). It may be necessary to send the patient's serum to a specific research laboratory where a standardized antigen and the appropriate negative and positive control sera are available. Persons working in an environment with a heavy exposure to an organic dust may develop precipitating antibodies without the presence of clinical disease.

A related disease can be seen with fungal colonization of the sinuses. Precipitating antibodies and IgE antibodies are usually found in allergic fungal sinusitis.

Allergic Bronchopulmonary Mycosis

Allergic bronchopulmonary mycosis (ABPM) occurs when fungal species reside in the airways of allergic asthmatics. Their proteins are very antigenic, and high titers of IgE and IgG antibodies are produced. Additionally, cell-mediated hypersensitivity is likely to occur. Persons suspected of having ABPM should have total serum IgE and precipitins (IgG antibody) to fungi determined, a skin test to fungal extract, a chest radiograph, and a total eosinophil count (25). Some research laboratories are capable of measuring specific IgE and IgG antibodies to *Aspergillus fumigatus*. At present, these tests generally are not available. The majority of patients with ABPM will react on skin testing to the offending fungal allergen. In most cases, *Aspergillus* species are responsible, but sensitivity to other fungal species should be determined if there is a high enough suspicion that aspergillosis exists and testing to *Aspergillus* species is negative. Similar syndromes

have been seen with other fungi, including *Candida*, *Curvularia*, *Stemphylium*, and *Helminthosporium*. Chest radiographs also are important in following the disease process. Thick-walled bronchi, infiltrates, and mucoid impaction may be seen. If the disease is not recognized early, progression to central bronchiectasis and pulmonary fibrosis with bulla formation may occur. Response to treatment can be followed with periodic chest radiographs and total serum IgE levels.

Hereditary Angioedema and Acquired C1 Esterase Inhibitor Deficiency

Complement Measurements

Approximately 15% of patients with HAE will have the antigenic presence of the C1 esterase inhibitor (C1EInh); assays utilizing an antibody to C1 esterase will indicate normal levels and will not reflect the functional impairment of the molecule. Therefore, if the patient is suspected of having HAE and has a normal immunologic assay for the C1 esterase inhibitor, a functional C1 esterase inhibitor assay can be obtained. Because of this inhibitor deficiency, there is excessive activation of C4 and C2. For this reason, serum C4 or C2 levels are the most convenient screening tests for diagnosing this condition. The C4 level is in the low normal or below normal range between attacks and always drops to a very low range during acute attacks. C2 levels exhibit similar behavior. If the C4 level is low normal or below, the diagnosis should be confirmed by measuring functional C1 esterase levels. In addition to low C4 and C1EInh levels in serum, patients with acquired C1EInh deficiency have low serum levels of C1.

Controversial Methodologies

To help diagnose allergic disorders, a number of controversial methods have arisen that have not utilized the scientific method to determine their validity. One such procedure is cytotoxic testing.

This test depends on the addition of food, chemical, and inhalant extracts to peripheral blood leukocytes, which are then observed for microscopic changes. This testing procedure has been demonstrated to be invalid, and federal agencies and many insurance companies have refused to pay for this testing. As a consequence, its use has declined in recent years, but other tests, without scientific merit, claiming diagnostic efficacy in food hypersensitivity continue to appear.

In an effort to find a lower, safer dose of allergen to be used for immunotherapy (desensitization therapy), a test (Rinkel method) that depends on an endpoint titration was developed. This concentration directed the clinician to a treatment dose that typically was considerably lower than the levels used in conventional immunotherapy. As a consequence, many patients on this form of therapy never reached a dosage that could benefit their symptoms. Very weakly sensitive individuals, however, could approach an effective concentration because of the higher concentration of allergen extract subsequently used for immunotherapy. Controlled studies have demonstrated the method to be ineffective in controlling symptoms of ragweed hay fever, while conventional immunotherapy resulted in improvement significantly greater than that of placebo-treated individuals.

Subcutaneous and sublingual provocation testing are used currently primarily by clinical ecologists in their treatment of allergic disorders and any of a variety of conditions that they consider to have an allergic basis, such as migraine headache or multiple somatic complaints. The therapy is based on the principle of finding a dosage of an allergen or chemical that provokes objective or subjective signs in the patient. A neutralization dose that is either slightly weaker or stronger than the provoking dose is then chosen and is injected subcutaneously in an effort to “neutralize” the symptoms. Others have proposed that sublingual solutions of chemicals or allergen extracts also can neutralize symptomatology. At present, no controlled studies have demonstrated that subcutaneous or sublingual neutralization techniques have any validity.

INTERPRETATION

Part of "66 - Allergic Conditions"

IgE-mediated Disease

Skin Test Size and Scoring

Generally, an epicutaneous skin test reaction is considered significant if it results in erythema (flare) greater than 20 mm in diameter and definitely is considered positive if there is an accompanying wheal. The advantages of the epicutaneous (prick) test are its ease and speed of application, infrequency of false-positive reactions, sensitivity, cost, and the rapidity with which results can be obtained. Intradermal skin test reactions generally are considered positive if the erythema has a diameter of greater than 21 mm with an accompanying wheal of 5 to 10 mm. If the erythema is greater than 30 mm with an accompanying 10 mm wheal, the reaction is definitely considered to be positive.

Epicutaneous skin testing is relatively sensitive and specific. Intradermal skin tests are more sensitive and less specific. Quantitative intradermal skin testing with serial dilutions is the standard method for hymenoptera venoms and is recommended for some drugs. Positive skin tests are not uncommon in individuals who do not have active allergic disease. Testing for hymenoptera venoms, drugs and foods should not be performed on individuals who have not experienced allergic reactions. It is important to verify that drugs and unusual extracts do not have nonspecific reactivity.

Despite the wide degree of popularity, there are some problems with skin testing. A mild degree of discomfort is experienced, and there certainly is intersubject variation in the skin reactivity that results in the wheal-and-flare response. Persons with dermatographism can develop erythema and wheals simply from the irritation of the skin with the test device. Patients with extensively inflamed skin, as in atopic dermatitis, may not have enough clear surface area for the test to be applied. Furthermore, patients with atopic dermatitis often have a minimal flare response to the skin test, making the results more difficult to interpret. In these cases, *in vitro* testing may be required. The major role for skin testing is in the determination of IgE-mediated hypersensitivity to inhaled aeroallergens, food allergens, insect venoms, pharmaceuticals that are proteins and penicillin. Most

other pharmaceuticals have not been standardized as skin test reagents, thus dramatically reducing the effectiveness of the skin test in diagnosing the mechanism of possible drug hypersensitivity reactions. Furthermore, skin testing is ineffective in diagnosing hypersensitivity to radiocontrast media or to nonsteroidal antiinflammatory drugs, and it is not capable of detecting persons sensitive to sulfite and other pharmacologic sensitivities.

Total Serum IgE

Total serum IgE should be reported as international units (IU) per milliliter of serum. Results should be interpreted using tables of values by age and country or region, because levels vary widely because of infections and infestations with parasites. Serum from newborns may contain very low levels of IgE. Therefore, two sets of standards may be required, one for the adult range and one for the range of IgE values seen in the pediatric population (26).

Specific IgE Antibody Scoring

Specific IgE antibody usually is scored by arbitrarily breaking down the total radioactivity counts into classes. The interpretation of the test results should be restricted to the class designation and not to the actual total radioactivity counts. The class determination is made by comparison of the allergen-specific IgE content in test samples to positive and negative controls. The most common scoring systems assign ranges from zero to four or zero to six. Zero indicates a clearly negative result and one usually an equivocal result. Scores of two and higher are increasingly positive. There is little correlation between severity of allergic manifestations and degree of positivity. There are a number of factors that can lead to false results including high levels of total IgE, the presence of low affinity IgE antibodies, low amounts of some allergenic proteins coupled to the solid-phase, and the presence of low, but clinically significant amounts of specific IgE antibody. The presence of specific antibodies of IgG, IgA, and IgM classes may decrease the levels of specific IgE antibody by competing for allergen sites, but usually does not reduce the level to negative. This effect can be overcome by the use of matrices that have a very high protein binding capacity, and the use of enriched extract preparations with interfering materials removed by dialysis or gel filtration.

Some manufacturers recommend the use of quantitative scoring systems, which are calibrated with either a total IgE binding curve or by IgE binding to a specific allergen, commonly birch pollen. These systems are somewhat misleading because they are not truly quantitative, and they are dependent on binding equilibrium. Comparison of specific IgE binding to calibration curves assumes that the magnitude and distribution of antibody binding affinities are similar in both systems. In addition almost all patients demonstrate an IgE response to multiple allergenic components and epitopes of allergenic components. One cannot assume that each of these components is present on the solid phase in sufficient excess to not be a limiting factor in binding capacity. An additional complication is the binding of IgE to epitopes that may be clinically insignificant, such as carbohydrate side chains and low affinity cross-reactions. These are very common in food allergens.

The standard commercial assays for specific IgE do not utilize negative controls for each allergen and are usually not performed in duplicate. Controls also may be necessary for verifying that certain sera do not bind to the matrix and do not exhibit high nonspecific binding (27). It may be necessary to dilute some high total IgE sera before measuring specific IgE antibodies. Methods that use larger amounts of serum and longer incubation times more commonly exhibit high non-specific binding. Because multiple specific IgE tests usually are performed on each serum and several sera are typically tested at one time, there are errors that can result from mistakes in pipetting serum or the solid phase as well as from spilling labeled antibody, losing or transferring solid phase, or sticking to tubes. These can be decreased substantially by performing the assay in duplicate and rejecting results that do not agree within 35%.

Correlation of Specific IgE Antibody and Skin Testing with Clinical Disease

Either skin testing or specific IgE antibody testing can be used to determine the etiology of allergic disease (28). Intradermal skin testing is the most sensitive and least specific method. Prick and puncture skin tests are somewhat less sensitive and more specific; and are the most commonly employed methods for determining etiology. Specific IgE antibody testing is significantly less sensitive and appears to be more specific than skin tests, when compared to provocation tests. There are some relative indications for specific IgE antibody testing as described in Table 66.4.

TABLE 66.4. ADVANTAGES OF *IN VIVO* AND *IN VITRO* METHODS FOR DIAGNOSING ALLERGEN-SPECIFIC HYPERSENSITIVITY

Advantages of Skin Testing	Advantages of *In Vitro* Testing
High specificity	Useful when skin disease (e.g., eczema or dermatographism) preclude skin testing
High sensitivity	Useful in epidemiological studies of atopy
Results with minutes of application	Useful for identifying cross-reactive allergens (RAST inhibition)
Low cost	

The correlation of specific IgE antibody results with skin test results is allergen dependent. For most reasonably well-characterized allergen extracts the range is from 60% to 98% correlation, with an average of 75% to 80% agreement. Allergen extracts used for skin testing and for preparing solid-phases for *in vitro* testing vary widely in potency and composition. The standardized potency extracts used for skin testing for common aeroallergens are not optimal for preparing solid phase reagents. There are no standards of potency for most extracts, which vary by orders of magnitude among both manufacturers and lots. Source materials are not always consistent and proprietary methods may be used in preparation. Fungal extracts are a particular problem with both identity of genus and species and variation among strains, and with continued passaging in culture. Although most extracts are glycerinated to control proteolysis and degradation, this does not always provide a high degree of stability. Diluted extracts may degrade rapidly and may need to be replaced daily.

The finding of the presence of specific IgE antibodies by either skin testing or laboratory testing does not establish the presence of allergic disease. A complete medical evaluation correlating the presence of IgE-mediated reactivity with clinical history, medical findings, and symptoms is necessary. In some cases, challenge testing may be necessary. A significant fraction of the population has IgE-mediated reactivity with either subclinical or inactive allergic disease. The presence of IgE against foods and hymenoptera venoms is common particularly in patients with no histories of adverse reactions. IgE antibodies against a series of related foods in patients with a single food sensitivity may reflect low affinity cross-reactivity without clinical manifestations.

Cell-Mediated Immunity

Patch Test Scoring

The patch test results are interpreted 72 hours after their application. Before removing the test strips, care must be taken to ensure that sites where the test materials were placed can be identified accurately. The sites must be analyzed by an experienced clinician who is adept at interpreting the differences between irritant and true allergic responses. If only mild erythema is present at the site, the reaction should be recorded as doubtful. A site with erythema and some induration and possible papule formation may be interpreted as weak. Strong reactions have erythema with papules and vesicles, and extreme reactions will have bulla formation. Irritant reactions, if strong, will have a sharply demarcated border of erythema, but in allergic reactions, the border will not be so clearly defined. It is very difficult, however, to distinguish between weak irritant and weak allergic reactions.

In Vivo Challenges

Spirometry

Besides skin testing for IgE-mediated disorders, respiratory challenge in association with spirometry offers one of the most sensitive, objective determinations of an individual's clinical reactivity to a given allergen or other provoking substance. A positive response usually is defined as a drop in the FEV-1 by 20% below baseline values. If a positive response is found to aspirin, cross-sensitivity to other NSAIDs is implied because of the pharmacologic ability of these compounds to inhibit cyclooxygenase. Because some NSAIDs are more potent in this ability than others, adverse reactions may be more severe with compounds other than aspirin. Patients should be advised to avoid all aspirin and aspirin-containing compounds and be given a list of currently available NSAIDs. They should be told to notify their pharmacy on the assumption that other NSAIDs may appear on the market in the future. Acetaminophen usually is tolerated but also may cause a drop in the FEV-1 when given at high doses to NSAID-sensitive individuals. Should a patient who has a positive aspirin challenge require a NSAID, a desensitization usually can be accomplished. The patient can then take the drug without adverse effects as long as the medication is taken regularly. If several days elapse between doses, sensitivity may recur. Persons found to be sensitive to sulfite and other food additives should be counseled about foods and beverages that contain the material. Accidental exposure may occur, however, and patients should have a ß-adrenergic inhaler or epinephrine for subcutaneous injection should a reaction occur. Persons with adverse reactions to environmental materials in the workplace should be relocated. Retraining may be necessary for extremely sensitive individuals if avoidance in the particular environment is not possible. Persons with a hypersensitivity pneumonitis must avoid the material to which they are sensitive or face the possibility of a progressive obstructive and possibly restrictive pulmonary disease.

Allergen challenge tests can be used as a definitive method for diagnosis of respiratory allergy. These methods are used routinely in Europe, but primarily restricted to research centers in the United States. Bronchial inhalation challenge should be performed by a standardized method (28), using spirometry to determine reactivity. Nasal inhalation challenges can use several methods to detect reactivity including rhinomanometry, sneeze reflex, mediator production and blockage. Conjunctival provocation is less expensive, because it does not require instrumentation.

Reproduction of Clinical Symptoms

In some *in vivo* challenges, it is difficult to obtain an objective measure as sensitive as spirometry. The method of choice for diagnosis of non–life-threatening food allergy is the double-blind, placebo-controlled food challenge (29). Skin testing and/or specific IgE antibody determination should be used in patients with histories of severe anaphylaxis. Before challenge testing the patient should be maintained on a strict elimination diet for the suspected foods, especially if the manifestations are of a chronic nature.

Sting challenge testing is considered by many to be the most reliable method to evaluate *Hymenoptera* venom allergy (30). Various studies show 25% to 60% of patients with recent histories of systemic reactions from insect stings react to an intentional sting challenge. This method is usually restricted to research centers and, for the most part, has been supplanted by skin testing.

Hypersensitivity Pneumonitis

Precipitins can be found in the serum of most individuals suffering from hypersensitivity pneumonitis. However, precipitins can be detected in the serum of as many as 50% of persons with similar but asymptomatic exposure. Therefore, clinical data including chest radiographs are necessary for establishing the diagnosis. Lymphocyte transformation to antigen is seen less often (16%), but this is not a readily available laboratory test and is done only in research centers at present. Definitive diagnosis may require extended removal from exposure and rechallenge. Antibody tests are of very limited value in the diagnosis of hypersensitivity pneumonitislike conditions because of exposure to small molecules including isocyanates, formaldehyde and other reactive chemicals. Challenge tests with reactive chemicals are potentially dangerous and may result in long-term or irreversible changes. Histories of exposure and spirometric and radiographic evaluation, and lung biopsy, when needed, should be sufficient for diagnosis of these cases.

Allergic Bronchopulmonary Mycosis

Diagnostic criteria for allergic bronchopulmonary mycosis (ABPM) include the presence of asthma, a positive immediate skin test reaction to the fungus, an elevation of the total IgE, and precipitins to the fungus. Additional unessential criteria include the radiographic findings of infiltrates and central bronchiectasis. The best screening test for allergic bronchopulmonary aspergillosis (ABPA) is the presence of a very high level of specific IgE against *Aspergillus* allergens. Tests for IgE antibody against single allergens that exhibit a high correlation with the presence of ABPA are under development. Responsiveness to therapy can be monitored by following the total serum IgE and the chest radiograph. Persons with sufficiently advanced disease may require continuous corticosteroid therapy. Remissions can be achieved in patients in whom the disease has been detected at an early stage, before irreversible changes have occurred.

Hereditary Angioedema

A patient with angioedema who had the onset of clinical symptoms before the age of 20 years and who has a low serum C4 should be suspected of having hereditary angioedema (HAE). If the C4 is low, an immunologic assay for the C1 esterase inhibitor (C1EInh) should be obtained. A low level confirms the diagnosis of HAE. In 15% of individuals with HAE, however, the immunologic assay may be normal, indicating the presence of an antigenic material in plasma that is functionally inactive. In cases of angioedema where the C4 is low and the immunologic assay for C1EInh is normal, a functional assay for C1EInh should be obtained. A low functional assay confirms the diagnosis.

In patients with angioedema beginning in the third decade or later, a C4 and C1EInh should be determined. Many cases of acquired C1EInh deficiency are caused by immune complexes activating C1 with the subsequent binding and consumption of C1EInh. If abnormal, a search for another underlying disorder should be sought, as indicated in the methodology section.

REFERENCES

1. Snyderman R. Mechanisms of inflammation and tissue destruction in the rheumatic diseases. In: Wyngaarden JB, Smith LH Jr, eds. *Cecil textbook of medicine*. Philadelphia: WB Saunders, 1988:1984–1998.

2. Sly RM. Mortality from asthma, 1979+1984. *J Allergy Clin Immunol* 1988;82:705–717.

3. Kalliel JN, Goldstein BM, Braman SS. High frequency of atopic asthma in a pulmonary clinic population. *Chest* 1989;96:1336–1340.

4. Zimmerman B, Feanny S, Reisman J, et al. Allergy in asthma. I. The dose relationship of allergy to severity of childhood asthma. *J Allergy Clin Immunol* 1988;81:63–70.

5. Van Bever HP, Stevens WJ. Suppression of the late asthmatic reaction by hyposensitization in asthmatic children allergic to house dust mite (*Dermatophagoides pteronyssinus*). *Clin Exp Allergy* 1989;19:399–404.

6. Pepys J. Occupational allergic lung disease caused by organic agents. *J Allergy Clin Immunol* 1986;78:1058–1062.

7. Butcher BT, Bernstein IL, Schwartz HJ. Guidelines for the clinical evaluation of occupational asthma due to small molecular weight chemicals. *J Allergy Clin Immunol* 1989;84:834–838.

8. Mathison DA, Stevenson DD. Aspirin sensitivity in rhinosinusitis and asthma. *Immunol Allergy Practice* 1983;11:340–349.

9. Nicklas RA. Sulfites: a review with emphasis on biochemistry and clinical application. *Allergy Proc* 1989;10:349–356.

10. Champion RH. Urticaria: then and now. *Br J Dermatol* 1988;119:427–436.

11. Frank MM, Gelfand JA, Atkinson JP. Hereditary angioedema: the clinical syndrome and its management. *Ann Int Med* 1976;84:580–593.

12. Gelfand JA, Boss GR, Conley CL, et al. Acquired C1 esterase inhibitor deficiency and angioedema: a review. *Medicine* 1978;58:321–328.

13. Portnoy JM, Moffitt JE, Golden DBK, et al. eds. Stinging insect hypersensitivity: A practice parameter. *J Allerg Clin Immunol* 1999;103:963–980.

14. Sampson HA. Role of immediate food hypersensitivity in the pathogenesis of atopic dermatitis. *J Allergy Clin Immunol* 1983;71:473–480.

15. Richerson HB, Bernstein IL, Fink JN, et al. Guidelines for the clinical evaluation of hypersensitivity pneumonitis. *J Allergy Clin Immunol* 1989;84:839–844.

16. Michel FB, Skassa-Brociek W, Segalen C, et al. Reproducibility of six methods of skin prick tests. *J Allergy Clin Immunol* 1986;77:222 (abstract).

17. Demoly P, Michel F-B, Bousquet J. In vivo methods for study of allergy: skin tests, techniques, and interpretation. In: Middleton E Jr, Reed CE, Ellis EF, etal., eds. *Allergy principles and practice*, 5th ed. St. Louis, MO: CV Mosby, 1998:430–439.

18. Homberger HA. Methods in laboratory immunology. In: Middleton E Jr, Reed CE, Ellis EF, et al., eds. *Allergy principles and practice*, 5th ed. St. Louis, MO: CV Mosby, 1998:417–429.

19. Ownby DR. Clinical significance of immunoglobulin E. In: Middleton E Jr, Reed CE, Ellis EF, et al., eds. *Allergy principles and practice*, 5th ed. St. Louis, MO: CV Mosby, 1998:770–782.

20. Fisher AA. Role of patch testing. In: Fisher AA, ed. Contact dermatitis. Philadelphia: Lea & Febiger, 1986;9–29.

21. Stevenson DD, Simon RA. Sensitivity to aspirin and nonsteroidal antiinflammatory drugs. In: Middleton E Jr, Reed CE, Ellis EF, et al., eds. *Allergy principles and practice*. 5th ed., St. Louis: CV Mosby, 1998:1225–1234.

22. Bush RK, Taylor SL. Adverse reactions to food and drug additives. In: Middleton E Jr, Reed CE, Ellis EF, et al., eds. Allergy principles and practice. 5th ed., St. Louis: CV Mosby, 1998:1183–1198.

23. Smith AB, Castellan RM, Lewis D, et al. Guidelines for the epidemiologic assessment of occupational asthma. *J Allergy Clin Immunol* 1989;84:794–805.

24. Cartier A, Bernstein IL, Burge PS, et al. Guidelines for bronchoprovocation on the investigation of occupational asthma. *J Allergy Clin Immunol* 1989;84:823–829.

25. Greenberger PA. Allergic bronchopulmonary aspergillosis. *J Allergy Clin Immunol* 1984;74:645–653.

26. Hamilton RG, Adkinson NF. Clinical laboratory methods for the assessment and management of human allergic diseases. *Clin Lab Med* 1986;6:117–138.

27. Hoffman DR. Comparison of methods of performing the radioallergosorbent test: Phadebas, Fadal-Nalebuff and Hoffman protocols. *Annals Allergy*, 1980;45:343–346.

28. Bernstein IL, Storms WW. Practice parameters for allergy diagnostic testing. *Annals Allergy Asthma and Immunology*, 1995;75:543–625.

29. Sampson HA. Food allergy. Part 2: diagnosis and management. *J Allerg Clin Immunol* 1999;103:981–989.

30. Blaauw PJ, Smithuis OLMJ, Elbers ARW. The value of an in-hospital sting challenge as a criterion for application or omission of venom immunotherapy. *J Allergy Clin Immunol* 1996;98:39–47.

67

Receptor Assays of the Clinical Laboratory

Barry G. England

James B. Smart

The coordinated behavior of all cells within a multicellular organism is necessary if the organism is to function normally. Intercellular communication is a prerequisite of coordinated cell behavior and must be maintained throughout growth and development of the organism, as well as during the period of adult homeostasis. Any serious breakdown in the communication system inevitably leads to malfunction of one or more cells within the organism, perhaps leading to a disease process, or a more serious disruption that may result in death. A variety of communication or signaling mechanisms exist within the body that include nervous connections, short and long distance chemical signaling compounds that are borne by blood or lymph, and intracellular and transcellular signaling compounds that perform an autocrine or paracrine communication function. It is the interplay between the various signaling compounds and their respective target cells that maintains normal cell growth and function. Loss of communication may lead to uncoordinated growth of a cell, or group of cells, that if not corrected, may result in the uncontrolled and aggressive growth of a subset of cells (such as cancer) with resultant injury to normal tissues. A hormone is “a chemical substance, formed in one organ or part of the body and carried in the blood to another organ or part. Depending on the specificity of their effects, hormones can alter the functional activity, and sometimes the structure, of just one organ or of various numbers of them.” Blood-borne hormones are equally accessible to all vascularized tissues, yet it was recognized very early by Bayliss (1) that hormone action was restricted to specific target tissues. This observation lead to the concept of a recognition site in the target cells that is capable of concentrating the hormone and an effector site that elicits the specific hormone action (2). Nonresponsive tissues lack the hormone binding and/or effector sites.

The past two decades have witnessed great advances in the knowledge of receptors and hormone-mediated receptor action in hormone responsive tissues. This chapter will be devoted to the discussion of a selected few of the host of cell regulatory or signaling substances (hormones, growth factors) and their receptors, that provide for specificity in hormone/growth factor interaction with cells and tissues within the organism. In this discussion, we will refer to the hormone or any other signaling substance that binds to a receptor as a ligand. Formation of a ligand-receptor complex in a cell initiates a sequence of reactions that elicits a function specific to that cell. Target cells have receptors that are specific for that cell type, although a cell may have many different receptors, each of which regulates a different process within the cell. The binding affinity of a receptor-ligand complex defines the strength of attraction between the ligand and the receptor, whereas binding *specificity* describes the ability of the receptor to recognize a specific ligand.

CURRENT CONCEPTS IN RECEPTOR STRUCTURE AND FUNCTION

Part of "67 - Receptor Assays of the Clinical Laboratory"

Cell growth and differentiation is regulated by a variety of chemical regulators that act via cell surface receptors or via intracellular receptors. These include hormones that are released into the blood and act on distant targets, and growth factors and cytokines that effect regulation of cell function through autocrine or paracrine mechanisms. Several regulators such as steroids, vitamin D, retinoic acid, and thyroid hormones interact with intracellular receptors, which bind to hormone dependent transcription regulatory regions (hormone response elements) on DNA and effect gene expression. Other regulators of cell function are hydrophilic molecules that exert their action upon the cell through their respective cell surface, or membrane-bound receptors. These include the peptide hormone products of the endocrine glands, and a host of other regulatory compounds, including growth factors that exhibit local control of cellular growth. Amongst these are the peptide class of regulators, known collectively as peptide regulatory factors (PRF). The nomenclature used in the identification of PRF has been based upon their biological action, or the assay system used in their original identification and isolation. For example, two growth factors that were isolated from transformed cells and shown to be capable of inducing the phenotypic transformation of untransformed cells, were identified as transforming growth factors (TGFα and TGFß) (3). Epidermal growth factor (EGF), gonadotropin releasing hormone (GnRH), somatostatin (SS), insulinlike growth factor I (IGF-I), insulinlike growth factor II (IGF-II), and the several fibroblast growth factors, are all growth regulating agents for which receptors have been found in normal and malignant

tissue (4). Cellular receptors can be conveniently classified into two general groups based upon their location within the cell: (i) those that are located within the cell (nucleus or cytoplasm), and (ii) those that are associated with the plasma membrane. Extensive investigations currently are underway to ascertain the role these molecules may play in the growth of normal and tumor tissue, and to examine their utility in determining the prognosis or treatment of endocrine responsive tumors.

Intracellular Receptors

The first steroid receptor was described by Jenson (5) who characterized the binding of tritiated estradiol in tissues that respond to estrogenic stimuli (estrogen target tissue). Radioactive estradiol administered to the animal was found only in target tissues and other gonadal steroids did not accumulate in estrogenic target tissue (uterus, vagina, and pituitary). Early characterization of the estrogen receptor relied upon the use of sucrose density gradient separation of nuclear and cytosolic cell fractions, which yielded semipure fractions of the estrogen binding protein. Further purification was accomplished by gel electrophoresis and column filtration methods. In these early receptor studies, it was observed that unoccupied receptor was recovered from the “cytosolic” fraction of the cell and that only steroid-receptor complexes were found in the nucleus. This observation led to the two-step theory of steroid receptor action. In step 1 the lipophilic steroids freely migrate through the plasma membrane into the cell where they bind to the unoccupied receptor and the complex then is transported rapidly to the nucleus where the step 2 is binding of the complex to the DNA (6).

Estradiol, and the other ligands that bind to intracellular receptors, usually are hydrophobic and lipophilic, a property that facilitates their movement through the lipid-rich cell membrane. These hormones alter the pattern of gene expression and are responsible for slower and lasting effects on the cell. Although it generally had been well accepted in the 1970s and 1980s that steroid hormone receptors were present in the cytoplasm of the cell, and that rapid translocation of the complex to the nucleus preceded gene expression; understanding of the nature and action of intracellular receptors continues to evolve. In target cells that previously had been exposed to hormone, the hormone-receptor complex is bound tightly to the nucleus and is no longer found in the soluble fraction (cytosol) of tissue extracts. It now is thought that unoccupied gonadal steroid receptors are present in the nucleus in a loosely bound form and are contained in the “cytosolic” fraction of tissue extracts because it is removed easily from the nuclear envelope with low salt buffers. Data verifying the intracellular localization of steroid receptors awaited the development of specific monoclonal antibodies against the steroid receptors. These studies yielded a wealth of information regarding the location of these receptors within the cell, as well as permitting immunoaffinity purification of these labile proteins. Immunocytochemical staining of frozen sections of human breast cancer tissue showed clearly that the gonadal steroid receptors, which appear in the cytosol after tissue homogenization, are localized within the nuclear compartment of the intact cell. Similar observations have been made using the technique of enucleation to minimize contamination between the cytoplasm and nucleoplasm (7). Thus, procedures capable of localizing gonadal steroid receptors within intact cells or isolated cellular components have shown that ER and PR are confined to the nucleus of the responsive cells (8). The exception seems to be that unoccupied receptors for glucocorticoids and mineralocorticoids are located in the cytoplasm and the ligand-receptor complex translocates to the nucleus. Therefore, current interpretation of the location and function of nuclear receptors is more complex than originally thought and must accommodate an increasing variety of nuclear receptors and subforms of several of those receptors.

The successful cloning of nuclear receptors during the 1980s and the introduction of molecular biology techniques to the study of nuclear receptors has greatly increased our understanding of the structure-function relationships among the various nuclear receptors, and their similarities and differences (9, 10, 11, 12, 13). This receptor superfamily comprises approximately 150 distinct proteins that show similarities in their molecular structure, that function via a similar ligand-dependent mechanism of action, and are thought to share a common ancestry. Their primary function is to reconcile the transcriptional response of target cells to the gonadal steroids (androgens, estrogens, and progestins), adrenal steroids (glucocorticoids and mineralocorticoids), thyroid hormones, retinoid hormones, or Vitamin D3. Members of the nuclear receptor superfamily (Fig. 67.1) share the following characteristics; (i) a variable amino-terminal transcription activating domain (TAD) that shows specificity of binding to the respective ligand, (ii) a central and well-conserved DNA-binding domain (BDB) that binds to specific hormone response elements (HRE) on the nuclear DNA, and (iii) a carboxy-terminal ligand-binding and dimerization domain (LBD) that interacts with various factors of the transcriptional machinery, and the DNA-binding domain (DBD) (14).

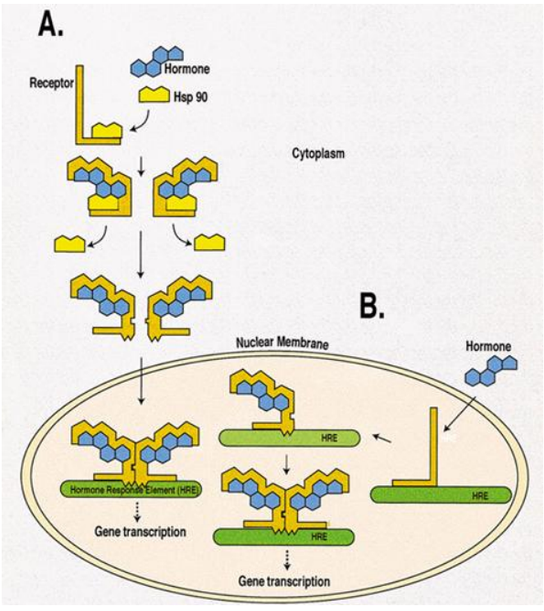


FIGURE 67.1. Members of the nuclear receptor superfamily contain (i) a variable amino-terminal transcription activating domain (TAD) that shows specificity of binding to the respective ligand, (ii) a central and well conserved DNA-binding domain (BDB) that binds to specific hormone response elements (HRE) on the nuclear DNA, and (iii) a carboxy-terminal ligand-binding and dimerization domain (LBD) that interacts with various factors of the transcriptional machinery, and the DNA binding domain (DBD). Type I nuclear receptors in the absence of the cognate hormone are maintained in an inactive state. Heat shock protein (Hsp) 90 (and perhaps Hsp70 and Hsp56) inhibitory proteins bind to the receptor, covering the DBD, and prevent binding to DNA. Ligand binding to the receptor induces a conformation change in the ligand-receptor complex that affects the release of the Hsp binding proteins. Loss of Hsp converts the receptor to a smaller, active form (4S) capable of binding to DNA. Hormone binding and the resultant dissociation of the inhibitory protein enable formation of homodimers, which bind with high affinity binding to the DNA HRE. The DBD of approximately 70 amino acids is highly conserved and forms two fingers (putative DNA-binding site) that are stabilized by a chelated Zn^{++} ion and interact specifically with HRE on the target genes. Activation of receptor-mediated gene transcription then occurs in response to the activated ligand-receptor homodimer complex binding to the HRE.

The nuclear receptors have been divided into three subfamilies within the nuclear receptor superfamily (15). Type I, classical or steroid receptors include those that bind androgens, estrogens, progestins, glucocorticoids, and mineralocorticoids. Type II receptors include those that bind the thyroid hormones, retinoic acid, and vitamin D3 (14). A third class of nuclear receptors (Type III) has been identified through cDNA clones that encode for polypeptides with structural features suggestive of cryptic steroid hormone receptors. These putative nuclear receptors that have not been linked with any known biologically active hormone or ligand, are described as orphan nuclear receptors (16, 17, 18).

Within the Type I steroid hormone receptors subfamily, there are two different forms for the following steroid receptors: glucocorticoid receptors α and ß, progesterone receptors A and B, and estrogen receptors α and ß. Type I nuclear receptors in the absence of the cognate hormone are maintained in an inactive state. Heat-shock protein (Hsp) 90 (and perhaps Hsp70 and Hsp56) inhibitory proteins bind to the receptor, covering the DBD, and prevent binding to DNA. Ligand binding to the receptor induces a conformation change in the ligand-receptor complex that affects the release of the Hsp-binding proteins (19, 20) from the 8S form of the receptor. Loss of Hsp converts the receptor to a smaller, active form (4S) capable of binding to DNA (6, 21, 22, 23). Hormone binding and the resultant dissociation of the inhibitory protein enables formation of homodimers,

which bind with high-affinity binding to the DNA HRE (17). The DBD of approximately 70 amino acids is highly conserved and forms two *fingers* (putative DNA-binding site) that are stabilized by a chelated Zn^{++} ion (24) and interact specifically with HRE on the target genes. Activation of receptor-mediated gene transcription then occurs in response to the activated ligand-receptor homodimer complex binding to the HRE. The mechanism by which activated receptor binding to DNA effects gene transcription, although incompletely understood, is under active investigation (15, 16, 17). It appears to involve a hormone- and ATP-dependent phosphorylation-dephosphorylation cycle (25). Some altered proteins appear to have the ability to “turn on” the transcription of certain proteins, without evidence of binding to DNA (26). Another highly conserved domain of approximately 250 amino acids, the ligand-binding domain (LBD), exhibits 70% to 95% homology for the same steroid receptor when examined in different species (27). Binding affinity of the receptor to DNA increases dramatically upon formation of the hormone-receptor complex.

Type II receptors of the nuclear receptor superfamily include

the thyroid hormone receptor, the vitamin D receptor (28) and the retinoid receptors (29). These receptors are not associated with Hsp molecules and bind to DNA even in the absence of ligand. Receptors bound to DNA in the absence of ligand interfere with the preinitiation complex and cause “silencing” of gene transcription. Binding of the ligand to the DNA associated Type II receptor reverses the “silencing” effect and initiates gene activation. Modulation of gene transcription often requires the presence of additional protein factors that bind to DNA response elements (enhancers, repressors, silencers). The specific transcription factors also may be activated by receptor phosphorylation. Structural changes of certain receptors within this superfamily have been described, which affect the ability of the receptor to regulate normal cell function, e.g., the variant T3 receptor, rat (c-erb-Aα2) binds DNA but lacks the ability to bind T3 or to activate gene transcription (30, 31, 32).

Neoplastic changes could be envisioned in which hormonal control of cell growth is lost because of the inability of an altered receptor to bind ligand, or alternatively DNA binding may be lost and ligand-binding ability maintained. Either of these changes may prevent the hormonal signal from reaching the specific gene and eliciting its subsequent action.

The identification of the orphan nuclear receptor subfamily has stimulated research into the discovery of new hormone response systems in which the receptor's ligand/function are currently unknown. The existence of a large number of potential new receptors offers the opportunity to develop novel therapeutic agents that may be efficacious in the treatment of a variety of illnesses. For a more complete review of this intriguing area of research see (18, 28).

Membrane-Bound Receptors

The membrane-bound receptors possess three distinct segments: (i) an extra cellular domain that includes a recognition site that exhibits ligand-binding specificity, (ii) a transmembrane domain, and (iii) an intracellular domain containing an effector or catalytic site. Membrane-bound receptors can be classified broadly into three different categories: allosterically activated enzymes; receptors coupled to an enzyme system via the G protein; and membrane-bound receptors that form a channel through which the transport of ions and other small molecules into and out of the cell is regulated. Water-soluble hormones, growth factors, or PRF occupy binding or recognition sites on the extracellular domain of membrane-bound receptors and generate an intracellular signal by eliciting a hormone-receptor response that is specific to the occupied receptor. Receptor-mediated responses may include activation of any of a number of receptor-associated enzymes such as: adenylate cyclase, guanylate cyclase, phosphodiesterases, protein kinases, phospholipases, and lipid methylases. Receptors also may elicit their response by interacting with stimulatory or inhibitory G proteins that are part of a complex allosteric enzyme system, or they may control the transport of specific ions or selected small molecules into or out of the cell. Other receptor actions may result from an interaction with select cytoskeletal elements, membrane lipids, chromatin proteins, or the receptor enzymatic activity may be regulated by hormone mediated phosphorylation or glycosylation. In contrast to intracellular receptors, activation of membrane receptors usually results in a rapid response by the target cell, usually within seconds or minutes.

Tyrosine Kinase Receptors

The single membrane-spanning tyrosine kinase receptors include: insulin, epidermal growth factor receptor (EGFr), platelet-derived growth factor receptor, fibroblast growth factor receptor, hepatocyte growth factor receptor, and Ephlike receptor protein kinases (33, 34). These receptors (Fig. 67.2) are monomeric molecules, with a single membrane-spanning moiety and tyrosine kinase activity on the intracellular domain. Upon binding to the ligand, the receptor dimerizes and undergoes conformation changes that activates the tyrosine kinase, autophosphorylating the tyrosine residues in the intracellular domain of the receptor. In addition, autophosphorylation of the tyrosine kinase receptor enhances receptor kinase activity toward cytosolic and membrane substrates that are external to the receptor. Signaling proteins bind to the phosphorylated tyrosine sites through a specialized domain, the src homology domain (SH2) that recognizes the phosphotyrosines in specific amino-acid sequences. The unique combinations provide binding specificity for the interaction of each receptor with its respective SH2-containing signaling proteins. Binding of the signaling proteins to the phosphorylated intracellular domain may elicit any of several actions; the recruitment of soluble proteins in the cytosol to the plasma membrane, allosteric activation or inactivation of specific enzymes, enhancement of mitogenic effects and cell growth, control of Ca^{++} metabolism, induction of gene transcription, activation of glycolytic enzymes, stimulation of cell motility, etc.

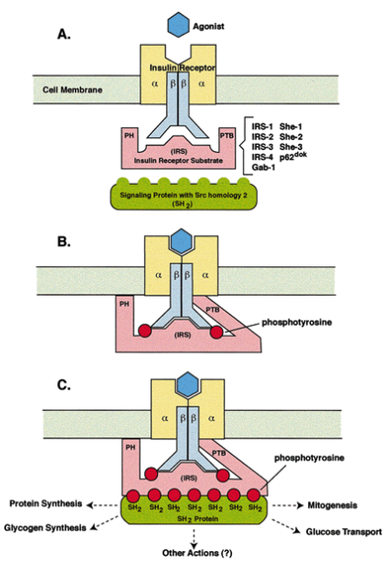


FIGURE 67.2. The single membrane-spanning tyrosine kinase receptors are monomeric molecules, with a single membrane-spanning moiety and tyrosine kinase activity on the intracellular domain. Upon binding to the ligand, the receptor dimerizes and undergoes a conformation change that activates the tyrosine kinase with autophosphorylation of the tyrosine residues in the intracellular domain of the receptor. Signaling proteins bind to the phosphorylated tyrosine sites through a specialized domain, the src homology domain (SH2) that recognize the phosphotyrosines in specific amino-acid sequences. Binding of the signaling proteins to the phosphorylated intracellular domain may elicit any of several actions; the recruitment of soluble proteins in the cytosol to the plasma membrane, allosteric activation or inactivation of specific enzymes, enhancement of mitogenic effects and cell growth, control of Ca^{++} metabolism, induction of gene transcription, activation of glycolytic enzymes, stimulation of cell motility, etc.

The insulin receptor also belongs to the family of receptors with intrinsic tyrosine kinase activity. However, this receptor is a disulfide bound heterotetramer with two α subunits and two ß subunits. The α subunits form the extracellular binding site for insulin and the ß subunits form the transmembrane and intracellular aspects of the insulin receptor molecule. Ligand binding to the extracellular binding site causes autophosphorylation of the receptor but does not create effective binding sites for the SH2-containing signaling proteins. Rather the phosphorylated intracellular domain binds to one of a family of insulin receptor substrate (IRS) phosphoproteins. Phosphorylation of tyrosines on the IRS protein creates docking sites for the various signaling proteins that transmit the signal downstream. The final cellular effects include actions such as the translocation of vesicles containing GLUT4 glucose transporters from the intracellular pool to the plasma membrane, activation of protein or glycogen synthesis, initiation of gene transcription, etc. Several IRS domains have been described that interact with sites on the receptor, with structural components of the plasma membrane, or with SH2 domains on signaling proteins. The pleckstrin homology (PH) domain recognizes polar lipids in the plasma membrane that are adjacent to the receptor, the phosphotyrosine binding (PTB) domain that recognizes phosphotyrosine in the amino acid sequence asparagine-proline-x amino acid-phosphotyrosine (NPXpY), and multiple phosphotyrosines that are bound by the SH2 domains of the signaling proteins. SH2 adaptors often have SH3 domains that recognize specific amino acid sequences of other intracellular signaling proteins (34).

IRS form a growing family of proteins that are phosphorylated, bind to the insulin receptor, and provide docking sites for the SH2-containing signaling proteins. Nine members of this fam-ily have been described to date (IRS-1, IRS-2, IRS-3, IRS-4, Gab-1, three Shc isoforms, and $p62^{dok}$). The PH and PTB domains of the IRS recognize phosphotidylinositides in plasma membrane locations adjacent to the insulin receptor and phosphotyrosine 960, which forms part of the NPXpY recognition site on the insulin receptor ß subunit just inside the membrane, respectively. The exact function of these IRS binding domains and their respective recognition sites is not presently known. Two different types of SH2-containing proteins have been described,

SH2-adaptor proteins and SH2-enzymes. Each of the two SH2 protein types consists of several different entities with their own unique actions that account for the multiplicity of cellular effects observed in response to insulin stimulation. Mutational changes in the insulin receptor are rare and unlikely to account for the insulin resistance that is observed with non–insulin-dependent diabetes nor has a single susceptibility gene for type 2 diabetes been identified. However, the multiplicity of insulin receptor signaling cascades that have been identified and genetic and nongenetic factors that can disturb insulin signaling could account for the wide variety of insulin sensitivities observed in diabetes [For a detailed description of the alterations in the insulin signaling cascades see Virkamaki (33).].

G Protein-Coupled Receptors

The guanine nucleotide-binding protein (G protein) coupled, seven-transmembrane segment, receptors comprise the largest family of membrane-bound receptors with an exceptional variety of agonists that utilize members of this family of receptors. Approximately 1,000 different receptor members of this superfamily have been identified since the first G protein receptor was cloned (35, 36, 37). Not only do many of the peptide and glycoprotein hormones, biogenic amines, lipids, nucleotides, and ions utilize members of this superfamily of receptors, they also are used to convert exogenous stimuli such as light, odors, and taste to intracellular signals. The receptors do not have inherent enzymatic activity, nor do they form an ion channel through the plasma membrane. The G protein-coupled receptors (Fig. 67.3) function as activators of G protein in the intracellular space. The function of the G protein is to transmit the receptor signal to effector molecules on the plasma membrane. The various heterotrimeric G proteins are made up of an α subunit (40 to 46 kd), a ß subunit (35 to 38 kd), and a γ subunit (7 to 9 kd). The ß and γ subunits, which are tightly bound and function as a single unit, are loosely bound to the α subunit forming a single protein complex. The α subunit contains a guanine nucleotide-binding site that binds either GDP or GTP. The G protein (αßγ heterotrimer), with GDP bound to the guanine nucleotide-binding site, is associated with the resting stage receptor. Binding of the hormone, or agonist, to the receptor induces a conformation change in the G protein complex, reducing the binding affinity for GDP, which leaves the guanine-binding site and is rapidly replaced with GTP. Upon binding GTP to the α subunit, the α subunit-GTP complex disassociates from the βγ complex, yielding two protein complexes that are capable of eliciting a cellular response by binding with effector proteins at the plasma membrane. The specific cellular action of the many elicited by agonist-G protein-coupled receptors depends upon the agonist, its cognate receptor, and the type of G protein and effector proteins that are activated. There are four families of G proteins: those that stimulate adenylate cyclase (G_s), those that inhibit adenylate cyclase (G_i), those that stimulate phospholipase C (G_q), and those that regulate ion channels (G_{12}). The activated G proteins carry signals along the plasma membrane but do not cross the cytoplasmic space. Rather, they regulate the concentration of small second messengers that can diffuse across the intracellular space, by the activation or inhibition of specific “second messenger” enzymes. The activated GTP-α subunits regulate intracellular concentrations of second messengers by controlling the cAMP enzyme (adenylate cyclase), and the phospholipase Cß enzyme [inositol 1,4,5-triphosphate (IP_3) and 1,2-Diacylglycerol (DAG)]. $G_{s\alpha}$ stimulates adenylate cyclase, Gi_{α} inhibits adenylate cyclase, and $G_{q\alpha}$ activates phospholipase Cß. $G_{ß\gamma}$ subunits regulate the activity of adenylate cyclase II and IV, phospholipase C subtypes ß1, ß2 and ß3, phospholipase A, K^+ and Ca^{++} channels and form links to the tyrosine kinase pathways (34). The second messenger DAG stimulates protein kinase C, a membrane bound enzyme that phosphorylates serine and threonine amino-acid–containing side chains in proteins. IP_3 induces the release of stored Ca^{++} from the endoplasmic reticulum (ER) by binding to a calcium channel in the ER. Calcium and cAMP are important intracellular mediators of a number of cellular events including gene transcription. Calcium and cAMP can both synergize or antagonize in their regulation of gene expression through actions on several transcription factors, including the cyclic AMP response element binding (CREB) protein. The dephosphorylated form of CREB, normally bound to the transcription response element, is inactive until serine-133 is phosphorylated. In general, cAMP promotes gene transcription through the protein kinase A mediated phosphorylation of serine-133, and calcium prevents transcriptional activation by the phosphorylation of serine-142. However, another calcium dependent kinase, calcium-calmodulin–dependent protein (CaMK IV), promotes transcription by phosphorylating serine-133. Thus, calcium can promote or prevent the activation of gene transcription.

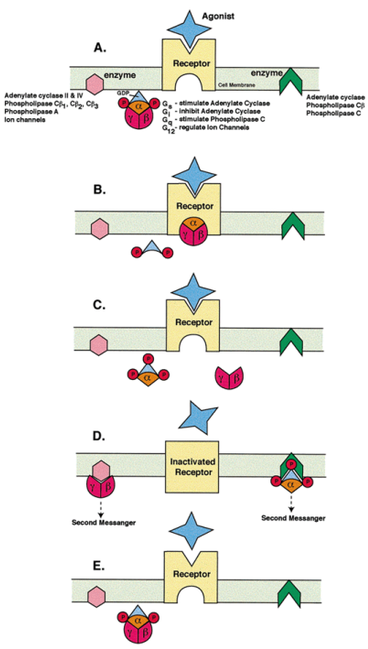


FIGURE 67.3. The G protein coupled, seven-transmembrane segment receptors comprise the largest family of membrane bound receptors. The G protein functions as a transmitter of the receptor signal to effector molecules on the plasma membrane. The various heterotrimeric G proteins are made up of an α subunit (40 to 46 kd), a ß subunit (35 to 38 kd), and a γ subunit (7 to 9 kd). The ß and γ subunits, which are tightly bound and function as a single unit, are loosely bound to the α subunit forming a single protein complex. The α subunit contains a guanine nucleotide-binding site that binds either GDP or GTP. The G protein (αßγ heterotrimer), with GDP bound to the guanine nucleotide-binding site, is associated with the resting stage receptor. Binding of the hormone, or agonist, to the receptor induces a conformation change in the G protein complex, reducing the binding affinity for GDP, which leaves the guanine-binding site and is rapidly replaced with GTP. Upon binding GTP to the α subunit, the α subunit-GTP complex disassociates from the βγ complex, yielding two protein complexes that are capable of eliciting a cellular response by binding with effector proteins at the plasma membrane. The activated GTP-α subunits regulate intracellular concentrations of second messengers by controlling the cAMP enzyme (adenylate cyclase), and the phospholipase Cß enzyme [inositol 1,4,5-triphosphate (IP3) and 1,2-Diacylglycerol (DAG)]. $G_{s\alpha}$ stimulates adenylate cyclase, $G_{i\alpha}$ inhibits adenylate cyclase, and $G_{q\alpha}$ activates phospholipase Cß. $G_{ß\gamma}$ subunits regulate the activity of adenylate cyclase II and IV, phospholipase C subtypes ß1, ß2 and ß3, phospholipase A, K^+ and Ca^{++} channels and form links to the tyrosine kinase pathways (34). The second messenger DAG stimulates protein kinase C, a membrane bound enzyme that phosphorylates serine and threonine amino-acid–containing side chains in proteins. IP_3 induces the release of stored Ca^{++} from the endoplasmic reticulum (ER) by binding to a calcium channel in the ER.

The deactivation or turning off of a hormone signal is as important to appropriate cellular function as is signal induction. This is true in the normally functioning cell and may be extremely important in neoplastic cells. Desensitization can occur rapidly within seconds to minutes, or it can occur gradually over the course of several hours or days. Some receptors are desensitized through the phosphorylation of the intracellular domain of agonist-stimulated receptor by specific enzymes. These phosphorylated receptors are able to bind agonist but are unable to stimulate the Gs protein. Removal of the phosphate groups by protein phosphatases will reverse the desensitization over time.

Receptors also can be desensitized through a process called receptor-mediated endocytosis. Agonist-stimulated receptors aggregate on the surface of the cell and are internalized in an endocytotic vesicle. The internalized vesicles eventually fuse with lysosomes where the agonist-receptor complex is degraded or a certain percentage are recycled and returned to the surface as active receptors. This method of recycling of receptors is not very efficient and *de novo* synthesis of receptors is necessary to restore receptor numbers to normal.

Receptors and Disease

Disease States (Steroids and Their Receptors)

Binding of a hormone to its receptor is required for receptor activation and subsequent biological action. Any molecule that can occupy the receptor-binding site without activating the receptor can prevent hormone action by competitive inhibition. Attachment

of a molecule, such as an antibody to an epitope, to a location on the receptor that is in close proximity to the binding site can prevent hormone action by steric hindrance. Either of these events would inhibit normal receptor function. In addition to these inhibitory actions on receptor function, any change in the molecular structure of the hormone or its receptor that prevents binding, or reduces binding affinity between the hormone and receptor, can reduce or completely prevent hormone action on the target cell. Each of these events has the potential of altering cell function. Clinical disease may result from an unwanted negative

effect of receptor function, or a form of disease treatment may be realized by regulating cell growth through the use of synthetic ligands that will prevent binding and subsequent biological action of endogenous hormones or growth factors. This type of therapy may be particularly useful in the regulation of hormone dependent cancers. A number of clinical examples of receptor mediated disease, or treatment of disease, have been described. Estrogen and progesterone receptor concentrations are widely used in the determination of prognosis and appropriate treatment regimens to be followed in several endocrine responsive tumors, with breast cancer being the best example. Long Acting Thyroid Stimulator (LATS) (38, 39, 40) stimulation of TSH is an example of a disease process in which antibody stimulation of normal receptor results in overproduction of a hormone stimulated gene product. Myasthenia gravis is a defect in which the acetylcholine-mediated nerve impulse transmission is impaired and is an example of the steric hindrance of the hormone-binding site by endogenous antiacetylcholine receptor antibody (41). Testicular feminization syndrome (TFS) and familial hypercholesterolemia (FH) are two examples of disorders resulting from receptor changes; the lack of function or absence of the androgen receptor in TFS, and of the low-density lipoprotein (LDL) receptor in the case of FH (42).

The first report implicating hormonal regulation of breast cancer was in 1896, when Sir George Beatson (43) reported the dramatically beneficial effect of oophorectomy in premenopausal women with advanced breast cancer. The first evidence for the existence of tissue receptors for estrogens was obtained *circa* 1960 when Jensen and Jacobson (5) and Glascock and Hoekstra (44) reported the specific uptake of ^{3}H-estrogen in target tissue after systemic administration of labeled hormone. Folca (45) administered tritiated estrogens to women with breast cancer and observed the uptake of radioactivity by the tumor. It was later shown that a significant accumulation of ^{3}H-estradiol occurred in tumors from patients who responded to hormonal therapy while those that were unresponsive to therapy lacked the receptor. The presence of progesterone receptor (PR) in human breast cancer tissue is indicative of estrogen responsiveness (46). The repertoire of prognostic cellular receptors that may be used as markers that are predictive of, or may be predictive of, tumor response to endocrine and chemotherapeutic treatment regimens is expanding and now includes both intracellular and membrane-bound receptors. The gonadal steroid receptors ER and PR are used widely in the prognosis and treatment of breast, endometrial, vulvar, and cervical cancer while active investigations are examining the potential role that receptors for many of the PRF may play in prognosis and treatment of a variety of cancers. Several of the small peptides, oncogene products, growth factors, and their respective receptors currently are being examined as potential candidates for improving prognostic accuracy and possible sites of drug interdiction in cancer treatment regimens. These include EGF, TGFα, TGFβ, luteinizing hormone-releasing hormone, and somatostatin; and the c-*erb*-B2 (AKA neu) oncogene product, as well as Ki67, a nuclear proliferation antigen.

Normal endocrine responsive, or target cells, contain intracellular or membrane-bound receptor sites for each of the hormones known to influence the growth and function of the target tissue. When malignant transformation of a cell occurs, the cell may retain all or part of the normal receptor population, or undergo a structural change in the receptor that alters its function. If the cell retains intracellular receptor sites, steroid control of cell growth and function probably will be maintained as in a normal cell. However, malignant transformation that induces structural alterations may cause the receptor to lose the ability to bind the steroid, or to bind to the HRE on cellular chromatin and thereby lose its transcriptional function. These cells may no longer recognize circulating hormones and lose endocrine, autocrine, or paracrine-mediated functions and fail to respond to endocrine therapy (47). The regulation of the growth and function of endocrine target tissues, e.g., breast, and male and female genital tissues, and endocrine-responsive tumor growth within these tissues, in many cases is dependent upon hormonal factors produced by endocrine glands, and by growth factors elicited by both normal and malignant cells. The roles that peptide hormones, gonadal steroids, autocrine and paracrine growth factors play in the regulation of growth and function of normal and malignant tissue are becoming more clear, almost on a daily basis, as are the roles of the increasing number of recently identified oncogene protein products. Many of these recent advances are being put to use in the clinical laboratory as aids in determining the prognosis and appropriate treatment of cancers of various endocrine responsive tissues.

Current clinical use of receptors includes the analysis of intracellular receptors for estrogens, progesterones, androgens, and glucocorticoids. Perhaps the best-known example of the clinical utility of laboratory analysis of receptor function is that of estrogen and progesterone receptors in predicting patient response to adjuvant therapy or endocrine therapy in breast cancer. The membrane-bound receptors for EGF, TGFα, TGFβ, and the oncogene protein product from the gene c-erb-B2, also have been employed as predictors of response to hormonal therapy and prognosis in carcinomas of the breast, prostate, uterus, ovaries and cervix, and may be of some assistance in treating glucocorticoid responsive leukemias. The basic molecular structures of intracellular and extracellular receptors have been determined by the work of a number of research groups, with continuing investigations that relate structure with function.

The clinical utility of tumor levels of ER is well accepted and widely used, even though only about 60% of breast cancers contain ER. Approximately two thirds of the ER(+) tumors respond to endocrine therapy and between 5% and 10% of ER(-) tumors also respond to endocrine therapy. Estrogen is known to regulate the production of the progesterone receptor and the determination of both steroid receptors in breast tumors increases the prognostic value of hormone receptor measurements. Approximately 70% of PR(+)ER(+) tumors, and 25% to 30% of PR(-)ER(+) tumors respond to hormonal therapy and the prognosis of receptor positive mammary tumors is better than receptor negative carcinomas. Because the measurement of steroid receptors in tumor tissue does not give absolute information regarding tumor response to endocrine or adjuvant chemotherapy, but rather provides an imprecise indication of the response that individual tumors will display, other factors undoubtedly play a role in the regulation of malignant tumor growth. Until the putative cell regulatory factors have been determined,

identification of ER or PR receptors in malignant tumors will continue to provide the best standard of care for predicting which patients should receive endocrine therapy or adjuvant chemotherapy.

DISEASE STATES (GROWTH FACTORS AND THEIR RECEPTORS)

Part of "67 - Receptor Assays of the Clinical Laboratory"

The protein mitogen, EGF, a 6 kd, 53 amino acid single chain polypeptide, plays an important role in regulating the growth of many ectodermal and mesodermal derived cells and is found in significant levels in the urine, saliva, tears, bile, prostatic fluid, seminal fluid, milk, and sweat. The receptor for EGFr has been identified in most normal cell types with an abundance of receptors in brain, thyroid, lung, liver, skin, placenta and fetal membranes. EGFr apparently is involved in fetal growth and has been implicated in a number of development events, including palate and skin differentiation, growth of follicles, eye opening, tooth eruption, lung maturation, gut and liver growth and differentiation of neurons. EGFr are present in normal mammary epithelium (48), on the basal layer of epithelial cells, in normal urothelium, and on cervical and vulvar squamous-cell surfaces (49). EGF binding sites have been identified in a variety of tumors including breast cancer cells (50), in bladder malignancy where EGFr were found throughout several layers of cells in the urothelium with a rich presence on the surface layer, and in psoriatic skin lesions. There appears to be increased expression of EGFr in some ectodermal-derived tumors over that found in contiguous normal tissue (51). Although it is well known that cell proliferation is the predominant response of normal cells to EGF, the relationship between increased EGFr and tumor growth has only recently come under intensive study. At the present time, there is no consensus on the clinical utility and prognostic value of EGFr determination in tumor tissue. Points of consensus appear to be that approximately 45% to 50% of breast cancer patients display EGFr (+) tumors, and there is a negative correlation between the levels of EGFr and the steroid receptors ER and PR. Perhaps the wide variety of methodologies used in the determination of EGFr tissue levels, the lack of uniform measures of positivity (cutoff values) of the test, and the relatively small number of patients studied, has hampered the clinical utility of this test. Many investigators are calling for standardization of EGFr methodologies, similar to what was done with the ER and PR methods, in an attempt to make the test more useful and facilitate interlaboratory comparisons. A wealth of information regarding EGF and EGFr (and other growth factors and their receptors) is available in a number of different reviews (52) and an abbreviated review follows in this chapter. This line of research certainly will lead to a better understanding of the endocrine, autocrine, and paracrine control of normal and malignant tissue, and may aid in treatment of all types of cancer. If the uncontrolled growth of cells that is characteristic of cancer is subject to some combination of endocrine, autocrine, or paracrine control, as seems likely, then understanding the systemic and local regulation of GF and GFr may provide new treatments for endocrine responsive cancers, and perhaps those that appear to be independent of endocrine control.

The virus growth factor EGF-like protein also binds to the EGFr and is thought to be used for cell access by the virus. Other members of the family of small proteins that bind to EGFr are transforming growth factor α (TFGα), platelet derived growth factor (PDGF), and amphiregulin. While EGF and TFGα bind with almost equal affinity, amphiregulin has a significantly lower affinity for the receptor. Amphiregulin is a weaker growth stimulator than either EGF or TFGα except for the stimulation of keratinocyte growth where it is equivalent to the other two proteins. Since the binding affinity of amphiregulin for the EGFr is much lower than the other two proteins, it is natural to assume that there is a primary binder for amphiregulin, but this site has not yet been found (53).

EGF and the TGFs α and ß (54) are capable of enhancing or inhibiting *in vitro* cell growth in human endometrial carcinoma cell line (RL95-2) cultures depending upon GF concentration and cell plating density. TGFß inhibits cell proliferation at both high and low plating density and induces the appearance of large cuboidal cells that are distinct from EGF or TGFα treated cells. EGF, TGFα and TGFß effects appeared to be similar in hCG producing tumors transplanted into nude mice (55). Low concentrations of EGF stimulated tumor growth and increased EGFr levels, whereas high EGF concentrations inhibited tumor growth and reduced EGFr in the tumors. TGFα mimicked the effects of EGF and actively competed for EGFr binding sites, whereas TGFß competed only slightly with EGF for EGFr binding sites. EGF and TGFα levels have been compared between extracts of malignant and nonmalignant tissues (56). Approximately 30% of ovarian, endometrial, and cervical carcinomas and 16% of breast carcinomas had elevated GF levels compared to normal tissues. In nonmalignant tissue, the mean GF concentration was 1.5±0.7 ng/mL, whereas, concentrations in endometrial, cervical, and breast carcinomas were significantly higher (4.2+1.5 ng/mL). Increased GF levels may indicate an autocrine or paracrine control of growth in neoplastic tissue through autologous control of GF production by the tumor. An action that is independent of endocrine control or perhaps loss of endocrine control of cell growth as the neoplastic process proceeds.

Distribution of EGFr appears to differ between normal and malignant urothelium. Messing showed that most (95%) normal urothelium contained EGFrs in the basal layer of the epithelial cells, whereas in patients with urothelial carcinomas there was a significant concentration of EGFrs on the surface as well as in the deeper layers, and EGFr density was correlated with tumor grade (57). Tumor growth and EGFr levels appear to be correlated, and an increase in the production of EGFr may enhance the autocrine or paracrine stimulation of cell proliferation and may thus relate to tumor growth and metastatic potential. Gene amplification and overexpression of EGFr and mRNA of the erb-B2 oncogene have been found in mammary carcinomas and the increased EGFr levels may provide a growth advantage in cells in which over-expression of the EGFr has occurred. In contrast, EGFr concentration and prevalence is lower in endometrial adenocarcinoma than in normal endometrium (58, 59). Further, tumor grade and its relationship to EGFr is confusing. Berchuck (58) reported no correlation for these two variables, whereas Reynolds (59) showed a clear decrease in EGFr levels with advancing

tumor grade. Reynolds (60) also has shown that EGFr can be regulated in endometrial adenocarcinomas by estradiol and progesterone. Endometrial cells cultured with estradiol showed a decrease in EGFr in normal and malignant tissue while progesterone treatment increased EGFr concentrations. Increasing density of EGFr sites on the cervical and vulval squamous-cell carcinomas has been related to the aggressiveness of this type of tumor. In particular, it was found that levels of EGFr >100 fM/mg of protein had poor prognosis while levels <100 fM/mg of protein had more favorable outcomes (relative to recurrence and five year survival) (49). ER concentration was inversely correlated with EGFr in normal and neoplastic breast and endometrial tissue, whereas IGF-I binding was not affected by estradiol or progesterone treatment (61), perhaps indicating that down regulation of EGFr is one of the biological effects of estradiol.

Certain oncogene products show striking similarities to some GF receptors. The c-erb-B2 oncogene protein product is a molecule in which the transmembrane and cytoplasmic regions are homologous to the 170 kd single membrane spanning receptor, EGFr. This glycoprotein has an extracellular domain that is highly protease resistant, heat stable and contains the EGF binding site, an intracellular domain that contains intrinsic tyrosine kinase activity (62), and a transmembrane portion that connects these two domains. The cytoplasmic region of the EGFr is homologous to the protein product of the v-erb-B1 oncogene. The c-erb-B2 is an oncogene that has been localized to chromosome 17, codes for a 185 kd transmembrane protein that, although structurally similar to EGFr, lacks part of the extracellular EGF recognition site.

It is noteworthy that while the capacity of the receptors increased, the affinity did not. This leads to the conclusion that it is simply an increase in the number of EGFrs and not a change in the actual receptor. In some cases, amplification of gene coding has been demonstrated. However, overexpression of this gene did not always result in increased EGFr. Other factors that might come into play here are: increased transcription of the gene (over other genes); reduced degradation of the mRNA allowing more “readings” by the ribosome; and reduced receptor degradation which might keep the cell “turned on” longer.

Growth factors with TKA stimulate cell growth and proliferation by attaching phosphates to proteins through the tyrosine or serine residues, effect cell proliferation, and may induce the phenotypic transformation of cells. The enzymatic activity is directed against several protein substrates, including the receptor itself. This latter activity includes autophosphorylation and evokes phosphorylation of a variety of cellular substrates, which is the initial step of signal transduction (63). TKA is constituently active in the oncogene protein product whereas EGFr requires the presence of EGF bound to the extracellular moiety of the receptor for kinase activity to be initiated.

Comparisons of GF receptors in normal endometrium and endometrial cancers are important because tyrosine kinase genes have the potential to mutate to an oncogene form and produce malignant transformations. Information gained by the determination of GF receptor concentrations and their binding affinities in normal and neoplastic tissues may lead to various new treatment regimens and provide prognostic information that is not currently available. Treatments may take the form of: (i) biologically inactive analogs of growth factors that are capable of binding to and inactivating the receptor, (ii) receptor-specific antibodies that recognize epitopes on the extracellular domains of the receptor and sterically hinder binding of the GF to the receptor, (iii) antibodies with attached cytotoxins that can be released into the cell following endocytosis, or (iv) regulation of GF receptor production with estrogens, progesterones, or their biologically inactive competitive inhibitors.

There appears to be a clear relationship between the overexpression of EGFr and clinically aggressive malignant disease. EGFr is a significant predictor of response to systemic endocrine treatment in metastatic breast cancer patients and levels of EGFr correlates with the aggressiveness of the disease (64). Cultured A431 cancer cells overexpress EGFr and EGF treatment of this cell type stimulates tumor cell proliferation, cellular protein phosphorylation, and matrix metalloproteinase (MMPs) secretion, suggesting that these effects may be mediated at least partly through EGFr (65). The complexities of the mechanism of action of EGFr may provide treatment opportunities, through the inhibition of signaling molecules, such as the EGFr family and their associated ligands, in a number of areas of cancer therapy, including proliferative, angiogenic, invasive, and metastatic aspects. Evidence of the potential benefits has been obtained from a multiplicity of laboratory experiments, although clinical trials are needed to evaluate the therapeutic advantages of such agents (66). In recent years, strategies to modulate either EGFr, or the downstream signal beyond the cell surface receptor have been investigated. One promising strategy involves the use of anti-EGFr monoclonal antibodies, either alone or in combination with conventional cytotoxic modalities such as chemotherapy or radiotherapy (67).

Because overexpression of TGF-α and its receptor (EGFr) also has been associated with aggressive disease, and a poor prognosis, the blockade of EGFr activation has been proposed as a means of anticancer therapy. Topotecan, a cytotoxic drug that specifically inhibits topoisomerase I in ovarian, breast, and colon cancer cell lines, which express functional EGFr and the monoclonal antibody C225, an anti-EGFr humanized chimeric mouse MAb that is presently in Phase II clinical trials in cancer patients, have been evaluated. Combination therapy showed an *in vitro* dose-dependent inhibition of growth when cancer cells were in a sequential schedule. Treatment with MAb C225 also markedly enhanced apoptotic cell death induced by topotecan. Almost complete tumor regression was observed in all tumor-bearing mice treated with the two agents in combination. These results support the evaluation of the combination therapy of topoisomerase I inhibitors and anti-EGFr blocking MAbs in clinical trials (68).

EGF and TGF ß1 exert opposite effects in most cells. MDA-MB-231, a breast carcinoma cell line that overexpresses EGFr, demonstrated EGF-mediated reduction in TGF ß1 levels in culture media collected after 11 days of growth. The converse was also true, as TGF ß1 reduced the secretion of EGF in media from cultured cells, suggesting regulatory relationship between EGF and TGF ß1 (69). In cultured cancer cells, EGF and TGF ß1 elicits an increase in type IV collagenases, stromelysin and urokinase-type plasminogen activator and thymidine phosphorylase/platelet-derived endothelial cell growth factor. Thus, EGF

and TGF ß1 appear to act as positive regulators on the invasion process of gynecological tumor cells through their stimulatory action on the motility of tumor cells, the expression of proteases and the angiogenic phenotype (70). Combination antitumor therapy aimed at the inhibition of mitosis and induction of apoptosis is also a potentially useful approach for cancer therapy. The cell growth inhibitor - CGP52411- and the naturally occurring inducer of apoptosis TGF ß1, alone and in combination decreased growth rate in two ovarian carcinoma cell lines (PEO1 and OVCAR3). In both cell lines, TGF ß1 enhanced apoptosis and reduced EGF stimulated cell growth. Dual therapy showed an additive effect in PEO1 (71). The observation that preexposure of tumor cells to EGF alters their response to cisplatin and lymphokine-activated killer (LAK) cytolysis, an action that depends upon the degree of EGFr expression, may prove helpful in preselection of patients for appropriate therapy (72). Luteolin (Lu) and quercetin (Qu), two flavonoids with potent antiinflammatory, antimutagenic, anticarcinogenic, and anticancer effects, inhibit EGFr tyrosine kinase activity and autophosphorylation of EGFr and EGF promoted growth of A431 cells and thus may be good anticancer and antimetastasis agents (65).

Thus, in summary, preclinical and clinical data have established a link between EGF receptor/ligand expression and poor prognosis in cancer patients. Whereas, recent progress in understanding the mechanism of EGFR action has aided in the design of more effective tyrosine kinase inhibitors with improvements in potency, specificity, and *in vitro* and *in vivo* activity. The recent availability of very specific, irreversible inhibitors of the EGFR family that provide unique pharmacological properties, exceptional efficacy and improved *in vivo* performance has led to promising clinical trials. A review by DW Fry (1999) addresses the targeting of the EGFR family for cancer therapeutics, and highlights promising kinase inhibitors that are in development (73).

METHODOLOGIES

Part of "67 - Receptor Assays of the Clinical Laboratory"

Intracellular Receptor Analyses

In 196l Folca (45) demonstrated that the uptake of [3H]-hexestrol *in vivo* by breast cancer tissues was higher in four patients who responded to adrenalectomy than in six others who did not respond to the surgery (74). This observation led to the development of *in vitro* techniques using [3H]-estradiol to study tissue binding in primary and metastatic breast tumors. Estrogen receptors in organs have been measured *in vitro* using either tissue slices or cell-free preparations. The first quantifications were performed by sucrose-density gradient (SDG) analysis, which demonstrates the molecular forms of the receptor (75). While these early studies concerned themselves with ER, the importance of understanding the role of all receptors in general has become apparent. This has led to a need to quantify and characterize receptors as a class and the last few years have witnessed a flurry of activity aimed at this goal. The longer history of study of ERs has produced a body of knowledge about these receptors that is not yet accumulated for the other receptors. Thus some of the specific comments in the methodology descriptions may only be applicable to ER. Even though these methods have general applications, the newer receptors (e.g., EGFr) suffer from the lack of accepted standards and protocols so quantitative comparisons are often not meaningful. Of particular importance has been the recent application of monoclonal antibodies to this endeavor, which has led to the development of techniques that are rapidly replacing the other methodologies.

In addition to new methodologies, our models also are changing. In particular, until recently, ERs were thought to reside in the cytosol in the unbound state and were transferred to the nucleus when they became activated through binding to estradiol. Studies with newer methodologies have shown us that these receptors in fact are all in or associated with the nucleus. Thus, when reviewing some of the more “classical” procedures it is necessary to keep in mind that the model of intracellular ER location has evolved. Finally, it should be noted, that the only assays currently performed in the clinical laboratory are the Dextran Coated Charcoal (DCC), immunoassays and Immuno Cytochemical Assays (ICA).

Specimen Collection and Preparation

Hormone receptors are heat-labile proteins. It is necessary that they be stored immediately at 4°C if the sample is assayed within an hour. Otherwise, they should be quick frozen and stored in liquid nitrogen. Dry ice may be utilized for short storage (a few days) if liquid nitrogen is not available. Lyophilization of cytosol and storage at 4°C also has been used. The receptor is pH sensitive with maximum stability at pH=7.4. Low protein content can lead to underestimation of receptors in the DCC assay. Thus, with a low concentration of receptors, a minimum of 3 mg/mL of protein is suggested, but for higher concentrations of receptors, the protein concentration can be as low as 1 mg/mL. Miller (76) has reported that inclusion of molybdate (10 to 20 mM) in the extraction buffer increases the assay levels of both PR and ER. A tumor with an ER or PR concentration of 3 to 10 fM/mg of protein or higher usually is considered estrogen receptor-positive. There is not complete agreement between laboratories regarding the exact cutoff value, some laboratories have suggested that recent information shows the value of 10 fM/mg is more clinically relevant.

Receptor inactivation can occur within intact tissues, e.g., through prolonged exposure of excised intact tumor tissue to high temperatures. Further, steroid receptors may be inactivated during the tissue handling procedures before and during the cytosol preparation through high temperatures or excess homogenization. Finally, steroid receptors may be inactivated after they have been partitioned into the cytosol fraction of the tissue during the receptor assay procedure. These cytosolic receptors may be inactivated by such agents as high temperature, high ionic strength, pH changes, exposure to active proteases or phosphatases, and oxidation of sulfhydryl groups.

Physical Separations

Centrifugation

DCC Assay. Determination of receptor concentrations using dextran-coated charcoal utilizes the adsorptive properties of

charcoal to remove small and large-molecular-weight compounds from solution at different rates (77, 78). Small-molecular-weight compounds are removed very quickly, leaving the larger compounds in solution. This procedure is very time and temperature dependent with dextran being used to partially coat the charcoal particles and reduce the adsorptive properties of charcoal thereby reducing the rate and extent of adsorption of large molecular weight compounds such as proteins. This has been the most widely used methodology for the separation of free and receptor-bound radioactive ligands in the quantification of most steroid receptors. It is a sensitive and popular assay procedure and has formed the basis for most of the quality assurance, or proficiency testing programs such as the College of American Pathologists and Southwest Oncology Group. The assay is performed by using increasing concentrations of radiolabeled hormone and a concurrent range of unlabeled hormone competitor to titrate the receptor to a saturation end point. Specific and saturable binding can be determined by subtracting the nonspecific (nonsaturable) binding from the total binding to both specific and nonspecific sites. Specific binding can be converted to the total number of binding sites, generally expressed as femtomoles per milligram of protein. The dissociation constant is a measure of the affinity of the receptor for the steroid and is approximately equal to the concentration of free steroid at which one-half of the receptors are saturated. Kilodalton is the reciprocal of kilo annum and also can be obtained from a Scatchard plot (79). The values obtained from the DCC method correlate well with those of the sucrose gradient methods (see below). It is sensitive, inexpensive and relatively easy to perform. A modification designed to hasten this assay is to use only two data points, and assume a linear relationship. This can give workable results but if more than one binding site is present on the receptor, the value for kilodalton is greatly underestimated (80).

Sucrose Density-gradient Ultracentrifugation. This was the first ER assay applied in a routine clinical setting. The measurable quantity here is sedimentation velocity of proteins as a function of MW and density. It is a long procedure because of the required centrifugation times (16 hour in a swinging bucket rotor). The cytosol is incubated with labeled hormone and centrifuged through a sucrose gradient. The layers are counted and compared to known markers [e.g., BSA (4.6S), and IgG (6.8S)] with established molecular weights. Resolution is limited but most tumors display the presence of an 8S-estrogen receptor and some also the 4S form. It has been suggested that only the 8S form of the receptor complex has predictive value for the endocrine responsiveness of a given tumor (77). The presence of molybdenum in the gradient increased the amount of receptor observed in the assay. This is especially true when the progesterone receptor is being measured. This method estimates the quantity and size of the receptor molecules, but it is a slow and expensive procedure.

Precipitation. The receptor is precipitated with protamine (81) followed by incubation with excess-labeled hormone. At 4°C the receptor sites bind to the labeled hormone leaving unbound-labeled steroid in solution. The method has a very high nonspecific background and tends to give low concentrations when compared to other techniques.

When using ammonium sulfate to precipitate the receptor fraction, Chen (82) reported that at 50% saturation 77% of the ER and 53% of the PR precipitated. Saturation of less that 50% resulted in a significant reduction in the recovery of receptor. ERs and PRs salted out in the ammonium sulfate pellet are stable and store better than does the cytosol preparation.

Chromatography

Gel Filtration and Immobilized Antisteroid Antibodies. This assay is infrequently performed and has no advantages over the other commonly performed assays. After incubation of the cytosol with labeled hormone, the mixture is passed through a Sephadex column with the receptor-hormone complex eluting in the void volume (83). Sephadex does compete with the receptor for E2, which can cause some dissociation of the receptor-hormone complex (about 15%) leading to a false decrease in the amount of receptor measured. Columns packed with gels containing bound anti-E2 have been used in a similar fashion with similar results. The primary advantage of the use of this immunoaffinity method is the ability to quantify receptor levels with samples where only small volumes are available.

Hydroxyapatite Assay (83). The hormone-receptor complex is adsorbed onto hydroxyapatite either in “batch” or column form and the free steroid and plasma contaminants are removed by washing. The assay is more reliable at low protein concentrations than the DCC method, and it can be performed on small sample sizes.

DEAE Anion Exchange Chromatography. Steroid receptors are acidic proteins that will bind to anion exchange resins such as DEAE-cellulose (84). The major problems of this method include the tendency of free E2 and testosterone to bind to DEAE at low ionic strength and the inability of receptors to bind to the column at high ionic strength. Analysis of molecular forms of the receptor and separations from nonreceptor binding proteins is easily achieved.

Controlled Pore Glass Beads. The receptors bind strongly to controlled pore glass beads, followed by saturation by labeled hormone and elution from the column by ethanol for analysis (85).

High-performance Liquid Chromatography (HPLC). Although equipment is expensive, the required solvents are inexpensive and good column separation usually can be obtained between the analyte and other compounds present in the sample. There are several different kinds of HPLC analyses with the differentiation being based on the type of column used.

i. High performance size exclusion chromatography (HPSEC) utilizes polyethylene glycol (PEG) gel beads to prepare columns, which separate mixtures based on their molecular size and configuration. This has the ability to separate the various aggregates that the receptor displays. These columns are usually run at lower pressures and employ aqueous buffers. It gives the same type of information as obtained in sucrose gradient method with better resolution and in less time (86).

ii. Reverse-phase liquid chromatography (RPLC): The elution of a receptor protein from a RP column requires high concentrations of organic solvent in the mobile phase, which may irreversibly denature the proteins (87). Affinity columns have also been employed with HPLC.

iii. High-performance hydrophobic interaction chromatography (HPHIC) is the mildest application of HPLC to the separation of protein molecules with retention of their biological activity. The stationary phase is nonionic and binds the hydrophobic patches present in the protein. Unlike most HPLC procedures, which employ denaturing solvents, this column allows the use of physiological pH buffers (88).

iv. High performance ion exchange chromatography (HPIEC). This methodology relies on the same principle as the DEAE method (see DEAE Anion Exchange Chromatography, above). The column is packed with an anion exchange resin. The elution gradient proceeds from low to high ionic strength, which can be as high as 500 mM Phosphate (88).

v. High performance chromatofocusing (HPCF) is a technique that separates proteins based upon their surface charge properties. The column is a weak ion-exchange resin and the gradient is again from low to high ionic strength (88).

Electrophoresis

Agar Gel Electrophoresis. The receptor-hormone is separated and found on the anodal side of the well, while two peaks representing steroid hormone binding globulin and free steroid are found on the cathodic side (89). The method compares favorably with other methods, particularly when the cytosol is treated briefly with DCC before electrophoresis to remove most of the free and albumin-bound steroid.

Isoelectric Focusing. Separation is accomplished through charge properties. In comparison with other methods, isoelectric focusing is sensitive and rapid (1.5 to 2 hours). Vollmer (90) describes an assay employing two different isotopes (125I and 3H) for the simultaneous determinations of ER and PR with correlation coefficients to the single isotope assay of 0.93 and 0.8, respectively.

Immuno Assays

The development of the monoclonal antibody has had a significant impact on the assay methodologies for receptors. To develop the antibody, the receptor is first isolated by affinity chromatography. The respective steroid or protein is bound to the column (usually a gel) and the homogenate containing the receptor passed through. The receptor binds to the immobilized ligand and, after washing the column, is eluted off (usually, through change in pH or ionic strength). This purified receptor then is employed to develop a monoclonal antisera using standard techniques.

Immunoradiometric Assays (IRMA)

A typical immunoradiometric assay (IRMA) would incorporate two monoclonal antisera, which bind to different sites on the receptor (91). One of the antibodies is attached to a solid phase to serve as the capture antibody. After incubation of the solid phase capture antibody with the homogenate, the unreacted (unbound) material is washed away. Then the second monoclonal antibody, which is labeled with 125I, is introduced. The solid phase is given another wash and then counted. This assay gives good correlation to earlier biochemical methods. The method is independent of binding of E2 to the receptors as that apparently does not affect the antigenic sites of the two monoclonal antibodies to the receptor.

Enzyme Immunoassay (EIA)

Recently, commercial monoclonal-based EIAs were made available (92) for both ER and PR. The assay methodology is identical to the radiometric method except that a horseradish peroxidase labeled second antibody is employed. The assay is read by development of o-phenylenediamine (OPD). Results are consistent with other receptor assays (93, 94), but with somewhat higher values. The sensitivity of the PR EIA falls off for values below 10 fM/mg (95, 96), while the ER assay can be extended to 5 fM/mg (94). To date no putative threshold value for the ER-EIA has won acceptance (97, 98).

Bound Receptor Assays

The continuing studies aimed at understanding the intercellular localization of the steroid receptors have already been discussed. This had an impact on the development of ER assays. Historically, nuclear estrogen receptors were considered to be bound receptors and those in the cytoplasm to be “free” receptors. Here “bound” receptors means complexed to their respective steroid. For ER and PR, two receptor sites appear to be present. For the ER, the site with the higher affinity for estradiol has the characteristics (concentration, affinity, and specificity) attributed to the receptor. The second site, called type II EBS, displays the steroid and tissue specificity but it is not the true ERs. It is present in higher concentration but displays a lower apparent affinity. There appears to be a close relationship between increased levels of type II receptors and true uterine growth.

The currently accepted models for depicting the intercellular location of receptor sites for the ER and PR show that both the bound and free receptors reside exclusively in the nucleus. The receptors for glucocorticoid (GR) appear to reside in both the cytoplasm and nucleoplasm. This leaves us with the task of understanding what the earlier ER assays were measuring in the cytoplasm, for both biochemical and cytochemical assays. In the biochemical assays, it is necessary to first prepare a cytosol. The preparation involves extensive homogenization of the tissue, freezing and centrifugation. This treatment artificially separates the total cellular ER into nuclear and cytosolic fractions. Many researchers also feel that the free receptors are not as tightly bound in or to the nucleus and the addition of buffer in this process causes the migration of free receptors into the cytoplasm. While this can explain what is biochemically measured in the cytoplasm, the preparation of slides may not be so destructive toward the cell and may require a different explanation. One possibility is that it is the type II receptors that are being measured here.

Biochemical assays of bound receptors are based on the principle that the receptor-estrogen complex dissociates slowly at low

temperatures and rapidly at higher temperatures. The nucleus contains both bound and free receptors. To effect an exchange with the exogenous labeled estradiol the nuclear fractions are incubated in the presence of labeled estradiol at elevated temperatures (usually 25°C to 37°C). The relative quantities of free to bound receptors vary according to the hormonal milieu at the time of tumor excision. In premenopausal or a postmenopausal tamoxifen treated patient, the unbound receptors should be lowest and bound receptors highest, while the opposite would be true in postmenopausal women. Thus it is sometime desirable to evaluate both occupied and unoccupied receptor sites. In these assays, conditions are controlled (e.g., temperature) to reduce the amount of exchange.

Sucrose Pad Nuclear Exchange Assay (99)

This approach incorporates centrifugation of tissue homogenates through 1.2 M sucrose, which removes the large amounts of lipid present. This effects reduction of the NSB from 80% to 10% to 40%. The lipid components sediment only to the buffer/sucrose interface, while the desired nuclear material continues through to the bottom of the tube. The SPA can be modified to optimally measure the type II EBS that is thought to be related to estrogen-dependent growth. The assay is very tedious requiring technical skill and cannot be used as a single point assay.

Protamine Sulfate Assay (100)

By incubating a KCL-extracted nuclear fraction at 25° to 30°C for several hours, the previously bound hormone can be exchanged, while at 4°C only unoccupied receptor sites bind to labeled hormone. A major disadvantage of this method is degradation of the receptor at the higher temperature due to the proteolytic activity found in human tumor nuclear extracts.

Hydroxyapatite Exchange Procedure (101)

Both bound and unbound receptors bind to hydroxyapatite (HAP). This is similar to the procedure with protamine sulfate, except that the labeled hormone is incubated with the (HAP) bound receptors at the requisite temperature. This method appears to be superior to the protamine sulfate assay.

Mersalyl Exchange Method (102)

Exchange of bound nuclear receptor complexes may be achieved at 0 to 4°C in the presence of molybdate and Mersalyl acid. Specific binding is subsequently measured by the hydroxyapatite procedure. This assay does not lead to receptor degradation or incomplete exchange as often occurs with temperature dependent exchanges.

NaSCN Exchange Assay (103)

This assay uses NaSCN during the exchange process, which solubilizes the receptors, while exchange is carried out at 4°C in the presence of NaSCN

Analysis of mRNA Using Polymerase Chain Reaction (PCR) Techniques

McGuire's group (104) describes a simplified PCR technique sensitive enough to detect ER mRNA in breast tumor specimens from 1 µg of total RNA. In a preliminary evaluation of this method on a small series of breast tumors, the ER message was found in tumors that were positive by the DCC assay. One tumor that was ER-negative by the ligand binding assay but positive for PR was examined by the PCR method and found to contain the ER mRNA. It is known that about 2% to 11% of breast cancers are apparently ER negative but PR positive in the DCC assay (105). It is tempting to speculate that because PR is an end product of estrogen action, one might speculate that an ER in these tumors is defective in ligand binding but capable of stimulating PR.

Membrane-Bound Receptor Analyses

The need to quantitate membrane receptors is a recent one. Analysis of membrane receptors accommodates techniques employed for cytosolic receptors with the added complication that we are now dealing with tissue. Preparation of the membrane sample is critical since it is necessary to eliminate residual cytosol, which, if present, will give artificially low results. This washing also releases most of the EGF receptor trapped in the nuclear pellet. For clinical use, one washing cycle may be adequate (62). EGFr values >10 fM/mg of protein appear to be significant and >100 fM/mg of protein is significant in cervical carcinoma.

Saturation and Displacement

The most common assays for membrane receptors are radiometric saturation and displacement assays employing 125I labeled ligands (106). As in the saturation assay for cytosolic receptors, increasing amounts of labeled ligand are reacted with fixed amounts of tissue sample. This also requires determination of NSB, which requires that each of the points be performed, both in the presence of a large excess of cold ligand and in absence of the cold ligand.

A variation of this is the displacement assay. Here a fixed amount of tissue and labeled ligand are equilibrated with increasing amounts of cold ligand. The separation of free from bound in tissue assays is usually done by centrifugation (15,000 g for 5 minutes), with the resulting tissue pellet being counted in a gamma-counter. Polyethylene glycol (PEG) has been utilized to enhance precipitation. Rapid filtration also has been utilized, with the membrane material being trapped on the filter and subsequently counted. Both the saturation and displacement assays are analyzed with Scatchard plots. Both methods give comparable binding capacities (or specific binding), but the displacement assay generally underestimates the association constant.

The saturating and displacement assays require multiple points and thus a sizable amount of tissue sample is necessary (100 µg of tissue for each assay tube) in order to reduce the tissue requirement and simplify the procedure a single point saturation assay has emerged (107). In this assay two points are needed: (i) the tissue sample with a fixed amount of labeled ligand, and (ii) the tissue with labeled ligand plus a large excess (usually about 100 fold) unlabeled ligand. The difference in binding between these two points, which is the specific binding, is converted into percent of total (B/T) and multiplied by the concentration of labeled ligand. This then is converted into femtomoles

of receptor/milligram of protein. The assay must be calibrated for each type of receptor and type of tissue to determine the optimum amount of tracer (usually the 50% point on a Scatchard curve) and cold ligand.

Immunoadsorption

Tissue receptors can be solubilized with Triton X-100. However, for some receptors (i.e., EGFr) Triton X-100 also deactivates the receptor (108). This effect appears to be Triton X-100 concentration dependent and thus utilization of this methodology requires awareness of this potential problem. EGF tissue receptors solubilized with Triton X-100 are immunoadsorbed onto microtiter wells coated with mouse anti-EGFr (109). From this point, either a radiolabel or an enzyme label can be introduced.

Use of a radiolabel requires incubating the captured EGFr with increasing concentrations of 125I-EGF in the presence and absence of excess unlabeled EGF. After washing the wells and counting, a Scatchard analysis is carried out.

When an enzyme is utilized as a label, a second anti-EGFr (MoAb) is incubated with the immobilized EGFr and is subsequently analyzed through bonding of this sandwich complex to a goat antimouse IgG labeled with peroxidase. A colorimetric signal is generated through treatment of the wells with o-phenylenediamine. The results are compared to a standard curve from which the binding capacity can be read. Solubilized A431 cells were used to prepare the standard curve. While this procedure does not yield a binding constant, because only a single point is needed for the unknown sample, the amount of sample tissue needed is thus greatly reduced.

Western Blot

A Western Blot procedure has been developed for EGFr (110). Purified membranes were solubilized with nonionic detergent and resolved by electrophoresis. This is followed by electroblotting onto nitrocellulose and rehybridizing with 125I-EGF. The receptor was identified by bands at 150 and 170 kd. The presence of 1% hemoglobin and 0.05% Tween 20 during the hybridization step optimized the radiogram signal.

Cytochemical Analysis

In these procedures, the object is to detect the receptors by attaching a label to the receptor that can be visualized. Because the cell is left intact, the methodology is applicable to both intracellular and membrane bound receptors. It is performed by fixing sections of the tumor tissue in graded ethanols, on uncoated glass slides, followed by incubation with the appropriate labeling material.

Ligand-Conjugate Label

Various different types of labels have been employed, the earliest being tritium tagged estradiol which is then read by autoradiography (111). Conjugating steroid molecules to proteins to which fluorescent molecules or peroxidase are attached allows the direct visual identification of receptor sites in the cells. Early studies showed an accumulation of label in the cytoplasm and this was assumed to be depicting the location of the receptor binding site, which was consistent with the steroid binding assays where ERs were found in the cytoplasmic fraction of the tissue extracts. Quantification of the binding sites is obtained by estimating the intensity of the fluorescence. Also demonstrated by this technique was the mosaic distribution of ER positive cells in the tumor. Because there is neither uniformity of distribution of binding sites, nor of intensities, a system was developed to give an average intensity. This was to provide some correlation to the DCC assay results. Using this averaging technique the results were comparable to the reference assay (DCC).

One major problem with the cytochemical assay is that it does not supply information about the affinity of the binding site (112). Without this information, it is difficult to show that the ER sites that were being measured by the biochemical assays were the same ones visualized by the cytochemical methods. In fact, we now have good evidence that the primary ER sites, which have the highest binding affinity, reside in the nucleus. This then leaves us the question of what the earlier cytochemical analyses were identifying in the cytoplasm. It has been suggested that the cytoplasmic estrogen binding sites represent secondary, low affinity sites. Scatchard analysis of the data of DCC assays shows two binding sites for estradiol with the Kd of the second site about 100 fold less that the Kd of the first.

Antireceptor Antibody Methods (47)

ICA is performed by fixing sections in graded ethanols, on uncoated glass slides, followed by incubation with labeled monoclonal antibodies. The bound antibodies can be visualized by the indirect immunoperoxidase method (113) or through the use of fluorescent or chemiluminescent labels. Using this technique in determining the presence of ER, it was observed that the staining was almost exclusively in the nucleus as opposed to the cytoplasm where earlier studies had observed it (Interestingly, earlier cytochemical work using a tritium label also located the receptor in the nucleus but apparently these data were not widely accepted). Further, the intensity of the nuclear staining correlated to the amount of ER determined by the sucrose density gradient analysis and detectable staining was always absent in the receptor-negative tumors. Presently, the method is only semiquantitative and it usually is considered as an adjunct to the biochemical assays. However, the continued development of intensity analysis will improve the quantitative aspects of this assay. As with the cytochemical analysis, ICA does not provide any quantitative information about the binding strength of the receptor.

One of the puzzling aspects of the application of steroid therapy to the treatment of breast cancer is the large number of positive ER tumors that do not respond to this therapy (20% to 40%) and conversely, the number of ER negative tumors that do respond to therapy (about 10%). The application of cytochemistry has highlighted one facet of this problem, which is the heterogeneity within the breast tumor. There can be within the tumor, benign tissue that is ER positive, while the rest of the tumor is ER negative. Conversely, an ER-positive sample might be so diluted with fibrous tissue that it gives a negative ER bioassay.

Another interesting application combines flow cytometry and

ICA to quantify PR and simultaneously perform a DNA assay (114). This allowed the assignment of PR levels in subsets of cells segregated by their DNA content. In T47D cell, PR was present throughout the cell cycle and levels doubled in G2 and mitosis.

Localization of Receptors

Receptors have been employed for *in vivo* visualization of tumors and metastases (115) or for the localization of receptors in organs and endocrine glands (116). Intestinal adenocarcinomas and some endocrine tumors express large numbers of high-affinity receptors for vasoactive intestinal peptide (VIP). Iodine-123 labeled VIP was shown to be useful in the successful localization of tumors and metastases of gastrointestinal adenocarcinomas, carcinoids, and insulinomas. The binding of the labeled VIP to the receptors was measured by the resulting radioactivity, which was 2 to 7 times background in these carcinomas.

Fluorescent Labels

The localization of receptors within a cell always has been more of a challenge. Recently, techniques based on the use of fluorescent labels have shown great promise (117). The development of the ability to fuse green fluorescent protein (GFP) with a variety of molecular species has opened a whole new area in monitoring events in the living cell. When fused to cellular proteins, GFP retains its fluorescent properties without altering the behavior of the protein allows the observation of detailed cellular events such as nuclear translocation of receptors. Htun (118) has fused jellyfish GFP with the rat glucocorticord receptor (GR) giving the resulting GFP-GR protein chimera. Using a variant of GFP that was resistant to photo bleaching, they were able to use time-lapse video to record the course of cytoplasm-to-nucleus translocation of GR in a single living cell. When not bound to a ligand, the GFP-GR resides exclusively in the cytoplasm. Upon binding to dexamethasone (DMS), the resulting complex moves directly to the nucleus. The rate of this translocation was dependent on the DMS concentration. Shortly after the translocation of the GFP-GR-DMS complex enters the nucleus, foci of bright fluorescence appear in an organized reproducible three-dimensional structure. When the GFP-GR is activated with the agonist RU486, which is known to bind with GR but not elicit observable activity, the RU486 activated GFP-GR translocates to the nucleus and in this case accumulates almost exclusively in a reticular pattern found throughout the nucleus.

In a similar fashion, the GFR fusion protein with androgen receptor (GFP-AR) was studied (119, 120). The GFP-AR was detected through out the cytoplasm in the absence of androgen. After incubation with the hormone, the fluorescence rapidly became localized in the nucleus. The mutant GFP-AR-del4, which lacks its carboxy-terminal function and thus has no androgen-binding function, remained localized in the cytoplasm after incubation with androgen. Luciferase activity (from the reporter gene TAT-tk-luc) was used as an indicator of the transcriptional activity of GFP-AR and AR in the presence of androgens or antiandrogens. Both demonstrated a dose dependence on the induction of luciferase activity, with GFp-AR being about one-half that of AR. Studies of the trafficking of GFP-AR in the presence of various ligands also was carried out. 100% localization to the nucleus never was achieved. For synthetic androgen and dihydrotestosterone at concentrations of 10-6 M, about 65% of the fluorescence was localized in the nucleus. Similar concentrations of testosterone yielded about 55% localization in the nucleus while the two antiandrogen ligands, cyproterone acetate and hydroxy-flutamide were at about 30% localized. Carey (121) studying the same GFP-GR and DMS translocation noted that various mutants of the enzyme Ran/TC4GTPase (Ran) when present inhibited the translocation of GFP-GR to the nucleus, and that those mutants of Ran also inhibited cell growth.

Fluorescence Resonance Energy Transfer (FRET)

Mutagenesis of the GFP protein sequence generates a variety of variant forms with differing brightness and spectral characteristics. One particularly useful form is the so-called “blue” variant (blue fluorescent protein of BFP) with an emission maximum at 445μm. This is easily differentiated from the GFP, which has an emission maximum at 511 μm. These two mutants can be used to label the ligand and the receptor separately. When these labeled proteins are used in conjunction with dual channel fluorescence microscopy, it is possible to observe the movement of the two different proteins in the same cell. Formation of the ligand-receptor complex can be visualized by the “overlapping” of the two colors giving rise to an intermediate color. The light microscope has a resolution of 2,500 Å, which really is not adequate to determine whether or not these two labeled proteins are interacting.

The resolution of the light microscope can be extended dramatically through the utilization of a technique known as fluorescence resonance energy transfer (FRET). This detects the increase in emission intensity, which is the result of a transfer of excitation energy from a fluorescent donor to a fluorescent acceptor molecule. Two important requirements are that the emission energy of the donor must overlap the activation energy of the acceptor and they must be spatially oriented for this to occur. The increase in the intensity of the acceptor fluorescence is also inversely related to the distance between the two proteins. Because of this FRET can only occur over a distance limited to approximately 20 to 100 Å. However, this technique only can be used to verify interaction. A negative result does not negate the possibility of interactions. FRET successfully observed the interaction between the pituitary-specific transcription factor Pit-1 and the c-Ets-1 receptor using BFP-Pit-1 and GFP-Ets-1 as the fluorescent species (122). This interaction, as well as the association of the Pit-1 protein and the estrogen receptor has been demonstrated previously through biochemical studies. The latter association, with Pit-1 and the estrogen receptor, was not verified in this FRET system.

Fluorescence Correlation Spectroscopy (FCS)

FCS is a technique that has been around for about 20 years. It is based on observing fluorescence emitted from as little as one molecule in a very small observation volume — the confocal volume. The fluorescence is initiated by a very narrow laser beam (approximately 0.2 μm) such that only one or two molecules are excited at a time. Experimentally, one observes (measures) the

intensity and duration of the fluorescent molecule or complex, which is contained in a dilute buffer, as it diffuses through observation region. The time that fluorescent species resides in the observation region depends on the rate of diffusion. An unbound fluorescent ligand would diffuse faster through the observation region than if it was bound to a receptor because the whole system gained mass and thus diffuses slower. If more than one ligand bound to the receptor it would diffuse even slower and would also yield a more intense fluorescence (because each ligand would add a fluorescent label). The diffusion time is compared to a previously determined standard, such as just the uncomplexed fluorescent ligand or receptor from which we should be able to get a good estimate on the number of ligands bound to one receptor.

The intensity of the fluorescence also should give us an estimate of the number of ligands as well, but rotation of the complex might reduce what we observe. If unlabeled ligand were now introduced into the mixture, it would displace some of the labeled ligands and such an observed particle would show a sudden increase in diffusion (less mass). This behavior provides a model from which we can estimate binding constants (123,124 and 125).

Wohland (126) compared the data obtained from an FCS study to that obtained from a radioligand binding analysis of the serotonin type 5HT3AS cell membrane receptor using the antagonist 1,2,3,9-tetrahydro-3-[(5-methyl-1H imidazol-4-yl)methyl]-9-(3 aminopropyl)-4-H-carbazol-4-one (GR-H). Several fluorescent labels were studied. They were: fluorocein (Flu), rhodamine 6G (6G), N-[7-nitrobenz-2-oxa-1,3-diazol-4-yl] (NBD) and the cyanine dye Cy5 (Cy5). To standardize the analysis determination of the unbound labeled ligand diffusion rates were obtained. The binding studies all showed a 1:1 ligand to receptor complex, even though the 5HT3AS receptor as a homopentameric structure would suggest the existence of up to five ligand-binding sites. The 1:1 ratio was consistent with the radioligand results. The binding affinity was dependent on the fluorescent label used. The GR-6G was particularly difficult to work with as it bound to surfaces and its complexes had a strong tendency to aggregate. Using the GR-Flu it was possible to make reasonable determinations of the mass of the receptor.

Molecular Biology

The tools of molecular biology are beginning to show up in receptor research (127). These tools are helping to fill a gap in our current methods. PCR techniques can be employed to indirectly identify the existence of receptors when there are no known selective ligands for that receptor. The also can be used to study what happens to the distribution and possible change (mutation) in receptors under changes in pathological states.

To study the function of a receptor *in vivo* generally has involved the development of an animal strain with the gene responsible for that receptor eliminated (knock-out). This is fraught with difficulties, and a long-time commitment. The knock-out sometimes proves to be lethal. A new approach involves the use of gene targeting by antisense oligodeoxynucleotides. This involves the development of a short single strand, which will bind to a complementary DNA or mRNA strand. This greatly inhibits the transcription or translation of that gene. This process has been given the label “knock-down.” Lai (128), in a study of the delta opioid receptors, employed this technique. There is good evidence that there are multiple delta receptors but radioligand binding has not been able to establish clear identities. By using antisense oligodeoxynucleotides, they were able to determine the existence of at least two subtypes of the delta receptor. The next level to study in the ligand-receptor scheme is to the level of second messenger coupling. Specifically, the formation of the complex initiates a “signal” that is sent generally to up-regulate or down-regulate a process. And this information is delegated primarily to the G-proteins. While it is not the intent of this section to pursue the receptor studies to this point, it is the obvious next step and the use of antisense oligodeoxynucleotides will play a big roll in that study (129).

Errors in Receptor Measurements

Receptors are thermolabile, pH and ionic strength dependent, and they can be mechanically destroyed during homogenization. Addition of reducing agents as dithiothreitol, molybdate, or glycerol to protect from oxidation can be helpful. Low values can result from too little protein in the sample. Another problem is interference by nonreceptor binders such as albumin, corticosteroid-binding globulin, and sex-hormone–binding globulin (SHBG).

INTERPRETATION

Part of "67 - Receptor Assays of the Clinical Laboratory"

Tissue Receptors and Cancer

Familial Hypercholesterolemia

Cholesterol plays an essential role in the normal function of nearly all cells and is required by all steroid-producing cells. Cholesterol, as are most steroids, is a lipophilic molecule that circulates in blood bound to binding proteins that increase steroid solubility. Cholesterol is packaged in lipoprotein complexes referred to as high-density lipoprotein (HDL), low-density lipoprotein (LDL), and very-low-density lipoprotein (VLDL). Much of the cholesterol in the circulation is bound to LDL (LDL-C), and in normal individuals LDL-C enters the cell by binding to LDL receptors (LDLr) located in the plasma membrane. Mutational changes in the LDLr gene produce structural changes in the LDLr that may impair binding of LDL-C to the cell surface and/or inhibit internalization of the complex into the cell. One consequence of the failure of cholesterol to enter the cells is that it builds up in the plasma, thereby producing high circulating levels of LDL-C. It is well accepted that elevated plasma levels of cholesterol are responsible for plaque formation in the coronary arteries leading to coronary artery disease.

Familial hypercholesterolemia (FH) is an autosomal-dominant disorder characterized by high circulating levels of cholesterol, and the tendency to develop coronary artery disease, tendinous zanthomas, and corneal arcus very early in life. Mutational changes in the LDL gene cause abnormal forms of LDLr to be produced in FH-affected patients and prevent, or impede, the normal uptake of LDL by target cells (42). About 1 in 500 people

express a single mutant LDL gene allele and it is known that patients with one normal and one abnormal LDL receptor gene express a lower than normal number of LDL receptors and show circulating cholesterol levels that are much higher than unaffected individuals. Patients with homozygous alleles for defective LDL receptors are very rare and exhibit very high plasma levels of cholesterol during childhood with attendant coronary artery disease. These patients usually die during adolescence of severe coronary artery plaque formation. Patients who are heterozygous for LDL receptor defects can be treated with drugs that lower cholesterol and stimulate the production of LDL receptors in liver cells. The combination of decreased liver cholesterol production, increased production of LDL receptors, and measures designed to reduce intestinal absorption of cholesterol can force circulating levels of LDL cholesterol into the normal range. Advances in the analytical procedures for detecting mutational changes in the LDLr gene are extremely helpful in management of FH patients and family members. For example, single-strand conformational polymorphism (SSCP) analysis allows rapid scanning of DNA fragments for any sequence variation in situations where high-throughput mutational analysis is required. Humphries et al. describe the application of this technique in the screening of patients with familial hypercholesterolemia for LDLr gene mutations (130). Many mutations in the LDLr gene have been described in various countries around the world and disease incidence appears to be higher in certain nationalities and geographical areas with elevated cardiovascular risk. However, the identification of these individuals is difficult because of the wide variety of mutations that exist in the country. The 42 mutations that have been detected in Italy account for only about 30% of FH cases, leading to the speculation that over 100 different mutations are present in Italy (131). It also appears that certain of these genetic changes are present in areas of relatively high incidence (FH-clusters) (21i) and may represent “founder effect.” Many mutations of the LDLr gene have been described and the specific point mutations, deletions, frameshifts, inheritance patterns and geographical distribution can be reviewed on line through the “Online Mendelian Inheritance in Man” web site (132).

Cancer

Clinically, endocrine-responsive tumors are assayed for receptor concentration and affinity. These include intracellular receptors for estrogens, progesterones, androgens, and glucocorticoids. Recent developments have implicated the membrane-bound receptor EGFr and its related oncogene product from the gene c-erb-B2 in predicting the response to hormonal therapy and prognosis in breast, endometrial, prostatic, certain renal carcinomas, and may be of some assistance in treating glucocorticoid responsive leukemias.

Breast Cancer

Breast cancer accounts for 30% of all cancers and 18% of all female cancer deaths. Although there is a 10% lifetime risk of breast cancer in females, there is a 90%, 5 year, survival rate if treatment is initiated in Stage I disease. Survival drops to 60% if the cancer has metastasized before treatment is instituted (133). Mammography and self-examination are the best methods of early detection in risk populations. Prognosis and treatment are based on tumor size, clinical stage, histologic grade, axillary lymph node status, and the presence and binding capacity of ER and PR, and more recently, the experimental use of growth-factor receptors and oncogene protein products. Generally, patients with nodal metastases have a worse prognosis than those without. However, about 25% to 35% of the so-called axillary-node-negative patients die from their disease within 10 years. A further problem using nodal status as a prognosis marker is the extensive and mutilating surgery that does not enhance survival. Patients are opting for more conservative treatment.

A good prognostic marker should provide information on the tumor-cell proliferation rates and metastatic potential, and it also should indicate what therapy is appropriate. Lastly, it should be attainable with a minimum of surgical intervention. In recent years, many potential biochemical markers have been described. These included steroid hormone receptors, growth-factor receptors, activated proto-oncogenes and proteolytic enzymes (134, 135).

Estradiol and Progesterone Receptors. While receptor assays identify patients who have a significant chance of successful estrogen treatment, the overall predictive value of the information obtained from receptor assays shows considerable disagreement in the literature. In 1977, Knight (136) showed that the absence of ER in breast tumors was associated with early recurrent disease. The role of ERs has been studied since then and still is not clearly understood. Using the DCC assay to look for the presence of ER and PR in 4,000 cases of breast cancer, Thorpe (137) found three categories or types. Those that were positive for ER and PR, or negative for ER while positive for PR were hormone responsive; negative ER and PR were nonresponsive, and positive ER with negative PR gave unpredictable responses (this latter category consisted mostly of postmenopausal patients).

One confusing aspect that arises when attempting to use ER content as a prognostic factor in breast cancer lies in the assays themselves. Some assays measure only unbound (to estradiol) receptor while other assays measure total ER (see Methodologies). Thus, Syne (75) noted that there were ambiguities of response to estrogen therapy in premenopausal women containing low levels of ER, as determined by a binding assay, (which primarily measures the “free” receptor). He questioned whether the low unbound receptor level in the cytoplasm was a result of the influence of high levels of circulating estrogens or whether the tumor truly lacks estrogen receptor. Conversely, the high ER content of a breast tumor from an elderly postmenopausal woman may indeed indicate a strongly ER-positive tumor or the lack of endogenous estrogen production needed to bind the usual amount of estrogen receptor to the DNA. Other studies have shown that in some so-called “false-negatives,” there is a lack of unbound ER, but they do contain bound ER. Conversely, some false-positives (patients with ER that do not respond to estrogen therapy) show that the tumors do contain unbound ER but no hormone bound ER.

McGuire (138) has shown that for the Stage 1 cancer patient, ER status and tumor size were the two most important prognostic factors. In stage 2 breast-cancer patients however, Progesterone

Receptor (PR) status appeared to be a better prognostic marker.

Other factors have been evaluated as prognostic indicators. Hill (139) tried grouping the occurrence of ER in premenopausal patients based on body mass. He found that women with a lean body mass have a greater proportion of negative ER tumors and a higher concentration of EGFr than heavier women. In males, with the exception of a predominance of centrally located lesions and a uniquely high frequency of positive hormone receptor status, carcinoma of the breast appears biologically similar to the disease in women, and treatment should be guided by similar principles (140).

While different methodologies for receptor assays are a factor in these contradictory observations, there are other operative variables which include: (i) different cutoff points used, (ii) different patient follow-up times, (iii) different patient populations with respect to menopausal status, disease stage and adjuvant treatment, (iv) different statistical test used to analyze the data, (v) small numbers of patients used in some studies, and (vi) nonstandard criteria used to assess date of first recurrence. Also, the comparisons of immunocytochemical and the newer receptor EIAs serve to point out the importance of cellular heterogeneity as a cause of variation (141, 142). Prognostic significance of ER values does not appear to depend upon the type of tissue biopsied for assay as ER values from primary tumors or metastatic tumors predict equally well (7).

Tamoxifen, [1-(4-ß-dimethylaminoethoxyphenyl)-1,2-diphenylbut-1-ene], is a synthetic nonsteroidal compound that shows varying degrees of estrogen agonist and antagonist activity. The antiestrogenic properties of this compound have led to its clinical use in the treatment of a number of human malignant tumors especially estrogen-receptor positive (ER+) metastatic breast cancer. Tamoxifen exerts its effects, at least in part, via the specific estrogen receptor proteins of estrogen responsive cells. A secondary, saturable, high-affinity tamoxifen-binding site is present in ER+ but not ER- tumors. Raam (143) recently has described a solid phase EIA for this secondary binding site. This secondary site is present at an order of magnitude greater than the concentration of the high-affinity estradiol receptor site measured in the same tissue. In receptor assays, tamoxifen is bound predominately by this secondary binding site in ER+ human mammary carcinoma cytosol, which is distinct from the classical estrogen receptor site. No saturable tamoxifen-binding sites were detected in any of the ER- tumors studied. Elucidation of the roles if any, of this tamoxifen-binding site in mediating the antitumor activity of tamoxifen in human breast cancer awaits further experimentation (144).

Valavaara (145) in the study of 113 postmenopausal patients with advanced breast cancer that were ER positive and treated with toremifene concluded that the concentration of the ER predicts the duration of the response, but not the rate of response. In this instance, PR status was not useful as a predictor of therapy response.

Growth-factor Receptors. In breast cancer, a strong inverse relationship between EGFr and ER status was found for EGFr values >10 fM/mg of protein. There appears to be an increased expression of EGFr in some ectodermal-derived tumors over that found in contiguous normal tissue. Expression of the EGFr appears to be increased in neoplastic tissue and usually is associated with increased aggressiveness of the disease. EGFr status has been used for classifying the prognosis of ER(-) tumors. ER(-)/EGFr(+) patients had a significantly worse prognosis than did patients that were negative for both receptors (104). Scatchard analysis showed that most tumors have two EGFr binding sites. A high affinity site with K_d = 2 nM/L and the second a low-affinity site with K_d = 9.5 nm/L.

TGFα, a peptide isolated from cultures of malignant cells, stimulated the growth of malignant tissues through binding to the EGFr. There appears to be an increased expression of EGFr in some ectodermal-derived tumors over that found in contiguous normal tissue. EGFr, or the closely related erb-B2 oncogene, induced gene amplification and mRNA overexpression have been found in mammary carcinomas. High concentrations of EGFr in breast cancer have been related to a poorer prognosis suggesting that it might be an index of greater aggressiveness of the disease. One study (146) concluded that the only significant marker for disease-free survival was the absence of the EGFr, but it did not correlate significantly with overall survival. Finally, EGFr status was a valid criterion for dividing ER-negative tumors into those with good and poor prognosis (i.e., ER-negative/EGFr-positive patients had a significantly worse prognosis that did patient's negative for both receptors).

Amplification of the c-erb-B2 oncogene had a positive relation to the number of axillary-node metastases as a prognostic marker. It was a stronger prognostic marker than PR status, tumor size, or ER status for both disease-free interval and survival (147). The presence of any of these oncogene products provides for a poor prognosis, perhaps because each is involved with the stimulation of cell growth. Further, permanent activation of TKA in the case of c-erb-B2 may indicate that it is involved with progression of the tumor and development of metastases. Measurement of gene amplification is relatively difficult and not suitable for the routine pathology laboratory. Preliminary data suggest that other oncogenes/oncoproteins can be used to provide prognostic information in breast cancer. Clair (148) has shown that breast cancers with high concentrations of the *ras* protein have a greater likelihood of developing recurrent disease than do tumors containing low concentrations of this protein. Similarly, amplification of both the *int*-2 and c-myc oncogenes have been associated with aggressive breast cancer. With gene amplification, it is noteworthy that while the capacity of the receptors increased, the affinity did not. This leads to the conclusion that it is simply an increase in the number of EGFrs and not a change in the actual receptor. However, overexpression of this gene did not always result in increased EGFr. Other factors that might come into play here such as increased transcription of the gene (over other genes); reduced degradation of the mRNA allowing more “readings” by the ribosome; and reduced receptor degradation, which might keep the cell “turned on” longer.

Endometrial, Cervical, and Vulvar Cancers

The incidence of endometrial adenocarcinoma is 11 to 12/100,000/year, and rising in developed countries (American Cancer Society, 1986). The majority of endometrial cancers are adenocarcinomas (80%) with the remaining 20% composed of

adenoacanthoma, adenosquamous carcinomas, papillary and clear-cell carcinomas (149). 5-year survival of patients with treatment initiated during stage 1 disease varies between 70% and 94%, depending upon the histological grade and receptor content of the tumor. 5-year survival drops to 55% to 88% with stage II disease, whereas the rates drop to 30% to 32% and 0 to 10% with stage III and stage IV disease, respectively (150,151,152,153,154 and 155). Cervical carcinoma occurs at the rate of about 10/100,000/year (156). Mass screening programs and early detection have been successful in reducing the incidence to about 3.6/100,000/year. Cervical carcinoma is primarily squamous, although cervical adenocarcinoma is increasing in incidence and currently makes up about a 10% of invasive cervical cancers (157). The 5-year survival rate for stage I disease is 62.8% to 100%, however if treatment is not instituted prior to metastases, survival rate drops below 10% (158). Vulvar carcinoma has an incidence of only 1 to 2/100,000/year, and a 5-year survival rate of about 90% (157, 158).

Growth, development, and maintenance of the female genital tract is controlled by sex steroids, estrogen, and progesterone receptors that are found in the highly differentiated target tissues of the normal female reproductive tract and growth factor receptors, e.g., EGFr. In the endometrium, estradiol stimulates mitosis and increases progesterone receptor concentration. The progesterone/receptor complex induces differentiation of the secretory function of endometrial cells. Normal endometrium contains higher concentrations of steroid receptors than those found in endometrial carcinoma and the malignant endometrium has a low PR to ER ratio (159). As dedifferentiation occurs in a carcinoma, it appears to be accompanied by a concomitant loss of receptors and a reduction in receptor concentration (118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162 and 163). Receptors of both types are found in highly differentiated tumors and poorly differentiated tumors while primary tumors with metastases usually lack them (160). Taken together, these data indicate that endocrine control of endometrial cancer may be lost or reduced because of modifications that occur during the malignant process. In tumors that lack or have low concentrations of estrogen and progesterone receptors, the rate of adnexal spread of the tumor is increased and the recurrence rate of stage I cancer is higher (149). Progesterone receptor, and not the estrogen receptor, appears to correlate significantly with tumor grade and histology with the greatest decline in receptor positivity occurring between grade 2 and grade 3 tumors.

The predictive value of receptors in determining the course of endometrial carcinoma currently is a topic of controversy. Differing cut-off values have been reported by different investigators and these numbers do not reflect those obtained in breast carcinoma. This is partly due to the absence of standardized assays and reference materials. Apart from that, the presence of receptors is a predictive factor.

Creasman et al. (164) first reported that estrogen and progesterone receptor-positive endometrial tumors were less aggressive than estrogen and progesterone receptor-negative tumors and longer survival time is associated with high levels of these receptors. Ensuing work (165, 166) indicated that the concentration of estrogen and progesterone receptors was prognostic. Palmer et al. (167), using cutoff values of 70 fmole ER/mg protein and 30 fmole PR/mg protein showed that stage, age, and each of the steroid receptors, were associated independently with survival. Recently, a number of reports confirming the utility of ER and PR levels on predicting survival time have been published (168,169,170 and 171). In contrast, Erlich described overall survival as superior for patients with progesterone receptor-positive tumors ($p < 0.001$) but could not find any statistical relationship between estradiol receptor status and survival. In a subsequent study of 309 malignant endometrial tumors, Kleine (172) likewise observed that estrogen receptor had no significant prognostic relevance over that obtained by PR determination. In an examination of survival in early endometrial cancer, Sutton et al. (171) determined, using a Cox proportional hazards model, that tumor grade, peritoneal cytological results, PR status, and age were the most closely associated with disease-free survival. In 1989 and 1990 reviews, Kauppila (159) and Podratz (173) both concluded from an examination of a number of studies of endometrial cancer that PR is the best receptor measurement for the prognosis of survival and response to progesterone therapy. In summary, a negative ER or PR both give similar information in identifying nonresponders and a positive PR provides better accuracy in determining patients that will respond to progesterone therapy. However, as many as 70% of patients with endometrial cancer that have receptor levels above the cutoff limits normally used to identify candidates for endocrine therapy fail to respond to progesterone therapy. Lack of response in PR positive tumors to endocrine therapy suggests the presence of a defect(s) in the cellular machinery distal to the steroid binding site, cellular heterogeneity within the tumor (i.e., some cells possess receptors and others do not), cellular control mechanisms that are hormone independent, or increased hormonal metabolism (174,175 and 176).

EGF is excreted in high concentration in the urine in a biologically active form. The presence of the EGF receptors in close contact with this milieu provides the opportunity for the EGF to incubate with normal, premalignant, and malignant transitional epithelium of the bladder for long periods of time. With such favorable conditions, it is not too surprising that the density of receptor expression reflects the degree of malignancy in the bladder. It also is worthwhile noting that the expression of EGFrs on the superficial cells often precedes the standard histological evidence for a preneoplastic state.

Attempts have been made to control the progress of malignancy through the use of toxins or radioisotopes attached to ligands or antireceptor antibodies. The idea is for the compound to seek out the receptor and deal a lethal dose of toxin or radiation. Other approaches are the use of unconjugated antireceptor antibody designed to block the receptor-binding site, and anti-EGF antibody intended to inactivate EGF. None of these has proven successful. Explanations for the lack of success focus on the object of the therapy. Thus, those aimed at EGF have to deal with a large, continuous, and fresh supply of EGF (excreted in the urine). It would be very difficult to nullify this nonending influx of EGF with outside agents. The methods aimed at the EGFr need to penetrate deep into the tumor mass to get all of the receptor sites.

It appears that the increasing density of EGFr sites on the cervical and vulvar squamous-cell carcinomas can be related to the

biological aggressiveness of these malignancies. In particular, those neoplasm's with levels of EGFr >100 fM/mg of protein show a poor prognosis while those with levels <100 fM/mg of protein had more favorable 5-year survival (49).

Prostatic Cancer

In 1983 Barkovitz (177) made the observation about carcinoma of the prostate (PC) that “an incidence of 50% is reported in 80-year old men — but it is extremely rare before age 40.” Localized PC can be cured by surgical removal or radiation therapy. When the general health and age of the patient or progress of the disease (metastases) prevents this course of treatment, hormone therapy is the next option. Treatment of prostate carcinoma (PC) through hormone therapy has been successful with up to 80% favorable response. Therapy commonly consists of androgen ablation through orchiectomy or treatment with estrogen, diethylstilbestrol (DES) or luteinizing hormone-releasing hormone (LH-RH) or its analogs. Orchiectomy often is combined with the administration of DES. Cyproterone acetate with DES has been used successfully (178). However, the tumor eventually becomes resistant to the hormone therapy and again progresses.

Analogous to the work in breast cancer, where receptors have been moderately successful in predicting the course of the disease, studies have been carried out to determine the utility of receptor concentrations and affinities to aid in the prognosis of PC. While receptors for androgen receptors (AR), ER, PR, EGF, LH-RH, and PRL have been quantified in PC tissue, their prognostic value has not been obvious. In fact, the presence or absence of receptors is not even a factor when PC is initially diagnosed because of the high success rate with hormone therapy.

The most studied prostate receptor is the intracellular AR. Genetically, the AR is a single unique gene product. The gene itself can be divided into eight coding exons and corresponds closely in organization to the human estrogen gene. It is X-linked and defects have been observed in males who carry only a single copy of the gene. The receptor is a 110 kd single protein chain, with three distinct regions; the ligand-binding region, the nuclear interaction region (containing two zinc fingers) and a third region that modulates the protein synthesis process. The hormone and DNA binding domains are analogous to the respective domains in other steroid receptors and much of these regions are highly conserved between the steroid receptors. In contrast to other steroid receptors, AR is present in considerably lower tissue concentrations, extremely labile, and susceptible to proteolytic breakdown. AR is capable of binding a large number of agonist molecules, but activation of the receptor appears to depend upon binding to testosterone or dihydrotestosterone (DHT). The affinity of the receptor for DHT is two-fold greater than for the unsaturated testosterone and the dissociation rate from AR for the latter is significantly higher than for DHT. It has been speculated the testosterone is functionally a prohormone for DHT and estrogens such as 17β-estradiol.

The AR has been found in both the epithelial and stromal cells. As with the other steroid receptors there has been an evolution of models depicting the location of the AR. In the nucleus, the AR is present in a salt-extractable and salt-resistant form. The presence of molybdate or dithiothreitol solubilized all of the nuclear AR. It has been speculated that the molybdate interacts with the DNA binding domain of the AR. This leads to the conclusion that the salt-resistant nuclear AR is that which is associated with DNA. It generally has been assumed that the AR also is present in the cytosol and Mulder (179) describes the presence of AR aggregates in the cytosol with sedimentation coefficients of 8S and 4S. The relative proportion of these two species is dependent on the presence of molybdate or salt where the 4S is favored. Immunocytochemical studies, however, find little evidence of AR in the cytosol and locate the receptor almost exclusively in the nuclear region (180). It should be kept in mind that this technique depends upon the antibody recognition and aggregates of the AR may hide or alter the antigenic site(s).

Androgens promote cellular growth and EGF regulates the growth of both normal and tumor epithelial cells of the prostate. There is evidence that ARs also act as secondary promoters through the increased expression of EGF receptors (also found in prostate tissue) on the cellular surface membrane because the binding capacity of EGF receptors increases in the presence of androgens. Both of these receptors appear to act through enhancing the cells sensitivity to growth factors. It is not clear whether the mechanisms follow autocrine or paracrine control. In addition, it should be noted that the presence of AR alone is not sufficient for stimulating growth. The other condition is the ability of the tissue to convert testosterone to DHT, which is the active species.

Other receptors found in prostate tissue are estrogen, progesterone, growth hormone, hCG, and prolactin (PRL) receptors. The expression of EGF receptors in benign prostatic hyperplasia (BPH) is about twice that found in prostate carcinoma (PC) (181), while the reverse is found for AR. In PC tissue, an increase in receptor content appears to be associated with a corresponding decrease in AR and loss of differentiation (182). Both EGF and androgen receptors are more commonly found in highly differentiated then in poorly differentiated tumors and generally not found after prolonged estrogen therapy. Patients treated with DES showed up to 50% decreases in hCG receptors after 7 days of treatment, and after six months of treatment with DES, levels had dropped to 25% of the original value (183). High- and low-affinity–binding sites for estrogens are found in normal prostate tissues, but only the only low affinity site is present in tissue from BPH or PC (184).

Kadar (185) reports that rat PC tissue contains receptors for LH-RH, somatostatin and PRL. Treatment with the LH-RH agonist, [D-trp6] LH-RH, or the somatostatin agonist, RC160, resulted in shrinkage of the tumor mass with a corresponding decrease in the above receptor expressions. The decrease for PRL receptors was the greatest (about 50%). The authors speculate that tumor growth is inhibited by down regulation of these receptors through binding to these agonists.

The current direction of study is toward finding prognostic indicators after therapy has been initiated. Benson, et al. (186) observed that 85% of patients with nuclear androgen binding of ≥26 fM/mg did not show progression of the disease after 3 years, compared to only 40% of patients with androgen nuclear binding of <26 fM/mg. However, in grade 4 lesions, androgen binding is not a useful indicator and the PC is generally resistant to hormone therapy. They concluded that, excluding grade 4, measuring the amount of androgen bound by the nuclear AR

was a gross predictor of response to endocrine therapy. But these numbers are not sufficient to make a determination to withhold therapy. In fact, the presence or complete absence of ARs does not predict the response, as 80% of AR-positive and 25% of AR-negative patients respond favorably to hormone therapy.

After relapses from hormone therapy, further evaluation of AR does not appear to provide any useful guidance. Here it is of interest that the concentration of DHT present in the tumor may be a prognostic marker. Connelly (187) reports that “patients whose tumors contained more than 2 ng DHT/g of tissue were more likely to experience remission after endocrine treatment than those whose tumors contained lower concentrations of DHT.”

There is no evident correlation between AR concentration and disease stage. The latter is best evaluated through determination of tumor size. One contributing factor to this observation is the large amount of heterogeneity of AR's in PC tissue. This may explain why response to hormone therapy and PC tissue AR concentrations are not necessarily parallel (188). The problem of heterogeneity also is dependent upon the method of obtaining samples. For example, while needle biopsy is a simple procedure, the sample probably is not representative of the whole tissue. Samples obtained by electro-resection render PC tissue unsuitable because of the heat generated.

There are significant errors in current assays, which often invalidate correlations between the amount of AR and therapy response. In addition to the problem of heterogeneity (188), other significant problems include contamination with a plasma protein; the presence of significant amounts of sex hormone binding globulin (SHBG), which competes for DHT; and the presence of large amounts of endogenous DHT. When there are large amounts of endogenous DHT, many of the receptor sites become bound *in vivo*. This requires doing an exchange assay, which is slow at low temperatures, while at elevated temperatures the receptor is unstable. The use of DHT as the label is limited because of potential metabolism of DHT during the assay procedure. More commonly, the synthetic steroid methyltrienolone (R-188) or 3H-mibolerone is utilized as the ligand. They do not bind to SHBG and are not subject to metabolism during the assay procedure. The disadvantage of R-188 is that it binds to both AR and PR. This is eliminated by swamping the PR with triamcinolone acetate.

Unfortunately, in spite of the high level of response to endocrine treatment of prostatic carcinoma, at some point all PCs progress to an androgen-resistant state at which time the disease continues to advance. Thus, the obvious need is to understand why these tumors become androgen resistant. Perhaps the receptor becomes inactive through genetic processes that result in truncation or production of chimeric forms of the receptor.

Hormone-Resistant States

Hormone-resistance disorders may be related to a defect in the hormone itself, although a receptor or postreceptor defect that prevents normal functional activity of the receptor usually causes them. Normal function of a receptor involves a number of intricate prereceptor, receptor and postreceptor events that are very clearly dependent upon the structure of each of the participating molecules. The change of a single nucleotide in the genetic DNA sequence can cause an amino acid substitution and a resultant structural change that, if located in a strategic position in the molecule, may reduce or completely inhibit receptor activity. The list of identified endocrine disorders that result from inherited receptor defects continues to increase. Amino acid, gene or chromosomal deletions are responsible for many of the known receptor defects, although point mutations are much more common and the amount of genetic polymorphism present in a normal population probably accounts for at least some of the normal physiological diversity in hormone action. There are several examples of heterogeneity in the structure and function of specific receptors as a result of point mutations. This is particularly true for the LDL receptor in which 10 different point mutational changes have been described (189). Nearly all of the genes that code for hormones, growth factors and their receptors will probably show the same degree of genetic heterogeneity as has been described for the LDL receptor. Although several receptor-deficient diseases have been described, this still is a new area of research in which the analytical procedures are still largely experimental. Moreover, they are performed by a relatively small number of research laboratories, and the clinical diagnosis, as a general rule, does not utilize laboratory analysis of receptor status. Difficulties that may be encountered in the analyses of receptor characteristics in hormone resistant disease are the procurement of appropriate specimens upon which to perform the analysis, nonuniformity of analytical procedures, the lack of standardization materials that permit interlaboratory comparisons, and the relatively low number of patients involved. However, because most of these diseases reflect genetic disorders, the mutational changes responsible for receptor or postreceptor defects will very likely be determined in the future in clinical laboratories equipped to amplify the appropriate DNA sequences by the Polymerase Chain Reaction (PCR) and determining polymorphisms using molecular biology procedures.

Insulin Resistance

Insulin is the primary hormone involved with glucose, lipid, and protein metabolism. Its actions are varied, ranging from: increased glucose transport, increased glycogen synthesis, and antilipolysis (initiated within minutes of insulin binding to the receptor), to the initiation of gene transcription and protein synthesis that occur some hours after the hormone has bound to the receptor. In diabetes mellitus, carbohydrate metabolism is compromised, either through insulin deficiency, or because of target cell refractoriness to insulin. The former condition is referred to as type I diabetes, or insulin-dependent diabetes mellitus (IDDM) and the latter as type II diabetes, or noninsulin dependent diabetes mellitus (NIDDM). Thus, diabetes, classified into two broad categories, is not a single disease entity, rather it appears to be a constellation of conditions that may result from any of a number of environmental insults and/or genetic defects. For example, in type I diabetes most patients suffer from the lack of insulin because of autoimmune destruction of the pancreatic ß cells, although, mutational changes in the insulin gene can produce a nonfunctional hormone that is incapable of activating its receptor, also causing a form of type I diabetes. Clinical signs

of type I diabetes are permanently elevated blood glucose levels, weight loss, polydipsia, polyuria, and glycosuria, conditions that can lead to ketoacidosis. Successful treatment of type I diabetes is based upon the replacement of normal physiological levels of insulin.

Type II diabetes is the most common form of diabetes (80% to 90% of all diabetics) with a very complex etiology. This condition characterized by elevated insulin levels normally occurs later in life. NIDDM was referred to as mature-onset diabetes because it tends to occur later in life. Because this form of diabetes tends to be a heritable disease, a family history of NIDDM is much more frequent in new cases of NIDDM, and is a greater risk factor than is true for IDDM. No single susceptibility gene for type II diabetes has been identified, nor are mutational changes in the insulin receptor alone likely to account for the insulin resistance observed with most NIDDM because these mutations are rather uncommon in the general population. However, genetic anomalies in this receptor do occur. An “online” source of information on this and other genetic information is “The Online Mendelian Inheritance in Man” (132). This Internet source currently lists 35 mutational changes in the insulin receptor that have been described in the literature. Regardless, it is certain that they (mutations of the insulin receptor) do not occur at the prevalence rate of type II diabetes and likely account for a small percentage of type II diabetes. A more likely target of mutational changes responsible for NIDDM comprises the multiplicity of individual proteins involved in the extremely elaborate insulin receptor signaling cascades. This myriad of interacting proteins quite obviously presents numerous opportunities for genetically derived changes in protein structure, changes which have the potential to perturb insulin signaling and account for a wide variety of the different insulin sensitivities observed in type II diabetes. In addition to the genetic component of type II diabetes, this condition is further complicated by environmental factors such as obesity and physical inactivity that are known risk factors for patients with a family history of NIDDM.

Mutational changes that have been reported in the insulin receptor include single point, and reading frame shift mutations in the insulin receptor gene, resulting in a variety of receptor defects. These defects include: (i) inhibition of tyrosine kinase activity, (ii) inhibition of receptor transport to the plasma membrane, (iii) decreased dimerization of the α and ß subunits, (iv) decreased binding affinity for insulin, (v) impaired posttranslational processing, (vi) inhibition of proreceptor proteolysis, (vii) facilitated constitutive activation of the receptor, and (viii) inappropriate placement of stop codons. In general, mutations in the α chains reduce or prevent insulin binding and changes in the transmembrane or cytosolic region of the receptor may inhibit receptor autophosphorylation, and subsequent tyrosine kinase activity. The post-receptor signaling cascades also may be disrupted by mutational changes that prevent or reduce expression of the insulin receptor gene through the slowing or total inhibition of gene transcription. Normal insulin signaling may be impaired by a reduction in receptor numbers at the plasma membrane or decreased binding affinity of the receptor. Any of these structure-function changes can cause diabetes, even if the patient is heterozygous for the genetic change because of the reduced number of functional molecules. This is particularly true later in life when metabolism begins to slow down and the fat content of the body increases, increasing the level of stress on glucose metabolism.

Androgen Resistance

The spectrum of androgen-resistant syndromes ranges from complete androgen insensitivity syndrome (CAIS), in which genetic males present as phenotypic females, through the incomplete, or partial form of the syndrome (PAIS) in which the predominant phenotypic form is a male but characterized by any of several examples of incomplete virilization. Other conditions such as ambiguous genitalia, primary amenorrhea, prostate cancer, spinal bulbar muscular atrophy, male infertility, rheumatoid arthritis, breast cancer, hirsutism, baldness, and acne also are found in individuals with CAIS. The spectrum of partial virilization ranges from hypospadias in severely affected individuals, to gynecomastia and azoospermia in mildly affected patients (190). Some phenotypically normal men with oligospermia have been found to suffer incomplete testicular feminization with qualitative deficiencies in the androgen receptor (191). Severe cases of this syndrome usually are associated with the complete absence of the androgen receptor (AR), whereas the incomplete forms of the disorder are associated with qualitative changes in the receptor, perhaps responding to point mutations in the androgen receptor gene.

A single copy gene in the X chromosome encodes AR and defective forms of the receptor are attributable to one of several mutational changes. In most cases, a single point mutation occurs, resulting in frameshifts or nucleotide substitutions that are responsible for inappropriately sited stop codons (192) or amino-acid substitutions that lead to structural changes in the ligand binding (193, 194, 195, 196, 197, 198, 199) domains of the gene. Amino-acid substitutions, resulting from single point mutations, account for the majority of AR defects that lead to CAIS, or to PAIS. The steroid-binding domain appears to undergo more mutational changes than other domains of the AR protein. For example, one mutation, N758T, in a linker region between the fifth alpha helix (H5) and the first beta strand (S1) displayed normal binding affinity to DHT but abnormal dissociation kinetics. Receptor transactivation capacity also was reduced about 50% (198). In another mutation (TYR 571 CYS) that occurred in the DNA binding domain, gene transcription was disrupted, but 5 α-dihydrotestosterone-receptor-binding characteristics were not affected. These data confirm that a single mutational substitution in the DNA binding domain of AR was related to the pathogenesis of CAIS (199). Single-point mutational changes in AR have been described that have caused frameshifts with resultant stop codons, abbreviated AR structural protein, and abnormal function in the transactivation of the receptor (200). In another example, a point mutation in the donor splice-site of the second intron caused the entire second exon to be lost, exons 1 and 3 directly linked to each other, and mRNA exhibited a premature termination between codons 598 and 599. This change produced a truncated AR protein with no *in vivo* function and with abnormally low levels of transcript and protein (201). In addition to the single-point

mutations described above, deletions of two or more base pairs have been described in CAIS cases in which androgen binding is absent.

A change in the number of CAG trinucleotide repeats in the AR ligand-binding domain of the gene also may affect AR activity (202). The CAG repeat in exon 1 is important for the transactivation function of AR, and consequently for many androgen-dependent processes (203). [Further, the transactivation domain of AR is of great interest because of its causal relationship with a fatal neuromuscular disease, spinal bulbar muscular atrophy (Kennedy's syndrome)]. In infertile males, 20% had an increase in the length of a polymorphic trinucleotide (CAG) repeat segment. The increased risk of male infertility associated with long CAG lengths is accompanied with reduced risk of prostate cancer. Conversely, short polyglutamine tracts are (204). Other mutations of the AR gene occur commonly in prostate cancers and appear to be significant for prognosis of the disease (205). Although the change in the number of CAG repeat segments is the most commonly reported mutational change in the polymorphic-trinucleotide-repeat segments, single point mutations do occur. The insertion of a single adenine (or equivalently, a GC-dinucleotide deletion) has been found in a large kindred of phenotypic females with CAIS. This change in exon 1 of the AR gene produced a frameshift at amino-acid 60 and premature termination of the receptor downstream of the mutation (206).

Two functionally different ARs (Kd1: 5.58 nM = mutant, Kd2: 0.06 nM = wild type) were described in one AIS individual, suggesting the presence of somatic mosaicism in the patient. In this patient, androgen action was expressed in somatic cells through the wild type AR, demonstrating that the AR can modulate *in vivo* androgen action and emphasizes the importance of evaluation of somatic mosaicism (207). Thus, somatic mosaicism may be clinically relevant in AIS. The possibility of functionally relevant expression of the wild-type AR needs to be considered in all mosaic individuals, and treatment should be adjusted accordingly (208).

In summary, androgen insensitivity syndrome (AIS) exhibits a wide spectrum of phenotypic and physiological manifestations of the inability to respond to androgenic steroids. These range from complete androgen insensitivity syndrome (CAIS), in which genetic males present as phenotypic females, through the incomplete, or partial form of the syndrome (PAIS) in which the predominant phenotypic form is a male but characterized by any of several examples of incomplete virilization. The inability to respond to the presence of androgenic steroids is because of mutational changes in the androgen receptor that render it inoperative, either through the inability to bind the steroid, to translocate, or to bind to DNA an effect transcription. The advent of molecular cDNA cloning procedures in recent years has vastly improved our understanding of the molecular structure and mutational changes that may occur in the androgen receptor and how it plays in the syndrome of androgen insensitivity. It has been suggested that because of the wide phenotypic heterogeneity and variability in AR gene mutations, proficient bioassays of AR function would improve diagnostic efficiency and comprehensive evaluation of male under-masculinization (209).

The extent to which biomedical science has enhanced our understanding of the molecular mechanics of AIS is impressive and in families with well-characterized AR mutations, carriers can be identified and prenatal testing offered. An unexpected finding is that an AR mutation also causes X-linked spinobulbar muscular atrophy (210).

APPENDIX

ABBREVIATIONS

NONSTANDARD ABBREVIATIONS

AR	Androgens receptors
BPH	Benign prostatic hyperplasia
DCC	Dextran-coated charcoal
DES	Diethylstilbestrol
DHT	Dihydrotestosterone
EGF	Epidermal growth factor
EGFr	Receptor for epidermal growth factor (EGFr)
EIA	Enzyme immunoassay
ER	Estrogen receptor
FH	Familial hypercholesterolemia
GF	Growth factors
GFr	Growth factor receptors
GnRH	Gonadotropin releasing hormone
GR	Glucocorticoid
HAP	Hydroxyapatite
hCG	Human chorionic gonadotropin
HPCF	High-performance chromatofocusing
HPHIC	High-performance hydrophobic interaction chromatography
HPIEC	High-performance ion exchange chromatography
HPLC	High-performance liquid chromatography
HPSEC	High-performance size exclusion chromatography
HRE	Hormone responsive element
HSP	Heat shock protein
ICA	Immuno cytochemical assays
IGF-I	Insulinlike growth factor I
IGF-II	Insulinlike growth factor II
IRMA	Immunoradiometric Assays
LATS	Long acting thyroid stimulator
LDL	Low-density lipoprotein
LH-RH	Luteinizing hormone-releasing hormone
moAb	Monoclonal antibody
NSB	Nonspecific binding
OPD	O-phenylenediamine
PC	Prostate carcinoma
PCR	Polymerase chain reaction
PDGF	Platelet derived growth factor
PEG	Polyethylene glycol
PR	Progesterone receptor
PRF	Peptide regulatory factors
PRL	Prolactin
RPLC	Reverse phase chromatography
SDG	Sucrose density gradient
SHBG	Sex hormone binding globulin
SPA	Sucrose pad nuclear exchange assay
SS	Somatostatin
TFS	Testicular feminization syndrome
TGFα	and TGFβ Transforming growth factors
TKA	Tyrosine kinase activity

REFERENCES

1. Bayliss WM, Starling EH. The mechanism of pancreatic secretion. *J Physiol* (London) 1902;28:325.

2. Nalbandov AV. *Reproductive Physiology: Comparative Reproduction Physiology of Domestic Animals, Laboratory Animals, and Man*. San Francisco: WH Freeman, 1958:195–197.

3. Green AR. Peptide regulatory factors: multifunctional mediators of cellular growth and differentiation. *Lancet* 1989;705.

4. Foekens JA, Peters HA, Portengen H, et al. Cell biological prognostic factors in breast cancer: a review. *J Clin Immunoassay* 1991;14:184–195.

5. Jensen EV, Jacobson HI. Fate of steroid estrogens in target tissues. In: Pincus G, Vollmer EP, eds. *Biological Activities of Steroids in Relation to Cancer*. New York: Academic Press, 1960:161–178.

6. Evans RM. The steroid and thyroid hormone receptor superfamily. *Science* 1988;240:889–895.

7. Greene GL, Sobel NB, King WJ, et al. Immunochemical studies of estrogen receptors. *J Steroid Biochem* 1984;20:51.

8. Green S, Gronemeyer H, Chambon P. Structure and function of steroid hormone receptors. In: Sluyser M, ed. *Growth Factors and Oncogenes in Breast Cancer* England: Ellis Horwood, 1987:7–28.

9. Miesfeld, RL. The structure and function of steroid receptor proteins. *Crit Rev Biochem Mol biol* 1989;24:101–117.

10. Tilley WD, Marcelli M, Wilson JD, et al. Characterization and expression of a cDNA encoding the human androgen receptor. *Proc Natl Acad Sci USA* 1989;86:327–331.

11. Arriza JL, Weinberger C, Cerelli G, et al. Cloning of human mineralocorticoid receptor complementary DAN: structural and functional kinship with the glucocorticoid receptor. *Science* 1987;237:268–275.

12. Greene GL, Gilna P, Waterfield M, et al. Sequence and expression of human estrogen receptor complementary DAN. *Science* 1986;231:1150–1154.

13. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM 1995 The nuclear receptor superfamily: the second decade. *Cell* 83:835–839.

14. McKenna NJ, Lanz RB, O'Malley BW. Nuclear receptor coregulators: cellular and molecular biology. *Endocrine Reviews* 1999; 20:31–344.

15. Wilson TE, Fahrner TJ, Johnson M, et al. Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* 1991;252:1296–1300.

16. Harding HP, Lazar MA. The orphan receptor Rev-Erba activates transcription via a novel response element. *Mol Cell Biol* 1993; 13:3113–3121.

17. Moraitis AN, Giguere V. Transition from monomeric to homodimeric DNA binding by nuclear receptors: identification of RevErbAa determinants required for RORa homodimer complex formation. *Molecular Endocrinology* 1999;13:431–439.

18. Giguere V. Orphan nuclear receptors: from gene to function. *Endocrine Reviews* 1999;20:689–725.

19. Joab I, Radanyi C, Renoir M, et al. Common non-hormone binding component in non-transformed chick oviduct receptors of four steroid hormones. *Nature* 1984;308:850–853.

20. Catelli MG, Binart N, Jung-Testas I, et al. The common 90-Kd protein component of non-transformed 8S' steroid receptors is a heat-shock protein. *EMBO J* 1985;4:3131–3135.

21. Milgrom E, Atger M, Baulieu EE. Acidophilic activation of steroid hormone receptors. *Biochemistry* 1973;12:5198–5205.

22. Baulieu EE. Antibodies against highly purified B-subunit of the chick oviduct progesterone receptor. *Biochem Biophys Res Commun* 1984;119:433–439.

23. Walter P, Green S, Greene G, et al. Cloning of the human estrogen receptor cDNA. *Proc Natl Acad Sci USA* 1985;82:7889–7893.

24. Sabbah M, Redeuilh G, Secco C, et al. The binding activity of estrogen receptor to DNA and heat shock protein (Mr 90,000) is dependent on receptor-bound metal. *J Biol Chem* 1987;262:8631–8635.

25. Orti E, Bodwell JE, Munck A. Phosphorylation of steroid hormone receptors. *Endocrine Reviews* 1992;13:105–128.

26. Ruh TS, Ruh MF, Singh RK. Nuclear acceptor sites: interaction with estrogen-versus antiestrogen-receptor complexes. In: Moudgil VK, ed. *Steroid Receptors in Health and Disease*. New York: Plenum Press, 1988:233–250.

27. Carson-Jurica MA, Schrader WT, O'Malley BW. Steroid receptor family: structure and functions. *Endocrine Reviews* 1990;11:201–220.

28. Baker AR, McDonnell DP, Hughes M, et al. Cloning and expression of a full-length cDNA encoding human vitamin D receptor. *Proc Natl Acad Sci USA* 1988;85:3294–3298.

29. Giguere V, Ong ES, Segui P. Identification of a receptor for the morphogen retinoic acid. *Nature* 1987;330:624–629.

30. Wang LH, Tsai SY, Cook RG, et al. COUP transcription factor is a member of the steroid receptor superfamily. *Nature* 1989; 340:163–166.

31. Lazar MA, Hodin RA, Darling DS, et al. Identification of a rat c-erbAalpha-related protein which binds DNA but does not bind thyroid hormone. *Mol Endocrinol* 1988;2:893–901.

32. Mitsuhashi NT, Tennyson GE, Nikodem VM. Alternative splicing generates messages encoding rat c-erbAalpha proteins that do not bind thyroid hormone. *Proc Natl Acad Sci USA* 1988;85:5804–5808.

33. Virkamaki A, Ueki K, Kahn CR. Protein-protein interaction in insulin signaling and the molecular mechanisms of insulin resistance. *J Clin Invest* 1999;103:931–943.

34. Gerhard M. *Biochemical Pathways, An Atlas of Biochemistry and Molecular Biology*, Spedtrum Akademischer Verlag (English Language Edition by John Wiley & Sons, Inc.), 1999.

35. Kiolakowski Jr, LF. GCRDb: a G-Protein-coupled receptor database. *Receptors Channels* 1994;2:1–7.

36. Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine Reviews* 1999;21:90–113.

37. Meisenberg G, Simmons WH. *Principles of Medical Biochemistry*. St. Louis, Missouri: Mosby, Inc., 1998.

38. Smith BR, Hall R. Thyroid-stimulating immunoglobulins in Graves' disease. *Lancet* 1974;2:427–431.

39. Zakarija M, McKenzie JM. Autoantibodies to the TSH receptor. *Exp Clin Endocrinol* 1991;97:165–169.

40. Zakarija M, McKenzie JM. Effects of thyrotropin and thyroid hormones in vivo on thyroid responsiveness to thyrotropin in vitro. *Endocr Res Commun* 1977;4:343–355.

41. Grob D. Myasthenia gravis: pathophysiology and management, retrospect and prospect. *Ann NY Acad Sci* 1981;377:xiii-xvi.

42. Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science* 1986;232:34–47.

43. Beatson GT. On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment, with illustrative cases. *Lancet* 1896;2:104–107.

44. Glascock RF, Hoekstra WG. Selective accumulation of tritium labeled hexestrol by the reproductive organs of immature female goats and sheep. *Biochem J* 1959;72:673–682.

45. Folca PJ, Glascock RF, Irving WJ. Studies with tritium-labeled hexestrol in advanced breast cancer. *Lancet* 1961;2:796–798.

46. Horwitz KB, McGuire WL. Specific progesterone receptors in human breast cancer. *Steroids* 1975;25:497–505.

47. Pertschuk LP. *Immunocytochemistry for Steroid Receptors*. Boca Raton, FL: CRC Press, 1990.

48. Dazzi H, Hasleton PS, Thatcher N, et al. Expression of epidermal growth factor receptor (EGF-R) in non-small cell lung cancer. Use of archival tissue and correlation of EGF-R with histology, tumor size, node status and survival. *Br J Cancer* 1989;59:746–749.

49. Pfeiffer D, Stellwag B, Pfeiffer A, et al. Clinical implications of EGFr in squamous cell carcinoma of the uterine cervix. *Gynecol Oncol* 1989;33:146–150.

50. Hanauske AR, Osborne CK, Chamness GC, et al. Alteration of EGF-receptor binding in human breast cancer cells by antineoplastic agents. *Eur J Cancer Clin Oncol* 1987;23:545–551.

51. Grimaux E, Romain S, Remvikos Y, et al. Prognostic value of epidermal growth factor receptor in node-positive breast cancer. *Breast Cancer Res Treat* 1989;14:77–90.

52. Fernandex-Pol JA. Modulation of EGF receptor protooncogene expression by growth factors and hormones in human breast carcinoma cells. *Crit Rev Oncog* 1991;2:173–185.

53. Fisher DA, Lakshmanan J. Metabolism and effects of epidermal and related growth factors in mammals. *Endo Rev* 1990;11:418–42.

54. Korc M, Haussler CA, Trookman NS. Divergent effects of epidermal growth factor and transforming growth factors on a human endometrial carcinoma cell line. *Cancer Res* 1987;47:4909–4914.

55. Miyachi Y, Terazono T, Nagao N, et al. Epidermal growth factor (EGF) receptors in human chorionic gonadotropin-producing tumor: transplantation in nude mice and the effect of EGF on tumor growth. *J Clin Endocrinol Metab* 1990;71:329–334.

56. Bauknecht T, Kohler M, Janz I. The occurrence of epidermal growth factor receptors and the characterization of EGF-like factors in human ovarian, endometrial, cervical and breast cancer. EGF receptors and factors in gynecological carcinomas. *J Cancer Res Clin Oncol* 1989;115:193–199.

57. Messing FM. Clinical implications of the expression of EGFr in human transitional cell carcinoma. *Cancer Res* 1990;50:2530–2537.

58. Berchuck A, Soisson AP, Olt GJ, et al. Epidermal growth factor receptor expression in normal and malignant endometrium. *Am J Obstet Gynecol* 1989;161:1247–1252.

59. Reynolds RK, Talavera F, Roberts JA, et al. Characterization of epidermal growth factor receptor in normal and neoplastic human endometrium. *Cancer* 1990;66:1967–1974.

60. Reynolds RK, Talavera F, Roberts JA. Regulation of epidermal growth factor and insulin-like growth factor I receptors by estradiol and progesterone in normal and neoplastic endometrial cell cultures. *Gynecol Oncol* 1990;38:396–406.

61. Llorens MA, Bermejo MJ, Salcedo MC, et al. Epidermal growth factor receptors in human breast and endometrial carcinomas. *J Steroid Biochem* 1989;34:505–509.

62. Dittadi R, Gion M, Brazzale A, et al. Radioligand binding assay of epidermal growth factor receptor (EGFr): causes of variability and standardization of the assay. *Clin Chem* 1990;36:849.

63. Goldschmitdt-Clermont PJ, Kim JW, Machesky LM, et al. Regulation of phospholipase C-g1 by profilin and tyrosine phosphorylation. *Science* 1991;251:1231–1233.

64. Neskovic-Konstantinovic Z, Nikolic-Vukosavljevic D, Brankovic-Magic M, et al. Expression of epidermal growth factor receptor in breast cancer, from early stages to advanced disease. *J Exp Clin Cancer Res* 1999;18:347–55.

65. Huang YT, Hwang JJ, Lee PP, et al. Effects of luteolin and quercetin, inhibitors of tyrosine kinase, on cell growth and metastasis-associated properties in A431 cells overexpressing epidermal growth factor receptor. *Brit J Pharm* 1999;128:999–1010.

66. Woodburn JR. The epidermal growth factor receptor and its inhibition in cancer therapy. *Pharmacology & Therapeutics* 1999; 82:241–250.

67. Huang SM, Harari PM. Epidermal growth factor receptor inhibition in cancer therapy: biology, rationale and preliminary clinical results. *Investigational New Drugs* 1999;17:259–269.

68. Ciardiello F, Bianco R, Damiano V, et al. Antitumor activity of sequential treatment with topotecan and anti-epidermal growth factor receptor monoclonal antibody C225. *Clin Cancer Res* 1999; 5:909–916.

69. Martinez-Carpio PA, Mur C, Rosel P, et al. Constitutive and regulated secretion of epidermal growth factor and transforming growth factor-ß1 in MDA-MB-231 breast cancer cell line in 11-day cultures. *Cellular Signalling* 1999;11:753–757.

70. Ueda M, Ueki M, Terai Y, et al. Biological implications of growth factors on the mechanism of invasion in gynecological tumor cells. *Gynecol Obstet Invest* 1999;48:221–228.

71. Edwards J, Bartlett JM. Induction of apoptosis and inhibition of cellular proliferation as a dual approach to treatment of ovarian cancer. *Internat J Oncol* 1999;15:583–588.

72. Nouri AM, Zubairi ST, Russell MV, et al. Concordance between tumor cell activation by epidermal growth factor and alteration of cell sensitivity to cisplatin and lymphokine-activated killer cell activity. *Oncology Reports* 2000;7:197–201.

73. Fry DW. Inhibition of the epidermal growth factor receptor family of tyrosine kinases as an approach to cancer chemotherapy: progression from reversible to irreversible inhibitors. *Pharmacol Therapeut* 1999;82:207–218.

74. Seibert K, Lippman ME. Hormone receptor assays in breast cancer. *J Clin Immunoassay* 1983;6:5.

75. Syne JS, Panko WB. New techniques for the measurement of estrogen receptors in human breast cancer. *J Clin Immunoassay* 1983;6:17.

76. Miller LK, Tuazon FB, Niu E, et al. Human breast tumor estrogen receptor: effects of molybdate and electrophoretic analyses. *Endocrinology* 1981;108:1369–1378.

77. Valenstein SL, Voigt W, Thompsen S, et al. Estrogen and progesterone receptor techniques for breast cancer. In: Ashkar FS, ed. *Radiobioassays*, Vol. 2. Boca Rotan, FL: CRC Press, 1983:53–64.

78. Korenman SG, Dukes BA. Specific estrogen binding by the cytoplasm of human breast carcinoma. *J Clin Endocrinol Metab* 1970; 30:639–645.

79. Scatchard G. The attraction of proteins for small molecules and ions. *Ann NY Acad Sci* 1949;51:660–672.

80. Nickolson S, Sainsbury JRC, Needhan KG, et al. Quantitative assays of EGFr in human breast cancer: cut off points of clinical relevance. *Int J Cancer* 1988;42:36–41.

81. Steggles AW, King RJ. The use of protamine to study [6,7-3H]-oestradiol-17ß binding in rat uterus. *Biochem J* 1970;118:695–701.

82. Chen YM, Vaughn CB. Ammonium sulfate fractionation and assay of hormone receptors. *Cancer Invest* 1989;7:231–235.

83. Chamness GC, McGuire WL. Methods for analyzing steroid receptors in human breast cancer. In: McGuire WL, ed. *Breast Cancer. 3. Advances in Research and Treatment*. New York: Plenum Press, 1979:149–197.

84. Santine DB, Sibley CH, Perriand ER. A filter assay for steroid hormone receptors. *Biochemistry* 1973;12:2412–2416.

85. Jensen EV, Greene GL, Closs LE, et al. Immunochemical probes for receptor structure and function. In: Bresciani F, ed. *Perspectives in Steroid Receptor Research*. New York: Ravin Press, 1980:23–36.

86. Pavlik EJ, Nelson K, van Nagell JR Jr, et al. Steroid receptor analysis by size-exclusion liquid chromatography: considerations for the clinical laboratory. *Clin Chem* 1985;31:537–545.

87. Regnier FE, Gooding KM. High-performance liquid chromatography of proteins. *Anal Biochem* 1980;103:1–25.

88. Wittliff JL, Allegra JC, Day TG Jr, et al. Structural features and clinical significance of estrogen receptors. In: Moudgil VK, ed. *Steroid receptors in health and disease*. New York: Plenum Press, 1988:287–312.

89. Wagner RK, Jungblut PW. Estradiol and dihydrotestosterone receptors in normal and neoplastic mammary tissue. *Acta Endrocinol* 1976;82:105–120.

90. Vollmer G, Helmchen B, Knuppen R. Simultaneous quantification of oestrogen and progesterone receptors by a ligand-binding assay in frozen sections. *J Clin Chem Clin Biochem* 1989;27:953–959.

91. Greene GL, Nolan C, Engler JP, et al. Monoclonal antibodies to human estrogen receptors. *Proc Natl Acad Sci USA* 1980;77:157.

92. Abbott Laboratories, Diagnostic Division, North Chicago, IL 60064.

93. Senekjian EK, Press MF, Blough RR, et al. Comparison of the quantity of estrogen receptors in human endometrium and myometrium by steroid-binding assay and enzyme immunoassay based on monoclonal antibodies to human estrophilin. *Am J Obstet Gynecol* 1989;160:592–597.

94. Andersen J, Bentzen SM, Poulsen HS. Relationship between radioligand binding assay, immunoenzyme assay and immunohistochemical assay for estrogen receptors in human breast cancer and association with tumor differentiation. *Eur J Cancer Clin Oncol* 1988; 24:377–384.

95. Wu JT, Wilson L. Progesterone receptor: stability studies and correlation between steroid binding assay and enzyme immunoassay. *Clin Chem* 1988;34:1987–1991.

96. Smyth CM, Benn DE, Reeve TS. An enzyme immunoassay compared with a ligand-binding assay for measuring progesterone receptors in cytosols from breast cancers. *Clin Chem* 1988;34:1116–1118.

97. Holmes FA, Fritsche HA, Loewy JW, et al. Measurement of estrogen and progesterone receptors in human breast tumors: enzyme immunoassay versus binding assay. *J Clin Oncol* 1990;8:1025–1035.

98. Clayton F, Wu J. The lability of estrogen receptor: correlation of estrogen binding and immunoreactivity. *Clin Chem* 1986; 32:1774–1777.

99. Syne JS, Markaverich BM, Clark JH. Estrogen binding sites in the nucleus of normal and malignant human tissue: II. Characteristics of the multiple nuclear binding sites. *Cancer Res* 1982;42:4449–4454.

100. Zava DT, Harrington NY, McGuire WL. Nuclear estradiol receptor in the adult rat uterus: a new exchange assay. *Biochemistry* 1976;15:4292–4297.

101. Garola RE, McGuire WL. An improved assay for nuclear estrogen receptor in experimental and human breast cancer. *Cancer Res* 1977;37:3333–3337.

102. Traish AM, Muller RE, Wotiz HH. A new procedure for the quantitation of nuclear and cytoplasmic androgen receptors. *J Biol Chem* 1981;256:12028–12033.

103. Bresciani F, Sica V, Weizs HH. Effect of NaSCN on receptor-estradiol interactions and application to assay total receptor ("filled" and "unfilled" sites) in tissues and tissue fractions including nuclei, by exchange at low temperature with 17ß-estradiol-3H. In: Bresciani F, ed. *Perspectives in Steroid Receptor Research*. New York: Ravin Press, 1980: 273–297.

104. Fuqua SA, Falette NF, McGuire WL. Sensitive detection of estrogen receptor RNA by polymerase chain reaction assay. *J Natl Cancer Inst* 1990;82:858–861.

105. Lee YTN. Variability of steroid receptors in multiple biopsies of breast cancer. Effect of systemic therapy. *Breast Cancer Res Treat* 1982;2:185–193.

106. Nicholson S, Sainsbury JR, Needham GK, et al. Quantitative assays of epidermal growth factor receptor in human breast cancer: cut-off points of clinical relevance. *Int J Cancer* 1988;42:36–41.

107. Formento JL, Francoual M, Formento P, et al. Epidermal growth factor receptor assay: validation of a single point method and application to breast cancer. *Breast Cancer Res Treat* 1991;17:211–219.

108. Carpenter G. Binding assays for epidermal growth factor. *Methods Enzymol* 1985;109:101.

109. Grimaux M, Laine-Bidron C, Magdelenat H. Immunoenzymetric assay of epidermal growth factor receptor. Application to breast tumor samples. *Tumour Biol* 1989;10:215–224.

110. Lin PH, Selinfreund R, Wharton W. Western blot detection of epidermal growth factor receptor from plasmalemma of culture cells using 125I-labeled epidermal growth factor. *Anal Biochem* 1987;167:128–139.

111. Stumpf WF, Roth IJ. High resolution autoradiography with dry-mounted, freeze-dried, frozen sections. Comparative study of six methods using two different compounds [3H]-estradiol and [3H]-mesobilirubinogen. *J Histochem Cytochem* 1966;14:274.

112. Chen PZ, Mei Z, Yao XY, et al. Selection of hormone-responsive advanced breast cancer with a cytoplasmic estrogen receptor assay. Analysis of 100 cases. *Cancer* 1989;63:139–142.

113. King WJ, Jensen EV, Miller I, et al. Immunocytochemical detection of estrogen receptor in frozen sections of human breast tumors with monoclonal anti-receptor antibodies. Proceedings of the 64th Annual Meeting of the Endocrine Society: San Francisco, 1982.

114. Graham ML II, Bunn PA Jr, Jewett PB, et al. Simultaneous measurement of progesterone receptors and DNA indices by flow cytometry: characterization of an assay in breast cancer cell lines. *Cancer Res* 1989;49:3934–3942.

115. Virgolini I, Raderer M, Kurtaran A, et al. Vasoactive intestinal peptide-receptor imaging for the localization of intestinal adenocarcinomas and endocrine tumors. *N Eng J Med* 1994;331:1116.

116. Rutherford RAD, Matsuda Y, Wilkins MR, et al. Identification of renal natriuretic peptide receptor subpopulations by the use of the non-peptide antagonist, HS-142-1. *Br J Pharmacol* 1994;113:931.

117. Day RN, Nordeen SK, Wan Y. Visualizing protein-protein interactions in the nucleus of the living cell. *Mol Endo* 1999;13:517.

118. Guiochon-Mantel A, Delabre K, Lescop P, et al. Intracellular traffic of steroid hormone receptors. *J Steroid Biochem Mol Biol* 1996;56:1.

119. Htun H, Barsony J, Renyi I, et al. Visualization of glucocorticoid receptor translocation and intranuclear organization in living cells with green fluorescent protein chimera. *Proc Natl Acad Sci USA* 1996;93:4845.

120. Georget V, Lobaccaro JM, Terouanne B, et al. Trafficking of the androgen receptor in living cells with fused green fluorescent protein-androgen receptor. *Mol Cell Endocrinol* 1997;129:17.

121. Carey KL, Richards SA, Lounsbury KM, et al. Evidence using a green fluorescent protein-glucocorticoid receptor chimera that the Ran/TC4GTPase mediates and essential function independent of nuclear protein import. *J Cell Biol* 1996;133:985.

122. Day RN. Visualization of Pit-1 transcription factor interactions in the living cell nucleus by fluorescence resonance energy transfer microscopy. *Mol Endocrinol* 1998;12:1410.

123. Widengren J, Rigler R. Review — fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces. *Cell Mol Biol* 1998;44:857.

124. Edelstein SJ, Schaad O, Changeux JP. Single binding versus single channel recordings: a new approach to study ionotropic receptors. *Biochemistry* 1997;36:13755.

125. Thompson VL. *Topics in Fluorescence Spectroscopy; Techniques* vol. 1, Lakowicz JR, ed., New York: Plenum Press, 1991:378.

126. Wohland T, Friedrich K, Hovius R, et al. Studies of ligand-receptor interactions by fluorescence correlation spectroscopy with different fluorophores: Evidence that homopentameric 5-hydroxytryptamine type 3AS receptor binds only one ligand. *Biochemistry* 1999;38:8671.

127. Parker RMC, Barnes NM. mRNA: Detection by in situ and northern hybridization. In: Keen M, ed. *Methods in Molecular Biology; Receptor Binding Techniques*, vol. 106, Totowa, New Jersey: Humana Press, 1999:247.

128. Lai J, Porreca F. Antisense targeting of delta opioid receptors. In: Raffa RB, Porreca F, eds. *Antisense Strategies for the Study of Receptor Mechanism*. New York: Chapman & Hall, 1996.

129. Raffa RB. In vivo antisense strategy for the study of second-messengers: application to G-proteins. In: Raffa RB, Porreca F, eds. *Antisense strategies for the study of receptor mechanism*. New York: Chapman & Hall, 1996.

130. Humphries SE, Gudnason V, Whittall R, et al. Single-strand conformation polymorphism analysis with high throughput modifications, and its use in mutation detection in familial hypercholesterolemia. International Federation of Clinical Chemistry Scientific Division: Committee on Molecular Biology Techniques. *Clinl Chem* 1997;43:427–35.

131. Cantafora A, Bertolini S, Calandra S. Familial hypercholesterolemia and mutations of the gene for low-density lipoproteins. In: *Italy Annali dell Istituto Superiore di Sanita* 1999;35:177–184.

132. OMIM, McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), 2000. World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/)

133. Hubbard EW. Tumor markers. *Diag Clin Testing* 1990;28:13–17.

134. Duffy MJ. Biochemical markers as prognostic indices in breast cancer. *Clin Chem* 1990;36:188.

135. Henry JA, Angus B, Horne CHW. Oestrogen receptor and oestrogen regulated proteins in human breast cancer: a review. *Keio J Med* 1989;38:241–261.

136. Knight III WA, Livingston RB, McGuire WL. Estrogen receptor as an independent prognostic factor for early recurrence in breast cancer. *Cancer Res* 1977;37:4669–4671.

137. Thorpe SM. Estrogen and progesterone receptor determinations in breast cancer. Technology, biology and clinical significance. *Acta Oncol* 1988;27:1–19.

138. McGuire WL. Prognostic factors in primary breast cancer. *Cancer Surveys* 1986;5:527–536.

139. Hill P. Leanness, peptide hormones and premenopausal breast cancer. *Med Hypotheses* 1989;28:45–50.

140. Digenis AG, Ross CB, Morrison JG, et al. Carcinoma of the male breast: a review of 41 cases. *South Med J* 1990;83:1162–1167.

141. Scheres HM, De Goeij AF, Rousch MJ, et al. Quantification of oestrogen receptors in breast cancer: radiochemical assay on cytosols and cryostat sections compared with semiquantitative immunocytochemical analysis. *J Clin Pathol* 1988;41:623–632.

142. Snijders MP, De Goeij AF, Koudstaal J, et al. Is immunohistochemical analysis of oestrogen and progesterone receptors in endometrial carcinoma superior to the radioligand binding assay? *J Pathol* 1990;161:129–135.

143. Raam S, Vrabel DN. Evaluation of an enzyme immunoassay kit for estrogen receptor measurement-report II. *Clin Chem* 1988;34:2053.

144. Sutherland RL, Murphy LC. The binding of tamoxifen to human mammary carcinoma cytosol. *Europ J Cancer* 1980;16:1141.

145. Valavaara R, Tuominen J, Johansson R. Predictive value of tumor estrogen and progesterone receptor levels in postmenopausal women with advanced breast cancer treated with toremifene. *Cancer* 1990;66:2264–2269.

146. Sainsbury J, Farndon JR, Needham GK, et al. Epidermal-growth-factor receptor status as predictor of early recurrence of death from breast cancer. *Lancet* 1987:1398–1402.

147. Varley JM, Swallow JE, Brammar WJ, et al. Alterations to either c-erbB-2(neu) or c-myc proto-oncogenes in breast carcinoma correlate with poor short-term prognosis. *Oncogene* 1987;1:423–430.

148. Clair T, Miller WR, Cho-Chung YS. Prognostic significance of the expression of a ras protein with a molecular weight of 21,000 by human breast cancer. *Cancer Res* 1987;47:5290–5293.

149. Ehrlich CE, Young PCM, Stehman FB, et al. Steroid receptors and clinical outcome in patients with adenocarcinoma of the endometrium. *Am J Obstet Gynecol* 1988;158:796–807.

150. Rauramo L, Gronroos M, Kyostila J. Prognosis of carcinoma corporis uteri managed by preoperative radium treatment. *Acta Obstet Gynecol Scand* 1968;47:517–527.

151. Gronroos M, Punnonen R, Rauramo L, et al. Problems in the treatment of endometrial carcinoma. *Ann Chir Gynecol* 1974;63:141–145.

152. Kauppila A, Gronroos M, Nieminen U. Clinical outcome in endometrial cancer. *Obstet Gynecol* 1982;60:463–480.

153. Tiitinen A, Forss M, Aho I, et al. Endometrial adenocarcinoma: clinical outcome in 881 patients and analysis of 146 patients whose deaths were due to endometrial cancer. *Gynecol Oncol* 1986;25:11–19.

154. Podratz KC, O'Brien PPPC, Malkasian GD, et al. Effects of progestational agents in treatment of endometrial carcinoma. *Obstet Gynecol* 1985;66:106–110.

155. Aalders JG, Abeler V, Kolstad P. Recurrent adenocarcinoma of the endometrium: a clinical and histopathological study of 379 patients. *Gynecol Oncol* 1984;17:85–103.

156. Horm JW, Asire AJ, Young JL Jr, et al. SEER program, cancer incidence and mortality in the United States 1973–81. In: NIH Publication No.85:1837, 1984.

157. Weiss RJ, Lucas WE. Adenocarcinoma of the uterine cervix. *Cancer* 1986;57:1996–2001.

158. Vesterinen E, Forss M, Nieminen U. Increase of cervical adenocarcinoma: a report of 520 cases of cervical carcinoma including 112 tumors with glandular elements. *Gynecol Oncol* 1989;33:49–53.

159. Kauppila A. Oestrogen and progesterone receptors as prognostic indicators in endometrial cancer. A review of the literature. *Acta Oncol* 1989;28:561–566.

160. Friberg LG, Kullander S, Persijn JP, et al. On receptors for estrogens (E2) and androgens (DHT) in human endometrial carcinoma and ovarian tumors. *Acta Obstet Gynecol Scand* 1978;57:261–264.

161. Vihko R, Janne O, Kauppila A. Steroid receptors in normal, hyperplastic and malignant human endometria. *Ann Clin Res* 1980;12:208–215.

162. Quinn MA, Pearce P, Fortune DW, et al. Correlation between cytoplasmic steroid receptors and tumour differentiation and invasion in endometrial carcinoma. *Br J Obstet Gynaecol* 1985;92:399–406.

163. Kleine W, Maier T, Geyer H, et al. Estrogen and progesterone receptors in endometrial cancer and their prognostic relevance. *Gynecol Oncol* 1990;38:59–65.

164. Creasman WT, McCarty KS, Barton TK, et al. Clinical correlates of estrogen- and progesterone-binding proteins in human endometrial adenocarcinoma. *Obstet Gynecol* 1980;53:363–370.

165. Kauppila A, Kujansuu E, Vihko R. Cytosol estrogen and progesterone receptors in endometrial carcinoma of patients treated with surgery, radiotherapy and progestin. Clinical correlates. *Cancer* 1982;50:2157–2162.

166. Chambers JT, MacLusky N, Eisenfield A, et al. Estrogen and progesterone receptor levels as prognosticators for survival in endometrial cancer. *Gynecol Oncol* 1988;31:65–81.

167. Palmer DC, Muir IM, Alexander AI, et al. The prognostic importance of steroid receptors in endometrial carcinoma. *Obstet Gynecol* 1988;72:388–393.

168. Martin JD, Haahnel R, McCartney AJ, et al. The effect of estrogen receptor status on survival in patients with endometrial adenocarcinomas. *Am J Obstet Gynecol* 1983;147:322.

169. Creasman WT, Soper JT, McCarty KS Jr, et al. Influence of cytoplasmic steroid receptor content on prognosis of early stage endometrial carcinoma. *Am J Obstet Gynecol* 1985;151:922–932.

170. Borazjani G, Twiggs LB, Leung BS, et al. Prognostic significance of steroid receptors measured in primary metastatic and recurrent endometrial carcinoma. *Am J Obstet Gynecol* 1989;161:1253–1257.

171. Sutton GP, Geisler HE, Stehman FB, et al. Features associated with survival and disease-free survival in early endometrial cancer. *Am J Obstet Gynecol* 1989;160:1385–1393.

172. Kleine W, Maier T, Geyer H, et al. Estrogen and progesterone receptors in endometrial cancer and their prognostic relevance. *Gynecol Oncol* 1990;38:59–65.

173. Podratz KC. Hormonal therapy in endometrial carcinoma. Recent results. *Cancer Res* 1990;118:242–251.

174. Kauppila AJ, Isotalo HE, Kivinen ST, et al. Prediction of clinical outcome with estrogen and progesterone receptor concentrations and their relationships to clinical and histopathological variables in endometrial cancer. *Cancer Res* 1986;46:5380–5384.

175. Senler TI, Dean WL, Murray LF, et al. Quantification of cytochrome P-450-dependent cyclohexane hydroxylase activity in normal and neoplastic reproductive tissues. *Biochem J* 1985;227:379–387.

176. Kauppila AJ, Isotalo HE, Kivinen S, et al. Short term effects of danazol and medroxyprogesterone acetate on cytosol and nuclear estrogen and progesterone receptors, 17β-hydroxy steroid dehydrogenase activity, histopathology and ultrastructure of human endometrial adenocarcinoma. *Int J Cancer* 1985;35:157–163.

177. Barkovitz GD, Brown TR, Migeon C J. Androgen Receptors. *Clin Endo Metab* 1983;12:155–173.

178. Rennie PS, Bruchovsky N, Goldenberg SL. Relationship of androgen receptors to the growth and regression of the prostate. *Am J Clin Oncol* 1988;11 (suppl 2):S13–S17.

179. Mulder D, van Loon D, de Boer W, et al. Mechanism of androgen action: recent observations on the domain structure of androgen receptors and the induction of EGF-receptors by androgens in prostate tumor cells. *J Steroid Biochem* 1989;32:151–156.

180. Tilley WD, Marcelli M, McPhaul MJ. Recent studies of the androgen receptor: new insights to old questions. *Mol Cell Endo* 1990;68:C7–C10.

181. Maddy SQ, Chisholm GD, Busuttil A, et al. Epidermal growth factor receptors in human prostate cancer: correlation with histological differentiation of the tumour. *Br J Cancer* 1989;60:41–44.

182. Davies P, Eaton CL, France TD, et al. Growth factor receptors and oncogene expression in prostate cells. *Am J Clin Oncol* 1988;11(suppl 2):S1–S7.

183. Namiki M, Kitamura M, Nonomura N, et al. Effect of estrogen treatment on testicular gonadotropin receptors in prostatic cancer patients. *Arch Androl* 1987;19:249–252.

184. Murphy JB, Emmott RC, Hicks LL, et al. Estrogen receptors in the human prostate, seminal vesicle, epididymis, testis, and genital skin: a marker for estrogen-responsive tissues? *J Clin Endocrinol Metab* 1980;50:938–948.

185. Kadar T, Redding TW, Ben-David M, et al. Receptors for prolactin, somatostatin, and luteinizing hormone-releasing hormone in experimental prostate cancer after treatment with analogs of luteinizing hormone-releasing hormone and somatostatin. *Proc Natl Acad Sci USA* 1988;85:890–894.

186. Benson RC Jr, Gorman BA, O'Brien PC, et al. Relationship between androgen receptor binding activity in human prostate cancer and clinical response to endocrine therapy. *Cancer* 1987;59:1599–1606.

187. Connelly JG, Mobbs BG. Clinical applications and value of receptor levels in Treatment of prostate cancer. *Prostate* 1984;5:477–483.

188. Demura T, Kuzumaki N, Oda A, et al. Establishment of monoclonal antibody to human androgen receptor and its clinical application for prostatic cancers. *Am J Clin Oncol* (CCT) 1988;11(suppl 2):S23–S26.

189. Baulieu EE. Hormones: a complex communication network. In: Baulieu EE, Kelly PA, eds. *Hormones: From Molecules to Disease.* New York: Hermann, 1990:140.

190. Griffin JE, Wilson JD. The syndromes of androgen resistance. *N Engl J Med* 1980;302:198–209.

191. Smallridge RC, Vigersky R, Glass AR, et al. Androgen receptor abnormalities in identical twins with oligospermia. *Am J Med* 1984;77:1049–1054.

192. Yaegashi N, Uehara S, Senoo M, et al. Point mutations in the steroid-binding domain of the androgen receptor gene of five Japanese patients with androgen insensitivity syndrome. *Tohoku Journal of Experimental Medicine* 1999;187:263–272.

193. Knoke I, Jakubiczka S, Lehnert H, et al. A new point mutation of the androgen receptor gene in a patient with partial androgen resistance and severe oligozoospermia. *Andrologia* 1999;31:199–201.

194. Melo KF, Latronico AC, Costa EM, et al. A novel point mutation (R840S) in the androgen receptor in a Brazilian family with partial androgen insensitivity syndrome. *Human Mutation* 1999;14:353.

195. Shkolny DL, Beitel LK, Ginsberg J, et al. Discordant measures of androgen-binding kinetics in two mutant androgen receptors causing mild or partial androgen insensitivity, respectively. *J Clin Endo Metab* 1999;84:805–10.

196. Peters I, Weidemann W, Romalo G, et al. An androgen receptor mutation in the direct vicinity of the proposed C-terminal alpha-helix of the ligand binding domain containing the AF-2 transcriptional activating function core is associated with complete androgen insensitivity. *Mol Cell Endo* 1999;148:47–53.

197. Uehara S, Tamura M, Nata M, et al. Complete androgen insensitivity in a 47,XXY patient with uniparental disomy for the X chromosome. *Am J Med Genet* 1999;86:107–111.

198. Yong EL, Tut TG, Ghadessy FJ, et al. Partial androgen insensitivity and correlations with the predicted three dimensional structure of the androgen receptor ligand-binding domain. *Mol Cell Endo* 1998;137:41–50.

199. Komori S, Sakata K, Kasumi H, et al. A substitutional mutation in the DNA binding domain of the androgen receptor causes complete androgen insensitivity syndrome. *Gyn Endo* 1999;13:327–332.

200. Gottlieb B, Vasiliou DM, Lumbroso R, et al. Analysis of exon 1 mutations in the androgen receptor gene. *Human Mutation* 1999;14:527–539.

201. Hellwinkel OJ, Bull K, Holterhus PM, et al. Complete androgen insensitivity caused by a splice donor site mutation in intron 2 of the human androgen receptor gene resulting in an exon 2-lacking transcript with premature stop-codon and reduced expression. *J Steroid Biochem Mol Biol* 1999;68:1–9.

202. Yong EL, Lim J, Qi W, et al. Molecular basis of androgen receptor diseases. *Ann Med* 2000;32:15–22.

203. Yong EL, Ghadessy F, Wang Q, et al. Androgen receptor transactivation domain and control of spermatogenesis. *Rev Reprod* 1998;3:141–144.

204. Wieacker PF, Knoke I, Jakubiczka S. Clinical and molecular aspects of androgen receptor defects. *Exper Clin Endo Diabetes* 1998; 106:446–453.

205. Zhu YS, Cai LQ, Cordero JJ, et al. A novel mutation in the CAG triplet region of exon 1 of androgen receptor gene causes complete androgen insensitivity syndrome in a large kindred. *J Clin Endo Metab* 1999;84:1590–1594.

206. Holterhus PM, Sinnecker GH, Wollmann HA, et al. Expression of two functionally different androgen receptors in a patient with androgen insensitivity. *Euro J Ped* 1999;158:702–706.

207. Holterhus PM, Wiebel J, Sinnecker GH, et al. Clinical and molecular spectrum of somatic mosaicism in androgen insensitivity syndrome. *Pediatric Research* 1999;46:684–690.

208. Ahmed SF, Cheng A, Dovey L, et al. Phenotypic features, androgen receptor binding, and mutational analysis in 278 clinical cases reported as androgen insensitivity syndrome. *J Clin Endocrin Metab* 2000;85:658–665.

209. Warne GL, Zajac JD, MacLean HE. Androgen insensitivity syndrome in the era of molecular genetics and the Internet: a point of view. *J Ped Endo Metab* 1998;11:3–9.

Section 11
Blood Bank/Transfusion Medicine

Blood Bank/Transfusion Medicine - Introduction

Harold A. Oberman

Section Chief

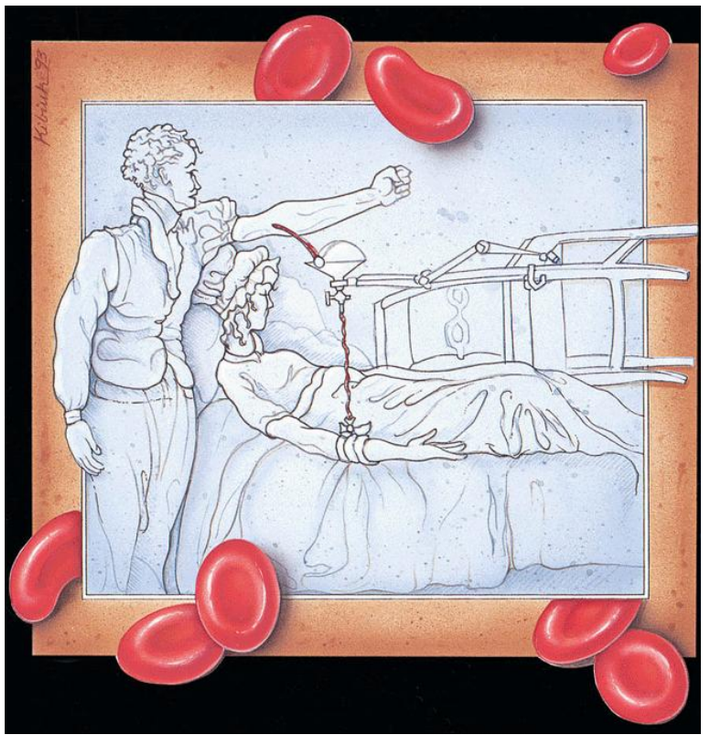


Figure.

The title of this section reflects the remarkable change in this area of clinical laboratory medicine in recent years. The term blood banking is far too restrictive, and its implied scope too limited, to characterize this subspecialty. The procurement and processing of blood for transfusion is only one facet of this field, as it has been modified to include a breadth of activities unforeseen only a few decades ago. The discipline has been clearly established as one that bridges the activities of the laboratory with those of clinical medicine. The uniqueness of this practice resulted in its recognition as a medical subspecialty by the American Board of Pathology, with the creation of a certifying examination 20 years ago. Similarly, the role of the medical technologist in this area has become more sharply defined, and also has resulted in the creation of subspecialty certification. The transfusion medicine medical specialist must translate the laboratory's results and activities to the patient's bedside. Therefore, it is insufficient for such an individual merely to be cognizant of immunohematology in a restrictive sense.

The following chapters address critical clinical activities such as apheresis procedures, blood component therapy, clinical consultation regarding adverse reactions to transfusion, the clinical significance of various blood group antibodies, and the transfusional management of various hematologic conditions. The laboratorian who assumes responsibility in this area must combine an understanding of clinical medicine with knowledge of the laboratory aspects of immunohematology.

68

Organization, Functions, Regulation, and Legal Concerns of Blood Banks

Harold A. Oberman

THE EVOLUTION OF BLOOD TRANSFUSION

Part of "68 - Organization, Functions, Regulation, and Legal Concerns of Blood Banks"

The Dawn of Blood Transfusion

Although blood has been considered as a therapeutic agent since ancient times, there was little rational, scientific basis for its use, nor was there full realization of its properties or of its dangers until the present century. Among the earliest references to blood transfusion are those in the Bible and in the writings of Ovid. At several points in the Scriptures there is admonition against transfusion (1), and this has resulted in the religious basis for the refusal of transfusion by such groups as Jehovah's Witnesses.

Ovid, in the eighth book of *The Metamorphoses*, describes restoration of youth to Aeson by the sorceress Medea through a “transfusion” of a magical potion. Unfortunately, she was unable to duplicate this remarkable feat. As other evidence of the mysticism associated with blood, Galen advised the drinking of the blood of a weasel for the cure of rabies, and Pliny described the drinking of the blood of dying gladiators as a cure for epilepsy (2).

Any rational approach to transfusion required an understanding of the circulation of the blood, and this was not available until the work of William Harvey in the early 17th century. By the middle of that century Richard Lower in London and Jean Denys in Paris were experimenting with blood transfusion in dogs, and both subsequently transfused humans with canine blood (3). Unfortunately, Denys, in 1667, transfused one of his subjects on three separate occasions, and the final transfusion eventuated in the first recorded example of a fatal acute hemolytic transfusion reaction. Even though a trial relieved Denys of guilt for this mishap (the subject's wife was subsequently found to have poisoned him), the controversy stemming from this case led to a multitude of decrees forbidding the practice, and resulted in the virtual elimination of transfusion for the next 150 years.

Nineteenth Century Blood Transfusion

The reawakening of interest in transfusion in the early 19th century was led by James Blundell, one of the most prominent obstetricians in London. He looked to transfusion as the most appropriate treatment for postpartum hemorrhage, and, in contrast to his 17th century forebears, reasoned that the father would be a more likely source of the blood than an animal (4). After the seemingly successful transfusion of a man dying of gastric carcinoma (5), Blundell undertook transfusions of patients with postpartum hemorrhage and consulted with other obstetricians who shared his enthusiasm for the procedure.

Transfusion during the 19th century was, at best, primitive. Among the indications cited in an 1849 publication were dyspepsia, esophageal stricture, dysentery, and prolonged fever (6). Moreover, the increased use of the technique without knowledge of blood groups or of immunology led to untoward complications. Therefore, by the last quarter of the century, popularity of transfusion waned, and substitutes for blood were sought. Among these substitutes were milk, alcohol, and saline solution.

The Modern Era

Landsteiner's recognition of the ABO blood group system at the outset of the 20th century permitted a rational approach to blood transfusion (7). Within the subsequent few years, Ottenberg recognized the value of reacting the blood of the donor with that of the recipient before transfusion to detect incompatibility (8). At the same time, the technique of direct transfusion was popularized, initially by vascular surgeons adept at anastomosing blood vessels of donor and recipient, and subsequently through the use of a variety of mechanical devices.

The problems of anticoagulation, preservation, and storage of blood had to be solved before indirect transfusion could be accomplished. The first World War served as an impetus to resolution of the first of these problems, as, near simultaneously, several investigators advocated citrate as an anticoagulant in 1914, and 2 years later Rous and Turner documented the value of adding sugar to the citrate as a preservative (9). This allowed the use of preserved blood for the transfusion of casualties by 1918 (10).

Use of anticoagulated stored blood for transfusion, so-called indirect transfusion, was slow to achieve greater popularity than direct transfusion because of the large amount of anticoagulant required in the early formulations. World War II, with the need

to transport blood for long distances, hastened development of improved preservative solutions so that by the early 1940s direct transfusion had all but vanished. Preservative solutions were modified initially through variations in the proportion of sodium citrate, citric acid and dextrose (11), and more recently through addition of sodium phosphate and adenine.

The development of organized centers for blood preservation also awaited the availability of reliable mechanical refrigeration equipment. The first blood banks did not appear until the mid-1930s, and not until 1937 was the first blood bank organized in the United States (12). World War II provided a remarkable impetus for blood banking. More than 13 million units of blood were collected during the war, principally by the American Red Cross, although much of it was used for fractionation into dried plasma.

The developments in this field in the years since World War II have been remarkable. One must appreciate that only the ABO, Rh, P, and MNS systems were known before 1945. Since that time, the understanding and complexity of these systems has increased and a multitude of new systems has been identified. In 1945, the antiglobulin reaction was introduced (13), and this facilitated detection of blood group antibodies.

In a contemporary development, stemming from attempts to meet the blood requirements of the Texas City, Texas, disaster, the American Association of Blood Banks was formed in 1947. While the major initial goal of the organization was to enhance blood availability through a national network of blood banks, of even greater significance were its additional contributions to safe and effective blood banking. The American Association of Blood Banks provided a focus for educational activities in the field, and its promulgation of standards of practice and its inspection and accreditation program were to play a major role in enhancing the safety of blood transfusion throughout the world.

The advent of plastic transfusion equipment, an innovation of the past 30 years, circumvented such complications of transfusion in glass bottles as pyrogen-related reactions and air embolism. It also permitted separation of blood components in a closed system, thereby providing the impetus for the current emphasis on component therapy. Plasmapheresis, first described in 1914 (14) was applied to large-scale plasma collection in the 1960s and, in the following decade, to treatment of disease. The latter has resulted in the technical subspecialty of therapeutic apheresis. By the end of the 1960s, in another major development with far-reaching clinical implications, the pathogenesis, treatment, and prevention of hemolytic disease of the newborn were defined.

During the past two decades, there has been increased emphasis on the infectious complications of transfusions, spurred by the development of tests for posttransfusion hepatitis and the advent of the AIDS epidemic. Continued refinement of these tests, enhanced by use of molecular biologic techniques, as well as pretreatment of some blood components to eliminate infectious organisms, has exemplified the yet unachieved effort to provide a blood supply free of risk of posttransfusion infection. Finally, the breadth of this field and the multitude of clinical issues it embraces have led to the recognition of Transfusion Medicine as a medical specialty.

BLOOD BANKS AND TRANSFUSION SERVICES

Part of "68 - Organization, Functions, Regulation, and Legal Concerns of Blood Banks"

Definitions and Functions of Blood Banks

The Standards of the American Association of Blood Banks defines a blood bank as an organization that collects, stores, and processes human blood. In contrast, a transfusion service tests the blood of the intended recipient and is concerned with the transfusion of the blood and its components. While this definition may be useful in distinguishing the roles of a community blood bank and a hospital transfusion program, it must be appreciated that it is an arbitrary and somewhat artificial distinction.

Blood banking embraces a variety of functions, including the recruitment and selection of blood donors; collection and typing of blood; preparation of blood components; testing of the blood for transfusion-transmitted diseases; distribution of the blood to the transfusing facility; pretransfusion testing and issuance of the blood to the patient; consultation on antibody-related problems and on indications for transfusion; and investigation of adverse reactions to transfusion.

Apheresis procedures play an important role, both in community and hospital blood banks. The technique is utilized for collection of blood components and for therapeutic procedures. Collection of peripheral blood stem cells by apheresis has facilitated both autologous and allogeneic bone marrow transplantation. A major therapeutic use of apheresis, or plasma exchange, is for treatment of thrombotic thrombocytopenic purpura/hemolytic uremic syndrome. This has allowed a successful therapeutic approach to these once-uniformly fatal conditions. Recent applications of therapeutic apheresis have included lipo-apheresis for reduction of low-density lipoprotein (LDL) cholesterol in patients with homozygous familial hypercholesterolemia and photo-pheresis for treatment of cutaneous T-cell lymphoma.

Many blood banks perform testing for resolution of disputed parentage, histocompatibility testing, and processing of bone marrow for autologous or allogeneic bone marrow transplantation. Finally, some institutions provide a program wherein outpatients come to the blood bank for their transfusional therapy, or, in some instances, are responsible for a program of home transfusion.

Therefore, the scope of the specialty is broad, embracing a variety of personnel, including donor recruiters, specialized administrators, nurses, technologists and physicians. In recent years the medical aspects have been grouped under the rubric of Transfusion Medicine.

Sources of Blood

Approximately 12.5 million units of blood are collected annually in the United States, and use of blood component therapy permits a much larger number of patients to be transfused. The number of donated units of blood available for transfusion purposes has declined in recent years for several reasons. Ill-founded concern regarding transmission of AIDS made people reluctant to donate in the mid-1980s. More recently, the more selective screening tests used to reduce the risk of transfusion-related infections have resulted in an increase in positive test results, resulting in loss of blood donors.

Most of the donated blood in the United States is collected by independent community blood banks, or by subsidiaries of the American Red Cross Blood Program. Many of the former blood banks are organized under the title of America's Blood Centers, formerly termed the Council of Community Blood Centers. These blood banks and the Red Cross program collect more than 90% of the blood transfused in the country, and hospital blood banks supply the rest (15). Centralization of the blood collection and processing function has been spurred by increased computerization of these functions as well as the advent of sophisticated tests for reducing the risk of transmission of transfusion-related infectious disease.

Fear of transfusion-transmitted infection, precipitated by the advent of AIDS, resulted in increased interest by the public in autologous blood provision and in directed donation programs, wherein a friend or relative provides blood for the patient. However, the introduction of more sensitive tests for Hepatitis C virus (HCV) and human immunodeficiency virus (HIV), accompanied by publicity emphasizing the increased safety of the blood supply, has occasioned a decline in both programs. Another factor in the decreased popularity of these programs is the failure of medical insurance carriers to reimburse such donations.

The Hospital Transfusion Service

Pretransfusion testing and issuance of the blood to the patient for transfusion is the responsibility of the hospital transfusion service. However, in some locales this function has been assumed by community blood banks. The latter centralized transfusion service is associated with cost savings in personnel and equipment for the hospital, as well as uniformity in the utilization of blood in the community. Therefore, it would be seen as a positive step in the current era of managed care. However, centralizing this function in the community or region likely inhibits the presence of an active transfusion medicine consultation program in the hospital.

The physician responsible for the hospital transfusion program not only is administratively responsible for all of the technical and medical policies and procedures of the laboratory, but also for serving as a consultant for any related problems (16). These responsibilities include designation of the source of the blood and components utilized in the hospital, the storage, processing and issuance of the blood and provision of a consultative service to enhance patient care. This includes responsibility for implementation of a quality-control program and for compliance with the requirements of accrediting agencies. The blood bank physician also may chair the transfusion committee of the medical staff and serves as an educational resource for the hospital's physicians on appropriate use of blood and components.

REGULATION OF BLOOD BANKS AND TRANSFUSION SERVICES

Part of "68 - Organization, Functions, Regulation, and Legal Concerns of Blood Banks"

The Standards of the American Association of Blood Banks (AABB) form the basis for the Association's accreditation and inspection program. The Standards were first published in 1958 in an effort to improve the quality and safety of blood transfusion. Prepared by the Committee on Standards, a group of knowledgeable physicians and technologists, each specified standard results from a consensus decision by the committee. It generally is agreed that the Standards are the most definitive statement of minimum performance guidelines of the practice of blood banking.

The AABB inspection program is a voluntary one, and is conducted every 2 years. Inspections are conducted as a prescheduled peer review undertaking, and there is an opportunity to correct deficiencies. Following correction of deficiencies a certificate of accreditation is issued. Therefore, the inspection should be perceived as an educational, as well as an accrediting, exercise.

Blood Banks also are inspected by the College of American Pathologists (CAP) as part of their extensive biannual inspection program. This program also is conducted by physicians and technologists, and is prescheduled. In contrast to the AABB inspection, it covers all hospital laboratories including the blood bank. Like the inspection of the AABB, the inspected laboratory is provided with a list of deficiencies and is given an opportunity for correction; therefore, the inspection also has a major educational component.

Federal responsibility for regulation of blood banks is vested in the Food and Drug Administration (FDA). Whereas only blood banks engaged in shipment of blood across state lines must be federally licensed, since 1975 all transfusion facilities have been required to register with the FDA. Regulation is based on conformity to the *Code of Federal Regulations* (CFR), published by the Center for Biologics Evaluation and Research of the FDA. All of the material pertaining to blood banking is contained in Title 21, Chapter 1 of the booklet. This forms the basis for the FDA inspection program, required of all licensed establishments.

In contrast to the peer-conducted AABB and CAP inspections, the annual inspections conducted by the FDA, which are necessary for maintenance of licensure, are unannounced. They are conducted by specially trained FDA personnel, who often lack any work experience in blood banking.

Similarly, the Department of Public Health of many states may conduct periodic unannounced inspections. This is based on statutory vestment of responsibility in such a department for the procurement, processing, distribution and use of blood, its components and its derivatives.

Finally, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) inspects all hospital laboratories, including blood banks, in the course of their annual or biannual inspection. This is a prescheduled inspection conducted by JCAHO-trained inspectors. The inspecting team includes a physician and nurse, as well as other individuals, one of whom may be a medical technologist. Like the inspections of the AABB and CAP, the JCAHO inspection includes a summation conference with the inspected institution and an opportunity to correct deficiencies before issuance of a final report. The type and number of deficiencies determines whether the institution will subsequently be reviewed annually or biannually.

Finally, specialty functions of blood banks also may be inspected. For example, those blood banks having a blood irradiator are inspected by the Nuclear Regulatory Commission. Similarly,

the bone marrow or stem cell processing function, as well as the histocompatibility testing area, is inspected by the Foundation for the Accreditation of Hematopoietic Cell Therapy (FAHCT). In addition, the blood bank may be inspected during external review of various hospital programs wherein they play an ancillary role, such as support of a trauma program.

LEGAL ISSUES IN BLOOD BANKING

Part of "68 - Organization, Functions, Regulation, and Legal Concerns of Blood Banks"

While a complete review of the subject of liability of blood banks is beyond the scope of the following discussion, it is appropriate to present a summary of some of the major issues confronting blood banks and transfusion services in this area. Of all clinical laboratories, the blood bank is most subject to legal action. Moreover, the variety of activities that may result in such actions is much wider than those present in any other clinical laboratory. In recent years, many claims have related to transfusion-transmitted infection, especially AIDS and hepatitis. Plaintiffs have sought damages on such disparate allegations as improper screening of blood donors, failure of compliance of the transfusion with clinically accepted guidelines, failure to inform the patient of the dangers of the transfusion, or failure to advise the patient of the alternatives to homologous transfusion.

Because the current litigious atmosphere promises increasing activity in this arena, the following comments are intended to serve as a framework to understand this important topic.

Malpractice

Malpractice is a lawsuit brought by an allegedly injured plaintiff against a defendant, in this instance a health-care professional, for breach of a recognized standard of care. In most instances, the defendants in a transfusion-related suit will include the director of the hospital blood bank, the hospital, the blood bank that provided the blood, and, very likely, the patient's physician. The independent acts of the patient's physician do not necessarily insulate the laboratory's responsibility, or vice versa (17).

A successful action requires demonstration of negligence on the part of the defendant, wherein there has been a breach of duty that has resulted in damages to the plaintiff. The plaintiff may be either a blood donor who has been injured pursuant to the donation procedure or a patient who has suffered an adverse reaction to transfusion or has acquired a posttransfusion infection.

The basic elements of such a lawsuit embrace proof that the defendant had a duty to the plaintiff, that the duty was breached, that this breach of duty caused the injury to the plaintiff and, finally, that there were compensatory damages.

The definition of the duty of the defendant to the plaintiff usually results in allegations that there was lack of compliance with the standard of care. Definition of the standard of care may be a point of contention. In some instances, this can be established by statements in such authoritative sources as the CFR of the FDA or the Standards of the AABB. In this regard, it is important to note that the words, “must” and “shall,” are carefully utilized in the Standards, in contrast to “may” and “should,” to distinguish mandatory practices from suggested practices, respectively. In other situations, expert witnesses are required to define and interpret this issue.

Expert witnesses also are used to determine whether or not the defendant breached the acknowledged duty and also to demonstrate that, but for that failure, the plaintiff's injury would not have occurred. Damages are awarded on the basis of medical costs, loss of earning power and pain and suffering; moreover, they commonly vary between jurisdictions.

Most states have legislation defining blood transfusion as a service, rather than sale of a product, in an attempt to protect blood banks from lawsuits based on the theory of strict liability. In the latter instance the defendant may be held guilty even in the absence of proof of negligence. For example, a blood bank could be liable for a patient acquiring a transfusion-related infection even if there is no test to detect that infection in the donor or in the unit of blood (18).

Consent for Transfusion

The blood donor and the patient receiving the transfusion must be informed of the risks of the procedure, and the patient also must be informed of the benefits and available alternatives. This must be done in a language that can be understood, and the subject must have an opportunity to ask questions and to refuse. Therefore, informed consent actually is a process wherein the subject, once informed, is allowed to make a choice.

In the instance of blood donation, the donor signs a form attesting to understanding of the performance of the venipuncture and removal of the blood or component, and also to his or her provision of truthful answers to the screening procedure. Patient informed consent is somewhat less straightforward. The manner of obtaining and documenting consent are subject to wide variation, and was the subject of a recent symposium (19). Recently, several states have mandated that written consent be obtained before elective transfusion through use of standardized forms.

In deciding how to comply with requirements for obtaining consent for a blood transfusion from an informed patient, one must consider when this should be done, who should be responsible for doing it, and how it should be documented. It is self-evident that the risks of transfusion must be discussed with the patient well in advance of the time of the transfusion if the patient is to have an opportunity to utilize such alternative procedures as autologous transfusion.

The physician responsible for the patient's care will be the one determining the need for the transfusion; therefore, this must also include responsibility for explaining the risks and benefits of the transfusion and detailing alternatives. This is best done at the time the need for a procedure requiring elective transfusion is first discussed with the patient. In most instances this function should not be performed by the hospital transfusion service.

The rationale for this activity is to provide the comprehending patient with an opportunity to understand the risks inherent in a potentially dangerous procedure, and, if appropriate, to take an alternative action. In addition, it allows the physician to demonstrate at a later date, if necessary, that the patient agreed

to accept the risk of the transfusion. Of course, the latter requires some form of documentation of the consent.

Documentation is necessary only to make it easier for the patient's physician to prove at a later date that the dialogue took place. In many instances this is accomplished by documentation on a special form or as a note in the patient's medical record. In some situations a signed form or a progress note in a medical record is considered unworkable or undesirable, and the information is provided to all patients through use of a brochure; however, it is necessary that the brochure advise the patient of a mechanism to ask questions or clarify areas of misunderstanding.

Refusal of Transfusion

From the foregoing discussion, it is evident that a competent patient has the right to refuse transfusion. This issue most often presents in the treatment of Jehovah's Witness patients who are unconscious and hemorrhaging or whose children require transfusion. The objection to transfusion is based on a strict interpretation of the recurrent biblical admonition against “eating” of blood.

In most instances when there is a serious threat to a minor's life if transfusion is not given, the juvenile court will order the transfusion even if this is contrary to the wishes of the parents. The court always will take into consideration the extent of the medical emergency and the extent of risk to the patient if transfusion is not performed. It is far less predictable that a court will order transfusion of an unconscious adult in a similar situation. In this situation, the court will attempt to balance the rights of the patient with the rights of society, may attempt to determine whether the patient expressed any wishes relative to transfusion when mentally competent, as well as the medical need for the procedure.

REFERENCES

1. Leviticus. 17:11,12.

2. Brown HM. The beginnings of intravenous medication. *Ann Med Hist* 1917;1:177–197.

3. Oberman HA. The evolution of blood transfusion. *Univ Mich Med Ctr J* 1967;33:68–74.

4. Blundell J. The after-management of floodings, and on transfusion. *Lancet* 1828;13:673–681.

5. Blundell J. Experiments on the transfusion of blood by the syringe. *Med Chir Trans* 818;9:56–92.

6. Routh CHF. Remarks, statistical and general on transfusion of blood. *Med Times* 1849;20:114–117.

7. Rosenfield RE. The past and future of immunohematology. *Am J Clin Pathol* 1975;64:569–579.

8. Oberman HA. The crossmatch: a brief historical perspective. *Transfusion* 1981;21:645–651.

9. Rous P, Turner JR. The preservation of living red blood cells in vitro. *J Exp Med* 1916;23:219–237.

10. Robertson O. Transfusion with preserved red blood cells. *Br Med J* 1918;1:691–696.

11. Loutit JF, Mollison PL, Young IM. Citric acid-sodium citrate-glucose mixtures for blood storage. *Q J Exp Physiol* 1943;32:183–202.

12. Fantus B. The therapy of the Cook county hospital. *JAMA* 1937;109;128–131.

13. Coombs RRA, Mourant AE, Race RRA. A new test for the detection of weak and “incomplete” Rh agglutinins. *Brit J Exp Path* 1945;26:255–266.

14. Abel JJ, Rowntree LG, Turner BB. Plasma removal with return of corpuscles. *J Pharmacol Exp Ther* 1914;5:625–641.

15. Drake AW, Finkelstein SN, Sapolsky HM. *The American Blood Supply*. Cambridge, MA: MIT Press, 1982.

16. Menitove JE, ed. *Standards for Blood Banks and Transfusion Services*, 19th ed. Arlington, VA: Am Ass Blood Banks, 1999.

17. Belsey R, Greene M, Baer D. Managing liability risk in the office laboratory. *JAMA* 1986;256:1338–1341.

18. Rabkin B, Rabkin MS. Individual and institutional liability for transfusion-acquired diseases. *JAMA* 1986;256:2242–2243.

19. Widmann FK, ed. *Informed Consent for Blood Transfusion*. Arlington, VA: Am Assoc Blood Banks, 1989.

69
Blood Collection

Alfred J. Grindon

DONOR RECRUITMENT

Part of "69 - Blood Collection"

It has been estimated that, of an American population of approximately 270 million people, at least 150 million may be eligible to be blood donors. However, only approximately 8 million of them actually donate, providing an average of 1.6 units of blood every year or a total of 12 million units of blood. Almost none of these donors receives payment because paid donors have been shown to be more likely to be positive for markers for infectious disease and to transmit hepatitis by transfusion, even when appropriate screening tests are performed to eliminate the apparently dangerous units. Despite genome testing for known viral agents, it remains inappropriate for donor recruiters to offer financial incentives (including gifts or raffle tickets) as an inducement to give blood.

The most common approach to blood donor recruitment, community responsibility, assumes that everyone who is able to give blood should do so, so that all who need it can obtain it without an implied recruitment burden on the potential or actual user of blood. Alternatively, some favor an approach using individual responsibility, with each person responsible for providing his or her blood needs, either by accruing blood donation “credits” against which needs can be drawn or by obtaining replacement donations from family and friends after blood has been used.

Of all eligible donors, almost half have given blood at some time in their lives, and half of these, or 25% of the total, have given within the past 4 years (1). These data do not suggest any lack of understanding of the donation process or of the need for blood on the part of the public at large. When those former donors who remain eligible are asked why they no longer give, the two most common responses are the lack of a personal request and the lack of a convenient donation site. Although support of blood donation through the media is important as a reminder, without a convenient site and a personal request, blood donation is unlikely to occur.

The typical successful blood drive involves a blood center recruiter working with a designated individual in a business (or community or school) to develop a campaign that should include (a) an informational campaign that includes a strong statement of support from the person in charge, (b) informational meetings of committees of employees, (c) the personal solicitation by committee members of all other employees before the date of the drive for pledges to donate and for a specific donation appointment time, and (d) a personal reminder to donors the day before or the day of the drive. In addition, those organizations that regularly have successful blood drives will have a program to recognize those who have worked to organize the blood drive as well as the donors themselves. From the donor's perspective, the most convenient site for blood collection is the workplace, utilizing a cafeteria or recreation facility and equipment brought by truck and set up by blood center staff. With smaller businesses that have insufficient space, a self-contained bloodmobile, typically holding four to six donor beds, can be used. Blood collections at blood centers or other fixed sites often depend on the recruitment of individual donors by telephone.

Particular characteristics are desirable for any blood donation site: The physical environment must be pleasant, clean, and uncluttered; furniture should have no sharp edges and be capable of thorough cleaning and decontamination. Adequate space must be provided for auditory and visual privacy in donor interviews.

Because it is easier to retain a regular blood donor than to recruit a new one, a positive donation experience is critical. Interpersonal skills on the part of collection staff are as important as technical knowledge. Although the frequency of infectious disease markers is clearly lower in repeat donors, this may be a reflection of prevalence; it is not clear that the incidence of such markers (and, by inference, the risk of transfusion-transmitted disease) is lower in repeat donors.

DONOR SELECTION: HISTORY

Part of "69 - Blood Collection"

Donors are selected by history, physical examination, and laboratory testing to protect themselves and the recipient of their blood. Donor selection criteria are based on the Code of Federal Regulations (CFR) of the U.S. Food and Drug Administration, the Standards for Blood Banks and Transfusion Services of the American Association of Blood Banks (AABB) (Standards) (2), state and local regulations, and internal procedures of organizations such as the American Red Cross.

Donor Safety

Because there is no tangible benefit to blood donation, it is important that the risk of the procedure be kept to an absolute minimum.

Standards require that donors be at least 17 years old.

Although some centers allow donation only to age 66, others have extended the age for donor eligibility to 70 or 75 years of age, with no apparent increase in adverse donor reactions. Donors are asked questions about their current state of health, to determine that they are healthy enough to give blood. Donors with a history of cancer or abnormal bleeding and pregnant women usually are deferred.

Blood donation frequency has been limited to no more than once every 8 weeks to maintain the iron stores of the donor. More frequent donation is acceptable where there is clear benefit to a specific recipient, such as with autologous donation or with programs designed to limit donor exposure by providing several units for a given recipient (typically a newborn infant) from the same donor.

Recipient Safety

Most questions asked of the donor are related to recipient safety, such as those related to a potential risk of transmitting infectious disease. Donors are deferred if they have a history of hepatitis or of a positive test for hepatitis, or if they were the only donor involved in a case of posttransfusion hepatitis. Because hepatitis A does not have a prolonged carrier state and childhood hepatitis is most likely to have been hepatitis A, it is reasonable to accept for donation those prospective donors who have a history of hepatitis before age 11.

Donors who know that they have antibody to hepatitis B surface antigen (HBsAg) are accepted because such individuals are immune but not infectious and may have developed this antibody as a result of vaccination. Donors with known antibody to hepatitis B core antigen (anti-HBc) are deferred because anti-HBc may be present without detectable surface antigen in an infectious donor early in the course of the disease and also because anti-HBc may serve as a marker for the presence of other infectious sexually transmitted diseases, such as non-A, non-B, and non-C hepatitis or retroviral agents that are as yet uncharacterized. Finally, donors are deferred for 12 months from sexual or household exposure to hepatitis, parenteral exposure to blood, incarceration for more than 72 hours, or exposure to tattooing or other potentially unsterile skin puncture.

Donors are deferred if they have engaged in practices known to increase risk of human immunodeficiency virus (HIV) infection, such as intravenous drug use, male homosexual or bisexual activity, or heterosexual activity with people at such increased risk. Donors also are questioned about a positive test for anti-HIV (3). In addition, donors from some parts of the world are deferred from time to time, as new HIV variants, poorly detectable by current tests, are recognized in those areas. Although the donor is asked to read a pamphlet that describes in detail the kinds of exposure leading to deferral, as well as signs and symptoms of acquired immunodeficiency syndrome (AIDS) or AIDS-related conditions, it is preferable to repeat some of this information orally to ensure communication to the illiterate or inattentive donor.

Transfusion malaria has not been seen in the United States with any frequency since the end of the Vietnam War. Nevertheless, it remains a cause of concern, as malaria eradication in endemic areas has become more difficult and drug-resistant malaria more of a challenge. For this reason, special criteria must be used for donors who have traveled to areas considered endemic for malaria by the Centers for Disease Control and Prevention. Such donors are deferred for 1 year if they are asymptomatic. Those with a history of malaria or natives of countries endemic for malaria must be asymptomatic for 3 years before being accepted as blood donors.

Patients with tick-borne parasitic diseases, such as Lyme disease and babesiosis, may have infectious agents in the blood, but infected individuals usually are so sick during such episodes that they would be otherwise ineligible to be blood donors. Transfusion is a significant mechanism of transmission of Chagas' disease (American trypanosomiasis) in Latin American countries, but the disease is not endemic in this country. Only rare cases of transfusion-associated Chagas' disease have been recognized in this country to date (4). Nevertheless, because of the possibility of a chronic carrier state, donors with a history of Lyme disease, babesiosis, or Chagas' disease are deferred indefinitely.

Donors are indefinitely deferred if they have received human growth hormone or dura mater transplants because of reports of transmission of Creutzfeldt-Jakob disease (CJD) from these products. The recognition that bovine spongiform encephalopathy (BSE) and variant CJD (vCJD) are probably caused by the same agent strain and produce the same neuropathological picture, the temporal association in Great Britain between BSE and vCJD, and concern about the theoretical transmissibility of vCJD to humans by transfusion (5) all have led to deferral of donors who have spent time in Great Britain, and to a lesser extent, continental Europe.

Donors recently immunized with live attenuated viral or bacterial preparations are deferred temporarily. Donors taking specific medications thought to be teratogenic in trace amounts are deferred until the medication has cleared: etretinate (Tegison), indefinitely; acitretin (Soriatane), 3 years; and finasteride and isotretinoin (Proscar or Propecia, and Accutane, respectively), for 1 month.

Finally, the donor is asked to indicate that he or she has read and understood the contents of provided informational pamphlets, understands the risks of blood donation, the potential for donor reactions, and consents to the process.

DONOR SELECTION: PHYSICAL EXAMINATION

Part of "69 - Blood Collection"

As part of the donor screening process, donors' vital signs are determined and the arms are examined for evidence of intravenous drug abuse. The following types of donors are deferred: (a) those who are febrile (temperature above 37.5°C), to reduce the potential for transmission of infection; (b) those who are hypertensive (blood pressure greater than 180/100 mm Hg); (c) those who have abnormalities of pulse rate (less than 50/min or more than 100/min) or rhythm indicative of underlying heart disease; and (d) those who are so small (less than 50 kg) that the removal of 450 to 500 mL of blood could lead to more frequent donor reactions than with larger donors.

A hemoglobin or hematocrit determination (or equivalent) is performed on every potential donor to ensure that donation does

not cause or aggravate anemia. Both the CFR and Standards have a minimum hemoglobin requirement of 12.5 g/dL; neither the FDA nor the Standards consider normal variations of hemoglobin levels with sex or altitude.

A sample of blood obtained from the earlobe will provide a value for hemoglobin or hematocrit of up to seven volumes percentage higher (6) than one obtained by fingerstick or venous sampling; donors may be accepted after earlobe sampling who would otherwise be deferred. Neither Standards nor the CFR specifies the sampling technique.

A copper sulfate hemoglobin screening method is frequently used. If a drop of blood has a hemoglobin content of 12.5 g/dL or greater, it will sink in a copper sulfate solution with a specific gravity of 1.053. Although this test is fairly sensitive, it is not very specific, and blood centers often test those who fail the copper sulfate screening test with a more specific test such as a microhematocrit (7). This testing combination is unlikely to accept as a blood donor someone whose hemoglobin is below acceptable levels, nor will it lead to the deferral of those whose hemoglobins are truly satisfactory. Portable spectrophotometers may also be used for hemoglobin testing. All techniques used must be adequately monitored with a quality control program.

Repeated blood donation can lead to depletion of iron stores, particularly in women of childbearing age who may already have marginal iron stores. Some workers have advocated protective measures, such as the use of specific tests for iron deficiency in blood donors, prolonging the acceptable period between blood donations for selected groups, or the use of oral iron supplementation for regular blood donors. Others conclude that although depletion of iron stores is common, morbidity owing to iron deficiency will not occur without the development of anemia. Thus, the hemoglobin screening tests are adequate to prevent disease. Both groups are concerned about the safety of the donor, and both recognize that women of childbearing age are frequently deferred because of anemia. Blood donor iron supplementation may become more common in this country, particularly if oral iron in a form nontoxic to small children becomes available.

COLLECTION PROCESS

Part of "69 - Blood Collection"

Blood collection techniques are designed to ensure the safety and efficacy of the product obtained. These techniques include positive identification of the donor and assigning a unique identifying number for the donation that is placed on the bag, the donor history card, and any pilot samples needed for laboratory testing. After finding a suitable vein, the phlebotomist prepares the skin by extensive cleansing with surgical soap in an area several inches around the site of intended venipuncture. This is followed by the application of an iodine-containing solution (typically PVP iodine, less sensitizing than tincture of iodine) to complete the process of making the skin as aseptic as possible (8). After this skin preparation, the needle is inserted and 450 to 500 mL of blood (less if the donor weighs less than 50 kg) is collected in a bag containing an anticoagulant-preservative solution. During collection, the blood is mixed regularly, either manually or with a mechanical agitation device, to ensure adequate mixing with the anticoagulant and prevention of formation of clots. When the collection is completed, tubing from the donor to the bag is sealed, after which the needle is withdrawn. At all times, attention is paid to the need to prevent anything other than donor blood from entering the preconnected sterile plastic bags.

Pressure is applied directly to the donor's venipuncture site to ensure stoppage of the blood flow, and the bag is prepared for transport to the component preparation laboratory after stripping the blood in the still-attached plastic tubing into the bag. This ensures that this blood has been adequately mixed with anticoagulant. The plastic tubing is then sealed into segments that remain attached and can be used for subsequent testing in the laboratory or in the hospital.

After a few minutes, an appropriate bandage is placed on the arm, the donor sits up for brief observation by the phlebotomist, and then is escorted to a refreshment area. Here the donor is encouraged to drink several glasses of liquid, while under observation for potential reactions. Donors are given verbal or written instructions after donation regarding care of their arms, the need to drink fluids to replace lost blood volume, and permissible activity levels.

PROCESSING OF DONOR BLOOD

Part of "69 - Blood Collection"

Each unit of donor blood must be tested for ABO group and Rh type, and blood from donors with a history of transfusion or pregnancy must be tested for unexpected antibodies. In addition, all donor blood is tested to determine its potential for transmission of infectious diseases. Required tests include a serologic test for syphilis, HBsAg, anti-HBc, anti-HIV-1/2, HIV p24 antigen, anti-human T-cell lymphotrophic virus I/II, and anti-hepatitis C virus (HCV). The results of these tests must be negative before the blood is released for transfusion. Some centers test for alanine aminotransferase, but this is not required. Genome detection by nucleic acid amplification techniques is now performed for HCV and (for most centers) HIV, in addition to traditional antibody detection. Although most needs for cytomegalovirus-safe blood are met by prestorage leukoreduction techniques (9), some clinicians may request blood found negative for anti-cytomegalovirus.

DONOR REACTIONS

Part of "69 - Blood Collection"

Donors may develop insignificant post-phlebotomy hematomas and local pain or paresthesias. Although some of these complications may require a visit to a physician for reassurance, they usually resolve within a brief interval.

A more difficult problem is the vasovagal reaction, seen in 1% to 3% of blood donors. Early symptoms and signs include pallor, nausea, and diaphoresis. Some donors become apprehensive and hyperventilate, producing signs of hypocalcemia. Often these early signs are recognized and the reaction reversed, with cool compresses and leg elevation, or, for hyperventilation, rebreathing of expired air. More severe reactions proceed to a loss of consciousness, associated with a significant drop in systolic blood pressure (rarely as low as 50 mm Hg) and a pulse rate ranging

from 40 to 60 per minute. Signs of severe reactions may include convulsive movements, vomiting, and (rarely) fecal or urinary incontinence. These reactions are frightening but inconsequential in otherwise healthy donors.

Although the primary cause of this vagal stimulation is thought to be psychological, it is associated with the volume of blood removed, the age and donation experience of the donor, and the ambience of the collection facility. Prevention includes assuring that underweight donors are not bled to excess and that additional attention is paid to young and first-time donors or donors otherwise apprehensive because early reactions often can be reversed. It is important to ensure that donors are under observation in the immediate post-donation period so that signs and symptoms are noted before fainting and potential injury occur.

Treatment involves keeping the donor supine, with legs elevated to enhance blood flow to the head. Recovery usually takes 5 to 15 minutes. Although some facilities are prepared to provide intravenous fluids, oxygen, or vagus nerve–blocking medications, these interventions are almost never necessary. The donor who has had a reaction should be advised that these reactions may recur within the next several hours, and extra caution should be used in driving or operating heavy machinery. Donors with severe reactions should be discouraged from attempting to give blood again in the near future.

SPECIAL COLLECTION ISSUES

Part of "69 - Blood Collection"

Blood Safety

Despite varied approaches to exclude those at high risk of infection from the donation process, there is serologic evidence that individuals infected with hepatitis and HIV-1 continue to donate. Continued effort must be made to reduce donation frequency by those at high risk. In addition to adequate privacy for the donor interview and verbal questioning about high-risk activity, accurate donor identification must be obtained for comparison with donors who have been deferred previously and whose identification has been entered into a permanent donor deferral record. Some centers use automated systems to perform such screening before donation. Sites other than the blood center should offer confidential and inexpensive infectious disease testing to minimize donation merely to obtain a test result. Finally, the donor should be given the opportunity to indicate confidentially after donation that his or her blood should not be used for transfusion. Opportunities for such confidential exclusion can be provided at the time of blood donation and by providing a telephone number that the donor may call afterward.

Staff Safety

Collection staff should be protected as much as possible from exposure to blood. Sharp objects must be placed in a container that allows disposal without fear of accidental puncture. Staff who are at risk of exposure to blood should be encouraged to be vaccinated for hepatitis B. Those preparing segments from the plastic tubing of the blood bag need to be protected from breakage and spraying of blood from these segments at the time of sealing.

Universal precautions (including the use of gloves by phlebotomists) are appropriate whenever dealing with patients, whether hospitalized or at the blood center for therapeutic phlebotomy or autologous donation. Collection staff is at minimal risk of infectious disease acquisition from the occasional exposure to blood of volunteer blood donors because there is no increase in frequency in such workers for markers for highly infectious agents such as hepatitis B (10). Nevertheless, gloves should be available for those phlebotomists who wish to use them and should be worn in every situation in which a phlebotomist has broken or abraded skin.

Directed Donations

Fear of infectious disease transmission has prompted many potential recipients of transfusions to seek blood from family and friends rather than from the general blood supply. No data exist to show greater safety of such donations. In fact, because of a higher percentage of first-time donors, directed units are likely to have a higher frequency of markers for infectious diseases. Despite concern that the availability of such blood tends to make the general blood supply less attractive in the eyes of the public, this practice is generally tolerated.

One special situation in which the use of a designated donor has merit is in the attempt to limit the number of donors to whom an individual recipient is exposed. For instance, for a pediatric recipient with limited blood requirements, several donations could be obtained from one donor, thereby limiting the exposure of the recipient to blood from other allogeneic donors. In special situations of this sort, the benefit to the recipient may outweigh the additional risk incurred by the donor with modification of established standards for blood donors, for instance, by decreasing the interval between donations and lowering of the hemoglobin standard.

Some indications for directed donation are prompted by medical need and not patient demand. These include plateletpheresis of siblings to provide products for a bleeding patient found to be refractory to platelets from random donors or the use of maternal platelets for neonatal thrombocytopenia.

Preoperative Autologous Donation

Patients may donate their own blood preoperatively for anticipated transfusion during elective surgery. The ideal candidate for preoperative autologous donation is someone who is otherwise in good health undergoing a procedure that is likely to require transfusion. Patients undergoing orthopedic or plastic surgical procedures often fall into this category.

There is more controversy regarding patients who have a significant systemic disease, such as patients scheduled to undergo coronary artery bypass procedures or patients with cerebrovascular disease. In these situations, a severe donor reaction with hypotension would be especially dangerous because of preexisting compromise of blood flow to critical organs. In fact, reactions leading to hospitalization occur 12 times more frequently (1/17,000) in autologous than in allogeneic donors (11). Nonetheless, such patients are often accepted as autologous donors (with the exception of those with significant aortic stenosis,

transient ischemic attacks, and active angina pectoris). Because the risk of disease transmission from allogeneic blood transfusion is constantly decreasing, it is difficult to justify significant risk to these patients.

Patients scheduled for operations in which blood is not expected to be used are not candidates for autologous donation. Autologous donation in pregnant women is particularly controversial. There is little benefit to those in whom bleeding at the time of delivery is not expected; moreover, there is a theoretical risk related to potential compromise of the fetal blood supply.

Preoperative autologous blood was 5% of the blood collected in 1997, after reaching a peak of 8% in 1992 (12). This reduction in autologous collections stems in part from the recognition of the general safety of the blood supply, and therefore this expensive procedure may be cost-effective only with younger patients undergoing operative procedures requiring several units of blood. Because blood from autologous donors has a higher frequency of markers for infectious diseases, it is not subsequently made available for allogeneic transfusion if not used by the autologous recipient.

When preoperative autologous transfusion is indicated, it is acceptable to modify the standards set for volunteer blood donors: the increased benefit to the patient justifies some modest increase in risk to the patient as a donor, such as a minimum hemoglobin of 11 g/dL rather than 12.5 g/dL and allowance of repeated donation at short intervals.

BLOOD COMPONENT PREPARATION AND STORAGE

Part of "69 - Blood Collection"

General Principles

The desirability of providing patients with those elements of blood that are needed and, to the extent possible, limiting exposure to unneeded components was recognized long before it was generally possible to practice such an approach. The widespread use of plastic bags in the late 1950s allowed centrifugal separation of the formed elements of blood in a closed, presterilized system without bacterial contamination. As a consequence, the often fatal septicemia resulting from introduction of bacteria into blood and blood components at the time of preparation or further manipulation has become rare. Sterile “docking” devices provide additional flexibility, allowing the coupling of two sterile plastic containers that were not connected at the time of manufacture. If, during component preparation, the system is opened aseptically without the use of a sterile docking device, the components can be stored for no longer than 24 hours at 1 to 6°C or 4 hours at room temperature.

Red Blood Cells

Cell Aging

Red cells in the body have a life span of approximately 120 days. Although minimal destruction of cells of all ages occurs normally, by far the more common pattern is the destruction of senescent red cells. This process is related in part to a loss of lipid in the red cell membrane and a loss of the ability of the membrane to maintain its shape. When red blood cells (RBCs) are stored *in vitro*, the process is associated with depletion of adenosine 5′-triphosphate (ATP). Length of storage of RBCs is related to the ability of the cells to maintain adequate levels of ATP. Energy is needed not only to drive the sodium-potassium pump but also to support the shape of the RBC and to prevent loss of the lipid from the membrane. Ideal storage *in vitro* would provide RBCs with a posttransfusion survival equal to, or better than, cells of a similar age remaining in the body. Because blood drawn from a normal donor contains RBCs with an average age of 60 days, it would be desirable to store those cells in *vitro* so that at the end of 30 days, or another 25% of the RBCs' normal life span, no more than the oldest 25% of the RBCs would die. In fact, it is possible to store RBCs for longer periods of time with greater survival of the remaining cells.

Measurement of RBC Preservation

Transfused RBCs circulating 24 hours after transfusion have a survival that is comparable with normal RBCs. It is therefore possible to use 24-hour posttransfusion recovery as an index of posttransfusion survival. Seventy percent of the RBCs transfused after a period of storage should survive in the circulation 24 hours after transfusion. Because cells from one individual have post-storage survival characteristics different from those of another, the use of a 75% 24-hour survival of transfused RBCs will ensure that at the end of the allowed storage period, most units of blood will have adequate recovery and survival despite individual variation.

In the evaluation of anticoagulant-preservative solutions, posttransfusion recovery and survival are currently assessed by the transfusion of radiolabeled autologous RBC. Although technically more difficult, it is probably important to utilize a double-radioisotope technique. The loss of some RBCs immediately after transfusion may lead to falsely high levels of recovery if a single-label technique is used.

Red Cell Function

The primary role of the RBC is the delivery of oxygen to tissues. Storage of RBCs *in vitro* leads to the depletion of 2,3-diphosphoglycerate (2,3-DPG), important for the release of oxygen from hemoglobin. Although this substance is regenerated within 24 hours after transfusion in storage-depleted RBCs, it would be ideal to maintain 2,3-DPG levels during storage. Rats, resuscitated after hemorrhage to half-normal hematocrits with 2,3-DPG–depleted blood, have compromised survival compared with those resuscitated with 2,3-DPG replete blood (13). However, it has not yet been possible to demonstrate in the human that absence of 2,3-DPG from transfused RBCs, even when such cells are given in massive amounts, has been associated with adverse effects, perhaps because of other compensatory mechanisms such as the Bohr effect.

Transfusion of blood toward or after the end of the allowable storage period has no apparent ill effect for the recipient, aside from a posttransfusion increase in hemoglobin catabolites, such as bilirubin and lactate dehydrogenase.

Anticoagulant-Preservative Solutions

Citrate has long been the anticoagulant of choice. It is readily catabolized in the body on transfusion. By chelating calcium, it inhibits several calcium-dependent steps of the coagulation sequence. It is commonly used in a sodium citrate/citric acid buffer and is provided in substantial molar excess to calcium. Anticoagulants such as ethylenediaminetetraacetic acid and heparin have been used from time to time but have deleterious effects on stored RBC over all but the shortest time span.

Dextrose was added to anticoagulant solutions because of the empirical recognition that it improved posttransfusion RBC recovery and survival. Buffering of the anticoagulant-sugar solution to a lower pH allowed heat sterilization without caramelization of the sugar. With greater awareness of the role of ATP and 2,3-DPG in RBC preservation and function, preservative solutions have been fashioned to optimize amounts of these materials in RBC during storage.

Standard anticoagulant-preservative solutions have long included a sodium citrate/citric acid buffer, adjusted to provide an initial pH of the stored blood of 7.5 to 7.6 (at 37°C), and dextrose. Although inorganic phosphate has a minimal effect on posttransfusion recovery, it has been shown to enhance maintenance of 2,3-DPG and ATP. Adenine prolongs RBC storage *in vitro* by serving as a substrate for synthesis of ATP. These elements provide the basis for a commonly used solution, CPDA-1, which contains a citrate buffer with phosphate, dextrose, and adenine.

Currently, collection container techniques allow increased amounts of adenine to be added selectively to the red blood cells, after removal of plasma or platelets, to provide still better RBC storage. The amount of adenine added is ultimately limited by the potential for precipitation of adenine metabolites in the kidneys after massive transfusion. Some adenine-enriched solutions contain a small amount of mannitol to reduce hemolysis after extended storage. Because these adenine solutions allow the separation and recovery of larger amounts of plasma and allow RBC to be stored for a longer period, they have become the favored anticoagulant-preservative solutions.

RBCs can be prepared simply by sedimentation, followed by separation of the supernatant plasma into an attached container. Most facilities that collect blood in substantial amounts separate platelets, fresh frozen plasma, or both, from the whole blood by centrifugation techniques soon after collection, leaving either CPDA-1 RBCs with a volume of 250 mL, a hematocrit of 70% to 80%, and 35-day dating or adenine-enriched (AS-1, AS-3, or AS-5) RBCs, with a volume of 350 mL, a hematocrit of 50% to 60%, and 42-day dating. Storage of RBC *in vitro* causes changes in plasma pH and potassium and RBC ATP and 2,3-DPG. The nature of these changes is shown in Table 69.1.

TABLE 69.1. BIOCHEMICAL CHANGES IN BLOOD STORED AT 5°C

	CPDA-1		AS-1[a]	
	WB	**WB**	**RBCs**	**RBCs**
Days stored	0	35	35	42
% Viable cells (24-hr posttransfusion)	95–100	79	71	80
pH (at 37°C)	7.6	6.98	6.71	6.6
ATP%	100	56	45	64
2,3-DPG%	100	<10	<10	<5
Plasma K^+ mmol/L	4.2	27.3	78.5	50

[a]As an example of adenine-enriched preservatives; AS-3 and AS-5 are not substantially different.

WB, whole blood; RBCs, red blood cells.

Modified from Vengelen-Tyler, V, ed. *Technical manual*, 13th ed. Bethesda, MD: American Association of Blood Banks, 1999.

Rejuvenation

RBCs severely depleted of ATP and 2,3-DPG during storage can regenerate these substances with the addition of “rejuvenating” solutions, such as one containing pyruvate, inosine, phosphate, glucose, and adenine. These solutions are potentially toxic and cannot be transfused, so after rejuvenation, the RBCs are washed and transfused or glycerolized and stored in the frozen state for subsequent transfusion.

pH

Because of continuing slow RBC metabolism of glucose to lactate during storage, the accumulation of hydrogen ions in the plasma lowers the pH over time. RBCs stored for 35 days in CPDA-1 may have a pH as low as 6.7 (37°C). There is no evidence that the massive transfusion of blood with a pH at this level has any deleterious clinical effect because of numerous *in vivo* mechanisms to compensate for acidosis.

Potassium

Potassium is gradually lost from RBCs during storage and increases in the plasma overlying whole blood at a rate of approximately 1 mEq/day. The potassium concentration is higher when accumulated in smaller amounts of plasma, such as with RBCs, in either CPDA-1 or adenine-supplemented preservative solutions. The infusion of RBCs containing potassium at these high concentrations (as much as 80 mEq/L toward the end of the storage time) is of no consequence in most clinical situations, including the transfusion of massive amounts of stored blood. There are concerns, however, about potassium levels in blood used for exchange transfusion of neonatal patients and for blood given to patients with renal failure. The concern is focused on the rate of infusion of hyperkalemic plasma rather than the amount of potassium infused, which is minimal in relationship to the total body potassium. Further, potassium-depleted RBCs will act as a sponge and reabsorb plasma potassium in hours. In these situations, provision of blood that is relatively fresh (stored for less than 7 days) provides levels of potassium that although elevated are safe for clinical use.

Washed RBCs

RBC washing can remove most (99%) of the plasma, so it can be used for the treatment of patients for whom some plasma components can be deleterious. For instance, some clinicians treat

patients with paroxysmal nocturnal hemoglobinuria with washed RBCs to prevent the infusion of complement components. Washing RBCs may be sufficient to prevent reactions in patients who lack IgA and who have had reactions to blood components containing these immunoglobulins. In a few cases, however, this procedure has been insufficient to prevent such reactions, and products from donors known to be IgA deficient are preferable. Finally, washing has been used to remove plasma containing elevated levels of plasma potassium, where this is thought to represent a hazard to neonatal patients.

Frozen RBCs

RBCs can be stored for 10 years in the frozen state. It is possible that little additional loss of posttransfusion recoverability is seen after several decades. Such storage requires either extremely rapid freezing of RBCs and maintenance at very low temperatures (obtained with liquid nitrogen freezers that maintain a temperature below -120°C), the addition of large amounts of cryoprotective agent, or some combination of these approaches (14). The inconvenience of transport of RBCs stored in liquid nitrogen has led to the widespread use of 40% wt/vol glycerol as an intracellular cryoprotective agent. The glycerol must be added to the RBCs slowly with mixing to minimize hemolysis. Glycerolized RBCs may be stored with mechanical refrigeration at below -65°C for as long as 10 years and may be transported in dry ice. After thawing, most of the glycerol in the RBCs must be removed to prevent posttransfusion osmotic hemolysis. Removal of glycerol is commonly performed by washing the thawed RBCs first with a hypertonic saline solution, followed by resuspension in normal saline. RBC recovery from a freeze/thaw/wash process must be at least 80%, and the RBC viability at 24 hours posttransfusion must be at least 70%. Because of cost, this technique is not widely used, but it is valuable for storage of RBCs of rare phenotype or for autologous transfusion situations when extended storage is needed.

Leukocyte Reduction

Background

It has been recognized for many years that transfused “passenger” leukocytes can cause deleterious effects in recipients of transfused RBCs. Even crude techniques such as removing a 50-mL buffy coat containing as many as 80% of the leukocytes of a unit were effective in preventing many transfusion reactions. For more complete leukocyte removal, RBCs were washed, in some cases with red cell rouleauxing agents to facilitate separation. A more sophisticated technique involved the filtration of stored, centrifuged blood through a microaggregate filter. This filter was designed to remove the aggregates of leukocyte and platelet debris found in RBCs on storage. Although widely used, this filter probably provided little beneficial effect except in patients on cardiopulmonary bypass because the lungs removed such microaggregates without demonstrable untoward effect.

Leukoreduction Filters

Current leukoreduction filters are effective in removing leukocytes, initially present in RBCs at a level of approximately 10e9, to a level of less than 5 × 10e6, with recovery of at least 85% of the red cells.

Platelets can also be filtered to a level such that a pool of six will contain 5 × 10e6 leukocytes but still meet Standards for platelet content. More commonly, leukoreduced platelet components are readily produced by newer apheresis processes to the same leukoreduction standards. Fresh frozen plasma prepared from platelet-poor plasma may contain sufficient leukocytes to warrant consideration of leukoreduction.

Leukoreduction to less than 5 × 10e6 of red cells and platelets will have several beneficial effects (15). Febrile transfusion reactions caused by infusion of leukocytes or cytokines are significantly reduced. This reduction is seen most clearly in patients who have received multiple platelet transfusions. Alloimmunization to platelets, manifested by the development of lymphocytotoxins and a platelet refractory state, is also significantly reduced. Disease transmission by agents found in leukocytes, such as cytomegalovirus and human T-cell lymphotrophic virus (and other herpes viruses, trypanosomes and toxoplasma, by inference) are reduced. Experimental evidence suggests that transfusion-transmitted bacterial infection should be reduced as well. Although leukoreduction is advocated in some countries to reduce the risk of vCJD, this is a hypothetical benefit at present.

The transfusion of allogeneic leukocytes has been known for years to have an immunomodulatory effect: such transfusions improve renal allograft survival and reduce the frequency of spontaneous abortion. There are increasing data, both clinical and in experimental animals, showing that allogeneic leukocytes cause immunomodulation that is manifested by postoperative infection (and perhaps by cancer recurrence) and that leukoreduction may reduce such infection and thereby shorten hospital stay.

Leukoreduction is preferably carried out in the blood center (prestorage) rather than at the bedside because leukoreduction is more consistently complete, with quality and process controls, and with better removal of white-cell fragments and reduced opportunity for cytokine development. Because leukofiltration can itself cause transient activation of cytokines, bedside leukoreduction has caused severe hypotensive reactions (16).

The use of current leukoreduction techniques has virtually eliminated use of older white-cell removal methods, including the use of microaggregate filters and washing of RBCs for leukoreduction.

Platelets

Platelets are the lightest of the formed elements of the blood. After slow centrifugation of whole-blood, platelet-rich plasma can be found at the top and can be readily separated. Subsequent hard centrifuging of this plasma will sediment the concentrated platelets, and the platelet-poor plasma can then be removed. Collection of blood into a closed system containing several preattached plastic bags allows this manipulation without opening the bag system; therefore, platelets in concentrated form can be prepared and stored at room temperature for several days after collection without fear of growth of bacteria introduced at the time of preparation.

The techniques used to provide consistently good platelet concentrates begin with collection of blood from the donor.

Blood collection should be fairly rapid and uninterrupted. The collection of blood over a prolonged period of time, or with interruptions, could lead to the initiation of coagulation on a scale too small to be detected but sufficient to cause the initiation of platelet activation. After collection, the temperature of whole blood is maintained above 20°C until platelets are prepared.

In the laboratory, optimal platelet recovery is best achieved by using a force as low as reasonably possible for the centrifugal preparation of the platelet-rich plasma. Although yield can be slightly increased by extending down into the lymphocyte-rich top layer of buffy coat, this usually is unnecessary. Furthermore, the increase in white cells from such a procedure adds to the metabolic storage burden and to a greater possibility of sensitization (unless white cells are removed pre–storage). The platelet-rich plasma is then recentrifuged to concentrate the platelets. This spin should be as hard as necessary to provide adequate concentration but not so hard that platelets are damaged in the process. The platelet-poor plasma is then expressed, leaving a final product volume of between 50 and 70 mL, depending on the type of container used.

Platelets may be stored for 5 days at room temperature by meeting critical storage requirements in areas of pH, plastic, and agitation. The pH of the platelet concentrate is inclined to fall as a result of normal metabolic activity during storage. Nevertheless, it must be maintained at levels at or above 6.2 to preserve posttransfusion function. Plastics with improved gas permeability can allow development of an abnormally elevated pH; this may also be deleterious to platelets.

The type of container used for platelet storage has special requirements. Polyvinyl chloride bags have in their formulation a plasticizer to provide flexibility. A common plasticizer, diethylhexylphthallate (DEHP), is leached from the polyvinyl into lipid-containing plasma, where it has a protective effect on red cell membranes, facilitating extended storage of RBCs. Because of the remote possibility of toxicity from the transfusion of this material and because of prolonged storage of DEHP in the body, manufacturers have sought formulations that would minimize its content (17). The plastics most often used for storage of platelets consist of polyolefin, which does not require a plasticizer, thinner polyvinyl with reduced DEHP, or polyvinyl with a different plasticizer, butyrye–trihexylcitrate, which is readily catabolized. Such approaches could reduce the current theoretical concerns related to the infusion of DEHP. In general, plastic designed for platelet storage provides better exchange of CO_2 and O_2 across the bag and therefore minimizes the reduction in pH in the stored platelet concentrate.

Another constant requirement for prolonged storage of platelet concentrates is agitation. Immediately after removal from the centrifuge, the concentrate should be allowed to stand undisturbed for an hour or two. After this, the component is placed on a mechanical device for continuous agitation. Some types of agitation are particularly harsh and provide poor platelet recovery. The common types currently utilized are to-and-fro, rotary, and elliptical agitation with the bag laid flat on the rotator. The most desirable type of rotator varies to some extent with the type of bag used (18). The temperature must be maintained between 20°C and 24°C, during storage.

Plasma at room temperature is a bacterial growth medium. Donors must not be bacteremic at the time of donation and platelets cannot be pooled in an open system to provide a typical dose for transfusion and then subsequently stored at room temperature. Even with these precautions, a few bacteria present at the time of collection may proliferate to unacceptable levels in platelets stored at room temperature. As a result, there is a maximum storage time of 5 days, even though survival characteristics of the platelets would permit their storage for longer periods with some containers.

Before platelets could be stored for 5 days, there was a considerable interest in their storage in the frozen state. The traditional and still preferred technique uses dimethyl sulfoxide (DMSO) rather than glycerol as a cryoprotective agent. However, platelets are more fragile than RBCs, and the recovery of platelets in the recipient after freezing and thawing is poor. Storage of platelets in the frozen state has little use today except when autologous platelets are obtained from patients with leukemia in remission and stored in the frozen state to provide support at the time of the next therapeutic induction. This may be especially helpful when patients are refractory to homologous platelets.

A small number of units of platelets prepared every month must be tested at the end of their storage time to be certain that a pH of at least 6.2 is maintained in all units tested and to confirm that the number of platelets present exceeds 5.5×10^{10} in 75% of the units tested. Some platelet preparation techniques result in the presence of substantial numbers of RBCs in the final product, giving it a pink tinge. The presence of more than 0.5 mL of RBCs is thought to be excessive. Rh-negative patients who are given repeated transfusions of platelets from Rh-positive donors with this level of RBC contamination will develop Rh antibodies, although this occurs in only 8% of Rh–negative patients. However, it is not difficult to produce products with much lower RBC contamination. Furthermore, if Rh-positive RBC-containing platelets must be given to Rh-negative recipients, Rh human immunoglobulin can be administered within 72 hours of the transfusion to prevent sensitization.

Fresh Frozen Plasma

The primary concern in the preparation of fresh frozen plasma is that it be frozen quickly, before deterioration of labile clotting factors. For this reason, the time of collection should be noted. The plasma must be placed in the freezer within 8 hours of this time and should be frozen solid in the next 2 hours (within 10 hours of the time of collection of the whole blood). Because the time to preparation of a frozen product and the time of freezing are important, rapid freezing techniques have been used. Plasma may be immersed in an ethanol/dry ice bath or other liquid heat exchange mixture or placed in a freezer designed for the rapid freezing of 200 to 250 mL of plasma. When frozen, there is little deterioration of labile clotting factors in fresh frozen plasma, so that it can be stored for a year below -18°C without appreciable loss of activity. Plasma frozen within 24 hours of collection has similar amounts of labile clotting factors (except factor VIII) and can be used in place of fresh frozen plasma for most purposes.

To minimize transfusion-transmitted disease, fresh frozen plasma is being replaced for some patients with products less likely to transmit infectious disease. Two such products are available

in this country. Pooled plasma, solvent/detergent-treated has been treated by a process that inactivates all lipid-encapsulated viruses, including HIV and hepatitis B virus. It has been shown to be safe and effective in treating conditions for which fresh frozen plasma would be used. Because of a theoretically increased risk of transmission of nonlipid-encapsulated agents resulting from the pooling process, other single-unit approaches to reducing the risk of transfusion-transmitted disease have been used. These include the European approaches of methylene blue, ultraviolet light, or pasteurization, and in the United States, donor retested fresh frozen plasma that has been stored until the donor can be retested after the window period of the infectious agents of concern (19). Plasma components are still used inappropriately for volume expansion. The risk of disease transmission, compared with pasturized albumin or plasma protein fraction, makes it undesirable for this purpose.

Cryoprecipitated Antihemophilic Factor

When fresh frozen plasma is thawed at 5°C, some material does not go into solution. The remaining precipitate contains much of the factor VIII [antihemophilic factor (AHF)] and fibrinogen, some of the factor XIII, and large multimers of von Willebrand factor from the original unit. The cold precipitation property of these proteins is important in the further manufacture of factor VIII preparations for the treatment of hemophilia. Cryoprecipitate can be separated from the precipitate-poor plasma by centrifugation of the plasma containing the cryoprecipitate. After centrifugation, the precipitate will adhere to the plastic bag sufficiently to allow transfer of the precipitate-poor plasma into another container. Between 7 and 15 mL plasma is usually left in the bag. Each cryoprecipitate is then refrozen, thawed when needed, and pooled for transfusion. In some facilities, small amounts of saline are added to the cryoprecipitate to facilitate its recovery. At least four units are assayed each month to determine factor VIII activity. Units tested must contain at least 80 units of AHF activity (compared with the approximately 200 to 250 units present in the original fresh frozen plasma), and at least 150 mg of fibrinogen.

Cryoprecipitate poor plasma (plasma from which the cryoprecipitate has been removed) has been shown to be useful for the treatment of thrombotic thrombocytopenic purpura, especially in patients unresponsive to therapeutic plasma exchange with fresh frozen plasma. The cryoprecipitation process removes from the plasma the large von Willebrand factor multimers that contribute to the pathophysiology of the disease, although the plasma transfusion allows the replacement of the metalloproteinase responsible for the cleavage of these multimers *in vivo* (20).

CYTAPHERESIS PRODUCT PREPARATION AND STORAGE

Part of "69 - Blood Collection"

Principles

Apheresis devices are available that allow the collection of an adequate dose of platelets or granulocytes from a donor at one setting. The most commonly used devices utilize a continuous flow of donor blood, using principles developed in the dairy industry for separation of cream.

The donor typically is connected to a machine by a single or double needle for a period of 1 to 2 hours, during which time blood is continuously collected, the desired component is separated by centrifugal techniques, and the unneeded components are returned to the donor through a second venous line. Both continuous and discontinuous flow devices may be used for the procedure. This process is carried out in a closed system; platelets collected by such techniques can be stored with agitation for 5 days at 20° to 24°C.

Donor Selection

The historical, physical, and laboratory criteria used for the selection of apheresis donors are similar to those of whole blood donors. In addition, there are some special concerns. Although these donors can be drawn as often as every 2 days (but no more than 24 times a year), the cumulative loss of RBCs must be less than 200 mL every 8 weeks and of plasma, less than 1 L (1.2 L for persons weighing more than 80 kg) every 7 days, with a maximum of 15 L in 12 months. If formed elements of the blood are removed on a frequent basis, the donor's safety must be protected by measuring levels of these formed elements during the course of a series of donations. Because two large veins are required for most procedures, the adequacy of venous access is important.

Plateletpheresis donors should not have taken aspirin-containing medication within the preceding 36 hours because aspirin blocks the generation of thromboxane A_2. This blockage inhibits the platelet-release reaction and renders the donated platelets less effective, particularly when the donor is the sole source of platelets for the recipient. The granulocytapheresis process is usually facilitated by the use of special agents given donors; these are considered below.

Untoward Consequences of Cytapheresis

Lymphocytes have a density similar to that of heavier platelets, and intensive plateletpheresis will remove large numbers of lymphocytes. This will deplete long-lived T lymphocytes and, in theory, could lead over time to a weakening of immune function. However, although reduction in lymphocytes and immunoglobulin levels has been noted (21), no clinical change in immune response has been seen. The removal of large numbers of platelets in a plateletpheresis procedure causes a drop in the peripheral blood platelet count of perhaps 30,000/μL. It is important, therefore, that the initial platelet count be adequate to support such a loss. For sequential donors, the postprocedure count from one plateletpheresis can be used as the initial count for the subsequent procedure.

The most common donor reaction during and after a cytapheresis procedure is the reaction to infusion of the anticoagulant sodium citrate, which can produce symptoms of hypocalcemia. Therefore, it is desirable to reduce the amount of infused citrate as much as possible. When such a reaction occurs, it is treated by slowing the infusion rate or providing oral calcium-containing antacids.

There are special concerns for donors asked to provide granulocytes collected by apheresis. To collect adequate numbers of

granulocytes, donors must be given a rouleaux-inducing agent, such as hydroxyethyl starch (HES). Because this material has been reported to cause both anaphylactic reactions and persistent localized skin reactions, donors must be asked about related allergies. In addition, products of this type are colloid volume expanders. Consequently, when given over several days, they can cause localized edema, headache, and other symptoms of fluid retention. The total product dose given to a donor should be monitored; it may be possible to reduce the dose for a given donor from one procedure to the next. A donor given a corticosteroid to improve granulocyte collection, particularly when the donor is used repeatedly over a short period of time, must be asked about underlying diseases that might be exacerbated by corticosteroids, such as peptic ulcer disease and tuberculosis. Granulocyte donors have also been given granulocyte colony–stimulating factor (G-CSF). This product improves granulocyte yield dramatically (22). However, it causes untoward effects: most donors given several doses have bone pain and flulike symptoms, amenable to analgesic treatment. Donors receiving G-CSF and subjected to large-volume apheresis for peripheral blood progenitor cell retrieval may develop significant thrombocytopenia (less than 100,000/μL). There is also a theoretical risk with G-CSF of the agent's stimulating neoplastic transformation or mobilizing such neoplastic forms.

There have been isolated reports of donor reactions related to the equipment. Kinks in tubing have been reported to cause hemolysis. Faulty rotating seals have caused particulate carbon to be found in collection containers (but without apparent ill effect for the donor). Cytapheresis devices generally have bubble detectors to guard against dangerous reactions resulting from air embolization during reinfusion.

Plateletpheresis

Indications

Patients refractory to random platelets are often responsive to platelets from donors matched at the HLA-A and -B loci or shown to be serologically compatible by cross-match techniques. Adequate amounts of such platelets from a single donor can be provided only with cytapheresis techniques. Furthermore, use of cytapheresis platelets reduces the number of donors to whom a recipient is exposed with each dose of platelets from six to one, producing a similar reduction in risk of transfusion-transmitted disease and in the potential for febrile transfusion reactions.

Most apheresis platelets are produced routinely to have fewer than 5 × 10e6 leukocytes. These leukoreduced apheresis platelets produce the beneficial effects of leukoreduction, including decreased febrile transfusion reactions, decreased alloimmunization, reduced disease transmission (particularly of cytomegalovirus), and a reduction in the effects of immunomodulation caused by transfusion of allogeneic leukocytes.

Collection Techniques

Platelet collection times vary from 1 to 2 hours. Most current procedures involve the use of a needle in each arm, although both intermittent and continuous techniques are available that use one arm. Each technique should be able to produce a product containing 3 $\times 10^{11}$ platelets in at least 75% of the units tested. Many of the products contain enough platelets to allow splitting into two doses. To be labeled leukoreduced, a plateletpheresis product must contain fewer than 5 × 10^{6} leukocytes.

Granulocytes Pheresis

Granulocytes are collected by procedures utilizing continuous flow techniques with differential centrifugation. Because granulocytes have densities only slightly less than RBCs, it has been difficult to provide a clean centrifugal separation. Therefore, it is customary to utilize a rouleaux-inducing agent, such as HES, to enhance the sedimentation of RBCs with centrifugal force. Even with the use of such an agent, it is difficult to obtain doses of granulocytes that meet standards without pretreatment of the cytapheresis donor with corticosteroids to increase the numbers of circulating granulocytes by mobilizing them from the bone marrow and by increasing the number in the circulating pool. Typical protocols include giving the corticosteroid in at least two doses, for instance, at 12 and 3 hours before beginning the cytapheresis procedure. In addition, donors may be given G-CSF, which induces leukocytosis in the donor and improves granulocyte recovery dramatically. Use of this approach, particularly if coupled with a rouleauxing agent and corticosteroid, produces levels of granulocytes that produce readily detectable increments in recipients. Standards requires 1 × 10^{10} granulocytes in 75% of the products tested; with HES, G-CSF, and corticosteroid levels of 7 to 8 × 10^{10} are readily achieved (23). Although this product may be lifesaving, its indications and clinical efficacy are still being defined.

Granulocytes may be stored at room temperature for as long as 24 hours after collection; however, they should be transfused as soon as possible.

Peripheral Blood Progenitor Cells

Modest modification of the granulocyte collection procedure allows collection of peripheral blood progenitor cells in donors pretreated with G-CSF. Typically, a donor selected as a candidate for autologous or allogeneic peripheral blood progenitor cell transplantation is premedicated for 5 days with G-CSF and then undergoes one or two large-volume (15 to 25 L processed) apheresis procedures (22). The donor/patient is subject to the side effects of G-CSF but avoids the bone marrow retrieval procedure.

Erythrocytopheresis

Apheresis technology can be used to obtain RBCs. Typically, two units of RBCs are collected at one 45- to 60-minute session but with the total volume removed not exceeding 571 mL. This procedure requires special considerations for donor safety (minimal hematocrit, 40; weight, more than 130 lb for males and 150 lb for females; donation, no more frequently than every 112 days) but allows collection of two units of difficult-to-obtain blood groups or from autologous donors, with one rather than two donor visits (24).

Therapeutic Cytapheresis

Cytapheresis techniques may also be used therapeutically to reduce dangerously elevated levels of leukocytes (greater than 100,000/µL, especially in acute myelogenous leukemia) or platelets (greater than 1,000,000/µl), before or as an adjunct to chemotherapy (25). In addition, RBC exchange by cytapheresis may be indicated for some patients with sickle cell anemia.

PLASMAPHERESIS

Part of "69 - Blood Collection"

Plasma can be obtained from donors by automated apheresis techniques to provide two units of plasma for preparation of fresh frozen plasma or to provide plasma for the manufacture of blood derivatives. A two-unit withdrawal can be performed as often as twice weekly on an ongoing basis, with the removal of as much as 1 L of plasma per week (1.2 L for persons weighing more than 80 kg). Individuals who donate plasma more frequently than every 8 weeks must have predonation serum protein determinations and periodic determinations of immunoglobulin levels.

Therapeutic Plasma Exchange

Automated plasmapheresis equipment is used for the therapeutic removal of plasma, using albumin and saline as replacement fluids (plasma products are used for patients with thrombotic thrombocytopenic purpura). With adequate venous access, a one-plasma volume (approximately 3 L) exchange can be performed in 90 minutes. This technique has been shown to be effective for patients with such conditions such as thrombotic thrombocytopenic purpura, Guillain-Barré syndrome, myasthenia gravis, and hyperviscosity syndrome (26).

REFERENCES

1. Drake AW. *Public attitudes and decision processes with regard to blood donation: final report and executive summary*. Springfield, VA: National Technical Information Service, 1978.

2. Menitove JE, ed. *Standards for blood banks and transfusion services*, 19th ed. Bethesda, MD: American Association of Blood Banks, 2000;20.

3. U.S. Food and Drug Administration. *Memorandum, April 23, 1992: Revised recommendations for the prevention of HIV transmission by blood and blood products*. Washington, DC: U.S. Food and Drug Administration, 1992.

4. Shulman IA. Intervention strategies to reduce the risk of transfusion transmitted *Trypanosoma cruzi* infection in the United States. *Transfus Med Rev* 1999;13:227–234.

5. Collinge J. Variant Creutzfeldt-Jakob disease. *Lancet* 1999;354:317–323.

6. Coburn TJ, Miller WV, Parrill WD. Unacceptable variability of hemoglobin estimation on samples obtained from ear punctures. *Transfusion* 1977;17:265–268.

7. Cable RG. Hemoglobin determination in blood donors. *Transfus Med Rev* 1995;9:131–144.

8. Vengelen-Tyler V, ed. *Technical manual*, 13th ed. Bethesda, MD: American Association of Blood Banks, 1999.

9. Smith DM, Lipton KS. Leukocyte reduction for the prevention of transfusion-transmitted cytomegalovirus (TT-CMV) (Association Bulletin 97-2). Bethesda, MD: American Association of Blood Banks, 1997.

10. Page PL. Risk of hepatitis B exposure in regional blood services. *Transfusion* 1987;27:242–244.

11. Popovsky MA, Whitaker B, Arnold NL. Severe outcomes of allogeneic and autologous blood donation: frequency and characterization. *Transfusion* 1995;35: 734–737.

12. *Comprehensive report on blood collection and transfusion. Executive summary*. Bethesda, MD: National Blood Data Resource Center, 1999.

13. Collins JA. The age and hematocrit of stored blood in determining the survival of rats after exchange transfusion and hemorrhage. In: Brewer GJ, ed. *Proceedings of the 4th International Conference on Red Cell Metabolism and Function, Ann Arbor, Michigan, September 14-17, 1977*. New York: Alan R. Liss, 1978.

14. Meryman HT. Frozen red cells. *Transfus Med Rev* 1989;3:121–127.

15. Blajchman MA. Transfusion-associated immunomodulation and universal white cell reduction: are we putting the cart before the horse? *Transfusion* 1999;39:665–670.

16. Zoon KC, Jacobson ED, Woodcock J. *Hypotension and bedside leukocyte reduction filters*. Washington, DC: Center for Devices and Radiological Health, U.S. Food and Drug Administration, 1999.

17. Rubin RJ, Ness PM. What price progress? An update on vinyl plastic blood bags. *Transfusion* 1989;29:358–361.

18. Cummings E. Platelet rotaters, infusion pumps and blood warmers. *Transfus Sci* 1989;10:199–206.

19. Bianco C. Choice of human plasma preparations for transfusion. *Transfus Med Rev* 1999;13:84–88.

20. Moake JL. Moschcowitz, multimers, and metalloprotease. *N Engl J Med* 1998;339:1629–1631.

21. Matsui Y, Martin-Alosco S, Doenges E, et al. Effects of frequent and sustained plateletpheresis on peripheral blood mononuclear cell populations and lymphocyte functions of normal volunteer donors. *Transfusion* 1986;26:446–452.

22. Caspar CB, Segar RA, Burger J, et al. Effective stimulation of donors for granulocyte transfusions with recombinant methionyl colony-stimulating factor. *Blood* 1993;81:2866–2871.

23. Dale DC, Liles WC, Llewelyn C, et al. Neutrophil transfusions: kinetics and functions of neutrophils mobilized with granulocyte colony-stimulating factor and dexamethasone. *Transfusion* 1998;38:713–720.

24. Shi PA, Ness PM. Two-unit red cell apheresis and its potential advantages over traditional whole-blood donation. *Transfusion* 1999;39:218–225.

25. Petoom F. Therapeutic cytapheresis. In: Westphal RG, Kasprisin DO, eds. *Current status of hemapheresis: indications, technology and complications*. Arlington, VA: American Association of Blood Banks, 1987:49–70.

26. AABB Hemapheresis Committee. *Guidelines for therapeutic hemapheresis*. Bethesda, MD: American Association of Blood Banks, 1995.

70
Pretransfusion Testing

W. John Judd

Pretransfusion testing comprises a series of policies and procedures including laboratory tests, the goals of which are to provide blood for transfusion that will have the optimal clinical effect without causing undue harm to the recipient. These elements of pretransfusion testing can be placed in one of three categories: those related to donor unit testing and processing, those related to patient sample collection and testing, and those that serve as a final check between the donor unit and intended recipient (Table 70.1). Proper performance of each element will enable the right unit of blood to be transfused to the right patient (1). This chapter addresses each element of pretransfusion testing. However, before embarking on such detailed discussion, it is appropriate to provide some background relative to the principles of immunohematology and the human blood groups. This background enables the reader to appreciate the rationale for the serologic procedures used in pretransfusion testing.

PRINCIPLES OF IMMUNOHEMATOLOGY

Part of "70 - Pretransfusion Testing"

Blood Group Antibodies

Blood group antibodies are proteins that react with antigens on the surface of red blood cells (RBCs). They can either be acquired naturally or through immunization with foreign RBCs (2).

IgM Antibodies

Naturally acquired antibodies are IgM proteins found normally in plasma from persons whose blood cells lack the corresponding antigen. They are stimulated by antigens present in the environment (3). Antibodies to A and B antigens are of this type. Bacteria constituting normal intestinal flora carry blood group A- and B-like polysaccharides. As these flora establish in the gut, they provide the immune stimulus for anti-A and anti-B (3). The antibodies thus formed react *in vitro* as hemagglutinins or hemolysins; their ability to do so is a reflection of their size and number of sites available for antigen-binding (4) (Table 70.2).

IgG Antibodies

Most blood group antibodies are IgG, which are immune in origin and do not appear in plasma/serum unless the host is exposed directly to foreign RBC antigens. The usual stimulating event is blood transfusion or pregnancy. IgG antibodies are smaller than IgM, have only two sites for antigen-binding, and may only coat RBCs rather than act as direct agglutinins.

All antibodies to RBC antigens, other than naturally occurring anti-A and anti-B, are considered unexpected. They can either be alloantibodies, directed toward non-ABO-system antigens absent on the RBCs of the antibody producer, or autoantibodies directed toward self-antigens. The latter may cause autoimmune hemolytic anemia. Unexpected antibodies in donor plasma may destroy recipient RBCs, whereas antibodies in the recipient may cause accelerated destruction of transfused RBCs. In pregnant women, such antibodies may cross the placenta and cause hemolytic disease of the newborn (5).

ANTIGEN-ANTIBODY INTERACTIONS

Part of "70 - Pretransfusion Testing"

Interaction between blood group antibodies can be observed either directly by examination of RBC and antibody mixtures for agglutination (or clumping) and hemolysis or indirectly by use of the antiglobulin test. These two serologic techniques are summarized in Table 70.3.

Direct Agglutination

Because of their size and number of antigen-binding sites, IgM antibodies react directly with RBCs in tests performed on glass slides or in test tubes. For the latter, RBCs are mixed with antibody, briefly incubated at room temperature, and centrifuged. The centrifuged tests are examined for agglutination and hemolysis.

Examination for hemagglutination may be performed either macroscopically or microscopically, but the latter is not encouraged because it leads to the detection of nonspecific reactivity. Shown in Fig. 70.1 are a series of tests manifesting varying degrees of agglutination. The strength of the observed reactions can be graded. The values given are based on a scoring system in common use (6).

Hemolysis

Anti-A and anti-B also cause direct lysis of group A (or B) RBCs, especially if tests are incubated at 37°C. This hemolysis results from the action of complement, a series of α and ß globulins that

act in sequence as enzymes (e.g., esterases) to attack the RBC membrane and cause holes to be formed through which the intracellular hemoglobin can escape (Fig. 70.2). For initiation of the complement cascade, two pairs of immunoglobulin heavy chains must be bound in close proximity to each other on the RBC surface. Initiation of the complement cascade readily occurs when a single pentameric IgM antibody molecule is bound; for it to oc-cur with IgG antibodies, there must be two IgG molecules bound in close proximity to each other. Further, not all IgM antibodies bind complement to RBCs and complement binding does not always proceed to complete RBC lysis by membrane attack complex. Rather, activated C3b may be cleaved to C3d, which does not initiate the binding of C5; hence, no membrane attack complex is formed. However, C3d remains bound to RBCs and can be detected by the antiglobulin test (see below) (4,5).

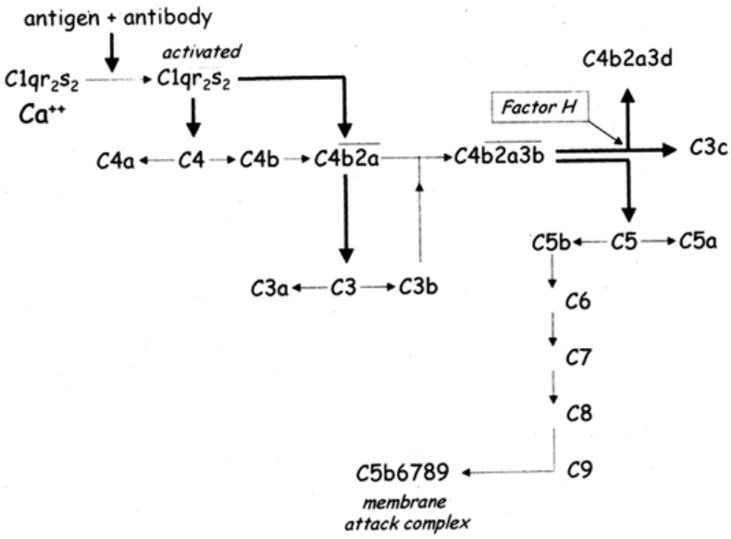


FIGURE 70.2. The complement cascade (classic pathway).

Antiglobulin Test

IgG antibodies and complement components bound to RBCs by either IgM or IgG antibodies are best detected by the antiglobulin test, in which antibodies raised in animals (usually rabbit or sheep) or prepared by hybridoma technology, are used to detect human IgG and complement bound to RBCs. The principles of this test are as follows:

1. Antibody molecules and complement components are globulins.
2. The injection of human globulins, either purified or in whole serum, into an animal stimulates the animal to produce antibodies against the foreign globulins. These antibodies are antihuman globulins; they can also be produced by hybridoma technology. The antihuman globulins that are important for blood group serologic work are anti-IgG and anti-complement (anti-C3d).
3. Anti-human globulin (AHG) will react with human globulins, either bound to RBCs or free in serum. Thus, RBCs must be washed free of unbound globulins before testing with AHG. This is crucial to the avoidance of false-negative tests owing to neutralization of AHG by unbound globulins.
4. Washed RBCs coated with human globulin are agglutinated by AHG.

Antiglobulin tests can be performed either indirectly after *in vitro* incubation of RBCs with serum or directly to demonstrate that RBCs are coated with globulins *in vivo*. The indirect antiglobulin test (IAT) is used to detect and identify unexpected antibodies in the serum of blood donors, prospective transfusion recipients, and prenatal patients. The direct antiglobulin test (DAT) is used to detect antibodies bound to RBCs *in vivo*; such antibodies may be seen in patients with hemolysis owing to autoantibodies or drugs, infants with hemolytic disease of the newborn, and patients manifesting an alloimmune response to a recent transfusion (5).

ANTIGENS, GENES, AND BLOOD GROUPS

Part of "70 - Pretransfusion Testing"

Approximately 300 distinct blood group antigens have been identified on human RBCs. Biochemical analysis has revealed

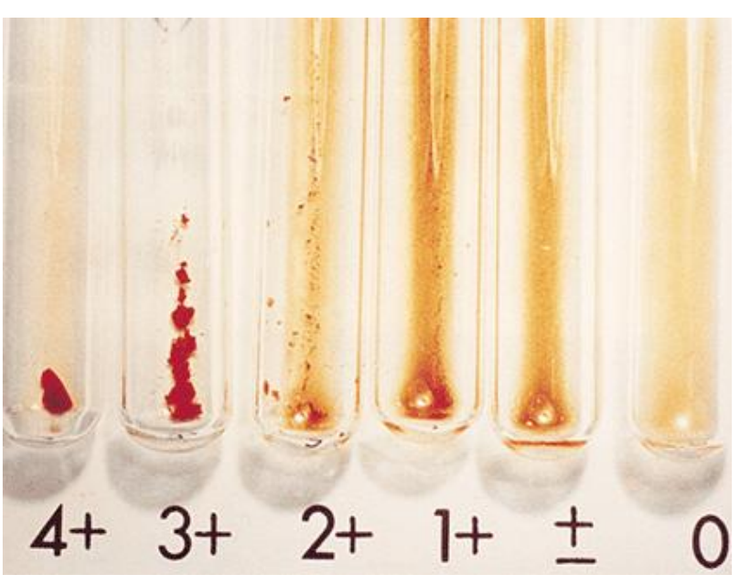


FIGURE 70.1. Hemagglutination reactions ranging from 4+ (strong positive) to 0 (negative). Based on the scoring system of Marsh (6).

that most antigen structures are either protein or lipid in nature. In some instances, blood group specificity is determined by the presence of attached carbohydrate moieties. The human A and B antigens, for example, can be either glycoprotein or glycolipid, with the same attached carbohydrate structure. With few exceptions, blood group antigens are an integral part of the cell membrane.

Several different notations are used and different concepts have been put forth to explain the genetics of the human blood groups. The presence of a gene in the host is normally reflected by the presence of the corresponding antigen on the RBCs. Usually, a single locus determines antigen expression, and there are two or more alternative genes or alleles (e.g., *a* and *b*) that can occupy a locus. Each individual inherits one allele from each of their parents. For a given blood group, when the same allele (e.g., *a*) is inherited from both parents, the offspring is said to be homozygous for *a* and only the antigen structure defined by *a* (a) will be present on their RBCs. When different alleles are inherited (i.e., *a* and *b*), the individual is heterozygous for *a* (and *b*) and both a and b antigens will be found on their RBCs.

In some blood group systems, several loci (or subloci) govern the expression of multiple blood group antigens within that system.

TABLE 70.1. ELEMENTS OF A PRETRANSFUSION TESTING PROCESS

Donor testing	
	ABO and Rh typing[a]
	Detection of unexpected antibodies to RBC antigens in plasma[a]
	Testing for infectious disease markers[a]
	Confirmation of ABO type, and Rh type of units labeled Rh negative
Recipient testing	
	Positive identification of patient before collection of blood samples
	Procurement and labeling of blood samples
	ABO and Rh typing
	Detection of unexpected antibodies to RBC antigens in plasma/serum
	Comparison of current test results with records of previous tests
Donor/recipient testing	
	Selection of RBC components that are compatible with the intended recipient's ABO group and with clinically significant unexpected antibodies, if present or previously detected
	Performance of tests to detect ABO incompatibility (and other known incompatibility) between selected RBC components and the intended recipient
	Positive identification of patient before transfusion
	Confirmation of concordance between donor unit label information and patient name and unique hospital identification number

[a]These tests, if performed by blood supplier, need not be repeated by transfusion service.

TABLE 70.2. CHARACTERISTICS OF BLOOD GROUP ANTIBODIES

Characteristic	Naturally Acquired	Immune
Immunoglobulin	IgM	IgG
Sedimentation constant	19 S	7 S
Molecular weight	900–1,000 kDa	150 kDa
Electrophoretic mobility	Between β and γ	γ
Bind complement	Often	Some
Placental transfer	Yes	No
Direct agglutinin	Yes	No
Hemolytic *in vitro*	Often	No
Example	Anti-A, anti-B	Rh antibodies

TABLE 70.3. TWO TYPES OF RBC SEROLOGIC TESTS

Direct Tests	Indirect Tests
Mix serum and RBCs	Mix serum and RBCs[a]
Incubate (room temperature, optional)	Incubate (37°C)
Centrifuge	Centrifuge, examine for agglutination and hemolysis (optional)
Examine for agglutination and hemolysis	Wash to remove unbound globulins
	Add antihuman globulin reagent
	Centrifuge
	Examine for agglutination
	Confirm negative tests with IgG-coated RBCs

[a]An enhancement reagent to promote antibody uptake may be incorporated here.

These loci are usually closely linked, located adjacent to each other on the chromosome. Such complex loci may contain multiple alleles and are referred to as haplotypes (2).

BLOOD GROUP SYSTEMS

Part of "70 - Pretransfusion Testing"

Some 200 antigens have been assigned by the International Society for Blood Transfusion (7,8) to 25 different blood group systems (Table 70.4). For a system to be established, the genes involved must be distinct from other blood group system genes and must either be polymorphic (i.e., two or more alleles, each with an appreciable frequency in a population) or the chromosome location must be known. Other antigens that do not meet these criteria have either been placed into collections, based primarily on biochemical data or phenotypic association, or into a series of either high- or low-frequency antigens.

TABLE 70.4. HUMAN BLOOD GROUP SYSTEMS

System Name	ISBT[a] Symbol	System Number	Number of Antigens	Chromosome Location	Gene Product(s)
ABO	ABO	001	4	9q34.1-q34.2	*A* = α-*N*-acetylgalactosaminyl transferase
					B = α-galactosyl transferase
MNS	MNS	002	43	4q28-q31	*GYPA* = glycophorin A; 43-kDa single-pass glycoprotein
					GYPB = glycophorin B; 25-kDa single-pass glycoprotein
P	P1	003	1	22q11-qter	α-galactosyl transferase
Rh	RH	004	45	1p36.13-p34	*RhD* and *RhCE*, 30–32-kDa multipass polypeptides
Lutheran	LU	005	18	19q13.2	78- and 85-kDa single-pass glycoproteins
Kell	KEL	006	23	7q33	93-kDa single-pass glycoprotein
Lewis	LE	007	6	19p13.3	α-fucosyl transferase
Duffy	FY	008	6	1q22-q23	38.5-kDa multipass glycoprotein
Kidd	JK	009	3	18q11-q12	43-kDa multipass glycoprotein
Diego	DI	010	18	17q12-q21	Band 3; 95–105-kDa multipass glycoprotein
Cartwright	YT	011	2	7q22	Acetylcholinesterase; GPI-linked glycoprotein
Xg	XG	012	1	Xp22.32	22–29-kDa single-pass glycoprotein
Sciana	SC	013	3	1p36.2-p22.1	60–68-kDa glycoprotein
Dombrock	DO	014	5	12p13.2-12p12.1	47–58-kDa GPI-linked glycoprotein
Colton	CO	015	3	7p14	28- and 40–60-kDa multipass glycoproteins
Landsteiner-Weiner	LW	016	3	19p13.3	37–43-kDa single-pass glycoprotein
Chido/Rogers	CH/RG	017	7	6p21.3	Complement (C4), glycoproteins adsorbed onto RBCs
Hh	H	018	1	19q13	α-Fucosyl transferase
Kx	XK	019	1	Xp21.1	37-kDa multipass glycoprotein
Gerbich	GE	020	7	2q14-q21	*GYPC* = glycophorin C, 40-kDa single-pass glycoprotein, glycophorin D, 30-kDa single-pass glycoprotein
Cromer	CROM	021	10	1q32	DAF (CD55); GPI-linked glycoprotein
Knops	KN	022	5	1q32	CR1 (CD35); single-pass glycoprotein
Indian	IN	023	2	11p13	CD44; 80-kDa single-pass glycoprotein
Ok	OK	024	2	19pter-p13.2	CD147; 35–69-kDa glycoprotein
Raph	RAPH	025	2	11p15.5	MER2; glycoprotein

[a]International Society of Blood Transfusion notations.

[b]Nonglycolipid antigens on red blood cells are either single-pass/multipass proteins or glycosylphosphatidylinositol (GPI)-linked proteins.

ABO Blood Group

Discovered by Karl Landsteiner in 1901 (9), ABO was the first human blood group system to be described. Three major alleles at the ABO locus on chromosome 9 govern the expression of A and B antigens (10). Gene *A* encodes for a protein (α-*N*-acetylgalactosaminyl transferase) that attaches a blood group specific carbohydrate (α-*N*-acetyl-D-galactosamine) and confers blood group A activity to a preformed carbohydrate structure called H antigen. Gene *B* encodes for an α-galactosyl transferase that attaches α-D-galactose and confers blood group B activity to H antigen. In both instances, some H remains unchanged. The *O* gene has no detectable product; H antigen remains unchanged and is strongly expressed on the RBCs (Fig. 70.3a).

These three genes account for the inheritance of four common phenotypes A, B, AB, and O. A and O blood types are the most common, and AB the least common (5). The *A* and *B* genes are codominant; that is to say when the gene is present, the antigen can be detected. The *O* gene is considered an amorph because its product cannot be detected. When either A and/or B antigens are present on the RBCs, the corresponding antibody(ies) should not be present in the serum or plasma. In adults, when A and/or B are absent from the RBCs, the corresponding naturally acquired antibody is present in the serum (2,3). This reciprocal relationship between antigens on the RBCs and antibodies in the serum is known as Landsteiner's law (5). Other ABO phenotypes do exist, but these are quite rare. Further, the A blood type can be subdivided, based on strength of antigen expression, with A1 RBCs having the most A antigen (3).

At the molecular level, *A* and *B* allelic cDNA are identical except for seven nucleotide substitutions resulting in four amino acid substitutions between A and B transferases. The majority of *O* alleles have a single nucleotide substitution close to the N-terminal of the coding sequence that shifts the reading frame of codons and results in translation of a protein lacking A or B

transferase activity (2). In addition to the ABO locus, expression of A and B is influenced by *H*/*h* and *Se*/*se* genes on chromosome 19. The *H* gene–specified transferase attaches fucose to paragloboside, a RBC membrane glycolipid. Paragloboside is synonymous with type 2 chains (Fig. 70.3a). The H antigen thus formed is the structure on RBCs that serves as the substrate for A and B gene–specified transferases described above. In rare individuals, absence of H (i.e., homozygous h) results in absence of H structures on RBC membranes and no substrate on which A and B gene products can act. The resulting phenotype is referred to as Bombay (or O_h) (2,3). The *Se* gene governs the addition of fucose to type 1 chains in the secretions (e.g., saliva, plasma, urine) to form water-soluble H-active structures (Fig. 70.3b). *Se* is present in approximately 80% of the population; depending on genes present at the ABO locus, the secretions of *Se* individuals will contain A and/or B antigens in addition to H antigens. *Se* gene products also interact with products encoded by genes at the LE (Lewis) locus also present on chromosome 19 (Fig. 70.3c).

Rh Blood Group

Currently 45 antigens are assigned to the Rh blood group system, although D is the most important. Other important Rh antigens are C, c, E, and e (5). Rh antigen expression is controlled by two adjacent homologous structural genes on chromosome 1 that are inherited as a pair or haplotype (11). The *RhD* gene encodes D antigen and is absent on both chromosomes of most Rh-negative subjects. The *RhCE* gene encodes CE protein. Amino acid substitutions at two different positions on the CE protein account for the Cc and Ee polymorphisms.

RBCs that carry D are called Rh positive, RBCs lacking D are called Rh negative. Approximately 85% of the white population are Rh positive and 15% are Rh negative (2). An additional consideration in pretransfusion testing is the existence of a weak form of D called D^u; this weak form of D can be detected by the IAT after incubation of RBCs with anti-D (5,12). Donor blood is tested for weak expression of D, by either the IAT or an equally sensitive method. However, it is not necessary to test the RBCs of apparent Rh-negative prospective recipients for weak expression of D; they can be transfused with Rh-negative blood.

ELEMENTS OF PRETRANSFUSION TESTING

Part of "70 - Pretransfusion Testing"

Donor Issues

Collection Facility

The initial ABO, Rh, and antibody detection tests on donor bloods, tests for infectious diseases, the interpretation of these tests, and correct labeling of donor units are functions normally carried out by a regional donor center. However, some hospital-based transfusion services continue to procure a portion of their blood needs from the population at large (allogeneic donors), and in recent years, efforts have been made for patients awaiting elective surgical procedures to deposit their blood for later use (autologous donors). In addition, some patients request that they receive blood from relatives or friends (directed donations). Regardless of the type of donor, each unit of blood must be subjected to tests for ABO, Rh, and unexpected antibodies, and for serologic markers of hepatitis B and C, human immunodeficiency viruses (HIV) 1 and 2, human T-cell lymphotrophic viruses type I, and syphilis (13). Donor blood must also be tested for alanine aminotransferase (ALT) that is elevated in hepatitis and other liver disorders. In addition, nucleic acid testing (NAT) is currently being performed to detect early evidence of exposure to HIV and hepatitis C.

The volume of testing that needs to be done at donor centers often necessitates use of automated equipment. Instrumentation currently in use is often based on microplate technology, and tests results may be interpreted electronically. ABO grouping entails testing both donor RBCs and serum/plasma. RBCs are tested with anti-A and anti-B, and the serum/plasma is tested with A_1 and B RBCs. Anti-A,B and A_2 RBCs may also be utilized,

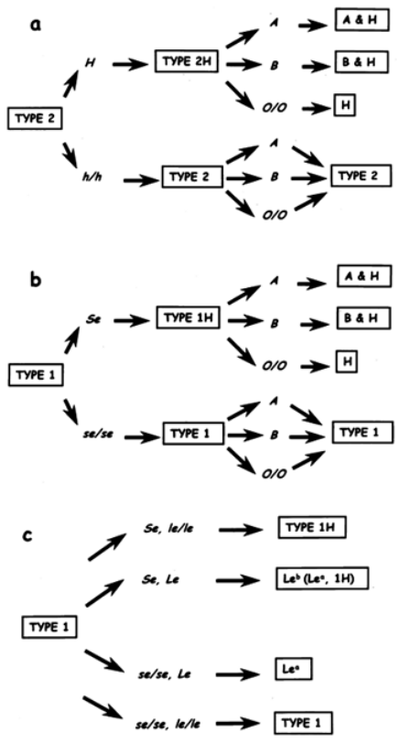


FIGURE 70.3. The interaction of the *ABO*, *H*/*SE*, and *LE* genes.

as results with these reagents serve to detect subgroups of A that may be nonreactive in direct tests with anti-A alone.

Rh typing is performed with anti-D. Because D is highly immunogenic (5), it is presumed that even weak expression of the antigen will evoke an immune response; consequently, donor RBCs that initially type as Rh negative in direct agglutination tests are tested further for weak D, usually by the antiglobulin test (Table 70.3). Only those units that are negative with anti-D by this method, or an equivalent procedure, can be labeled Rh negative (13). All straightforward D-positive and weak D-positive bloods are considered Rh positive.

Transfusing Facility

The ABO group of all units of whole blood or RBCs and the Rh group of those labeled Rh negative must be confirmed before transfusion (13). This must be done using a sample from an attached segment after the original ABO and Rh label has been affixed. Only tests with reagent antisera and donor RBCs need to be done; units labeled group O can be tested with anti-A,B alone (5). For D typing of units labeled Rh negative, only direct tests with anti-D are necessary; testing for weak expression of D is not required nor is repeat testing for unexpected antibodies or for markers of infectious diseases (13).

PATIENT ISSUES

Part of "70 - Pretransfusion Testing"

Sample Collection

The importance of the following measures cannot be overstressed because the major cause of fatal, hemolytic transfusion reactions is ABO incompatible transfusion resulting from patient/sample misidentification.

Requisition

Forms requesting blood and blood components must contain the first and last names and unique hospital registration number of the intended recipient (12). The name of the requesting physician, gender and date of birth of the patient, clinical diagnosis, and transfusion or pregnancy history are additional useful information.

Patient Identity

The collection of a properly labeled blood sample for pretransfusion testing from the correct patient is critical to safe blood transfusion. The person collecting the sample must positively identify the patient (5). This is facilitated through use of a wristband containing the patient's full name and unique hospital registration number that remains attached to the patient throughout the hospitalization. The information on the requisition form should be compared with that on the wristband; blood samples should not be collected if there is a discrepancy. In the absence of a wristband, the nursing staff should identify the patient; this should be documented on the requisition form. Nursing staff should be reminded to attach a wristband to the patient to validate patient identity at the time of transfusion. In an emergency, a temporary identification number should be used and cross-referenced with the patient's name and hospital identification number once they are known.

Labeling

Blood samples must be drawn into correctly labeled stoppered tubes. The tubes must be clearly labeled at the bedside with the patient's first and last names, the patient's unique hospital identification number and the date of collection (13). The phlebotomist should initial or sign the requisition form so that there is a means of identifying the person who collected the sample. By filing the requisition form with the patient's medical records, a permanent record is made of the phlebotomist's name (5).

Confirmation of Sample Identity

On receipt of blood samples for pretransfusion testing, the information on the label must be compared with that on the requisition. A new sample must be obtained whenever there are discrepancies or if there is any doubt about the identity of the sample. It is unacceptable to correct an incorrectly labeled sample.

Type of Sample

Either serum or plasma may be used for pretransfusion testing, but most workers use serum to avoid introducing small fibrin clots into serologic tests. Such clots may be mistaken for agglutination. Fibrin clots may also form when samples from heparinized patients are collected into nonanticoagulated tubes. These samples will clot properly after the addition of protamine sulfate (5). Some workers prefer to use serum rather than plasma for compatibility testing to facilitate detection of antibodies that primarily coat RBCs with the C3d component of complement. Bound C3d will not activate the lytic phase of the complement cascade but can be detected with AHG reagents containing anti-C3d (4,5). Ethylenediaminetetraacetic acid (EDTA), citrate, and other commonly used anticoagulants chelate calcium ions that are essential for complement activation. However, as discussed later, the use of AHG reagents containing anti-C3d for compatibility testing is not mandatory (5).

Age of Specimen

To ensure that the specimen used for compatibility testing is representative of a patient's current immune status, serologic studies must be performed using blood collected no more than 3 days in advance of the transfusion when the patient has been transfused or pregnant within the preceding 3 months or when such information is uncertain or unavailable (12) because both transfusion and pregnancy may stimulate alloantibody production. From a practical standpoint, it is simpler to stipulate that all pretransfusion samples must be collected within 3 days before RBC transfusions rather than determine whether a patient has been recently transfused or pregnant.

Storage

Blood samples used for compatibility testing, including donor RBCs, must be kept at 1° to 6°C for at least 1 week after each transfusion (5). This ensures that appropriate samples are available for investigational purposes should adverse responses to transfusion occur.

ABO Typing

Both reagent antisera and RBCs for ABO typing are available commercially. Anti-A and anti-B are monoclonal antibodies prepared by hybridoma technology. Reagent RBCs are usually suspended in a preservative medium containing EDTA, which chelates calcium ions essential for complement activation, thereby preventing lysis of the RBCs by lytic anti-A and anti-B (5). ABO grouping is performed using a direct agglutination technique (Table 70.3) (14). RBCs are tested with anti-A and anti-B, and the serum or plasma with known A_1 and B RBCs. Use of anti-A,B and A_2 RBCs is optional but generally considered unnecessary when ABO typing potential transfusion recipients (5). The expected findings for each of the four common ABO phenotypes, are shown in Table 70.5. When interpreting the results of ABO grouping tests, it is important to note the reciprocal relationship that exists between the absence of A and/or B antigens on RBCs and the presence of the expected anti-A and/or anti-B in the serum. If there is conflict between cell and serum ABO tests, group O blood should be provided for transfusion until the discrepancy is resolved and the ABO group of the patient can reliably be determined (13).

TABLE 70.5. EXPECTED REACTIONS OF THE FOUR COMMON ABO PHENOTYPES: RESULTS OF BLOOD TYPING TESTS

Blood Type	Anti-A	Anti-B	A_1 RBCs	B RBCs
O	0*	0	4+	4+
A	4+	0	0	4+
B	0	4+	4+	0
AB	4+	4+	0	0

O, no agglutination; 4+, strong agglutination (see Fig. 70.1).

Rh Typing

Only tests for D are performed routinely on the RBCs of prospective transfusion candidates. There are two different types of reagent anti-D available for Rh typing. High-protein reagents are prepared with human IgG anti-D diluted in bovine albumin and other substances that potentiate agglutination. Their final protein concentration may be greater than 20 g/dL. Such a high protein level is needed to potentiate IgG antibody reactivity so that positive and negative tests can be recognized almost instantaneously using a direct agglutination technique. Low-protein reagents (protein content less than 7 g/dL) are a blend of monoclonal IgM and monoclonal/polyclonal human IgG anti-D. With monoclonal-polyclonal anti-D blends, the IgM component causes direct agglutination of Rh-positive RBCs and the IgG component permits detection of the weak expression of D by application of the antiglobulin test (5). Only direct tests with anti-D are required on patient samples; the test for weak D (D^u) is not necessary (5,13). To avoid incorrect designation of an Rh-negative recipient as Rh positive because of autoantibodies or abnormal serum proteins, a control system appropriate to the anti-D reagent in use is required (13). For low protein anti-D, a concurrent negative test with anti-A and/or anti-B is considered an appropriate control system. Apparent AB Rh-positive samples should be retested concurrently with anti-D and an inert Rh control reagent (14).

Tests for Unexpected Antibodies

Methods for detecting unexpected antibodies in the serum or plasma of prospective transfusion recipients must be those that detect clinically significant antibodies (13). An IAT after 37°C incubation of patient's serum or plasma with reagent RBCs that are not pooled is usually required. When unexpected antibodies are detected, they must be identified (details are given in reference 5). If antibodies of specificities known to cause accelerated destruction of transfused incompatible RBCs are identified, then blood that lacks the corresponding antigen and is cross-match compatible by IAT should be selected for transfusion.

Several options exist regarding the selection of methods for pretransfusion antibody detection (Table 70.6 and Table 70.7). Decisions relative to these options are within the purview of the blood bank medical director. They should be made based on the type of patient served, the causes and frequency of previous significant antibody-mediated transfusion reactions, the availability of resources, and with the realization that no one method will detect all clinically significant antibodies.

TABLE 70.6. ACCEPTABLE METHODS FOR PRETRANSFUSION ANTIBODY DETECTION

	Serum	RBCs	Incubation	AHG
Saline	2–3 drops	1 drop	30–60 min, 37°C	IgG/PS
Albumin	2–3 drops	1 drop	15–30 min, 37°C	IgG/PS
Low ionic strength saline	Equal volumes		10–15 min, 37°C	IgG/PS
Gel test	25 µL	50 µL	15 min, 37°C	IgG
Polyethylene glycol	Equal volumes		15–30 min, 37°C	IgG
ReACT system	1 drop	1 drop	10 min, 37°C	None
Solid-phase adherence	1 drop		10–15 min, 37°C	IgG
Low ionic polycation	100 µL		1 min, RT	None

AHG, antihuman globulin; IgG, immunoglobulin G; PS, polyspecific AHG; RT, room temperature.

Tube Test Methods

Various RBC suspending media or additives are used either to enhance antibody uptake or to potentiate the agglutination phase of antibody-antigen interactions. Low ionic strength saline (LISS) solution (15), normal saline, or RBC preservative (modified Alsever's solution) are used as RBC suspending media. Bovine serum albumin (22% or 30% wt/vol), LISS additives, and polyethylene glycol are commonly added directly to serum-RBC mixtures (5,16). In the low-ionic polycation technique, serum is incubated with RBCs suspended in LISS solution to enhance antibody uptake (17). The second phase of the reaction is facilitated by aggregating RBCs with the polycation Polybrene® (hexadimethrine bromide). Aggregation is reversed with sodium citrate, but agglutinates formed by antigen-antibody interactions are not dispersed.

Because of enhanced antibody uptake in a low-ionic environment, use of either LISS for RBC suspension or LISS solution additives permits adequate detection of clinically significant antibodies after short incubation times (10 to 15 minutes). In contrast, incubation times of 30 to 60 minutes are required when albumin or saline tests are employed (5,14). For antiglobulin testing, either anti-IgG or polyspecific AHG, containing anti-IgG and anti-C3, may be used (5). However, use of polyspecific

AHG leads to the detection of an inordinate number of unwanted positive tests, and its use in routine antibody detection tests is not recommended (18,19).

Column Technologies

A gel test for detecting RBC antigen-antibody interactions was first described in 1990 by Lapierre and colleagues (20). Cards consisting of six microcolumns, each containing a viscous barrier below which is an agarose gel suspended in anti-IgG are commercially available. Atop each card is an incubation chamber in which reagent RBCs and test plasma are dispensed; the cards are incubated at 37°C, then centrifuged. As the RBCs pass through the viscous barrier, they are separated from the serum/plasma and come into contact with anti-IgG. If the RBCs become coated with antibody during incubation, they will be agglutinated by anti-IgG. The agglutinated RBCs become trapped in the gel; unagglutinated RBCs pellet to the bottom of the microcolumn. A procedure for the gel test is depicted in Fig. 70.4.

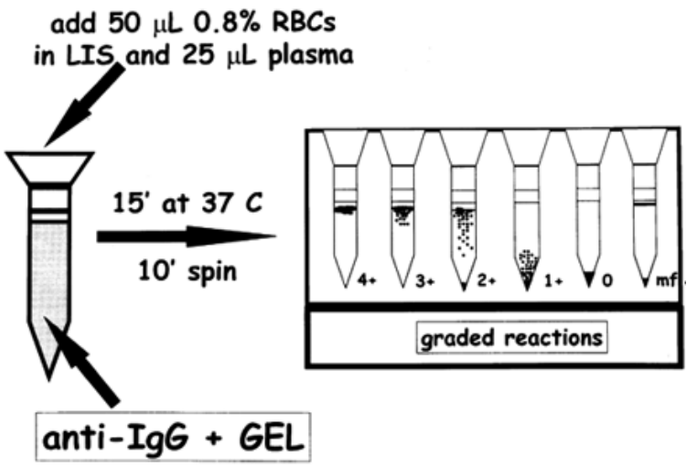


FIGURE 70.4. The gel test for detecting unexpected antibodies.

The gel test has proven to be equivalent to standard tube technologies (strength LISS, polyethylene glycol) for the detection of unexpected antibodies. The sensitivity and specificity of gel for potentially significant antibodies are 92% and 96%, respectively. This compares with 98% (sensitivity) and 90% (specificity) for a tube LISS procedure (21). Column technologies offer several advantages over conventional test tube procedures:

1. Simplified testing, as follows:
   a. Dispense measured volumes of test plasma/serum and RBCs (in LISS) into gel card incubation chambers.
   b. Transfer cards to an incubator (15 minutes) and then to a centrifuge (10 minutes).
   c. Read and record results.
2. When compared with conventional tube tests, there are no centrifugation/reading for direct agglutination after incubation, no addition of antiglobulin reagent, no need to validate negative tests utilizing IgG-coated RBCs. Omission of these manipulations (especially the associated, repetitive transfer of test tubes between racks and centrifuges) provides significant hands-on time savings.
3. Increased time savings through batch testing; a technologist can perform 36 to 48 antibody screens in approximately the same amount of time it takes to process 12 samples by conventional tube techniques (21).
4. Increased reproducibility of results, through use of measured volumes of reactants, elimination of the washing process, less subjective reading of tests.
5. Stability of reactions, facilitating validation of results by second technologist. With conventional tube methods, there is only one opportunity to read reactions; with the gel test, reactions can be read as many times as needed as long as 48 hours after centrifugation.
6. Reduction in the detection of cold agglutinins (anti-I/HI -M, -P_1, -Le) of no or doubtful clinical significance (seen during our initial evaluation of the gel test last year), with consequential reduction in the number of samples requiring antibody identification.
7. Can be automated or semiautomated using liquid sample handling devices, thereby further reducing hands-on time.

Two other column technologies are worthy of mention here. The BioVue (Ortho Clinical Diagnostics, Raritan, NJ) procedure is similar in many respects to the gel test except that fine glass particles are used instead of gel to separate agglutinated from unagglutinated red cells. This product is not currently available in the United States but is utilized in other parts of the world. The ReACT System (Gamma Biologicals, Houston, TX)

TABLE 70.7. OPTIONS IN PRETRANSFUSION TESTING

Option	Laboratories Performing
ABO/Rh	
Recipient ABO red cell tests with anti-A, B	50.9%
Recipient ABO serum tests with group A_2 red cells	8.4%
Detection of weak D on recipient red cells	Unknown
Antibody screen	
Immediate spin test	65.7%
Room-temperature incubation	6.9%
Polyspecific antihuman globulin, containing both anti-IgG and anti-C3d	31.4%
Pretransfusion direct antiglobulin test or autocontrol performed concurrently with screening tests for unexpected antibodies	55.4%
Microscopic examination of antiglobulin tests	52.8%
More than two reagent red cell samples for 61% detection of unexpected antibodies	61%
37°C reading, for direct agglutination after 37°C incubation.	86.7%

Laboratories' Performing percentages from, Maffei LM, Johnson ST, Shulman IA, et al. Survey on pretransfusion testing. *Transfusion* 1998;38:343–349.

is based on the affinity of staphylococcal proteins for immunoglobulins. In this procedure, RBCs and serum/plasma are incubated in a low-ionic environment and centrifuged through a column of protein G, which has a high affinity for the Fc fragment of human IgG. If antibody is bound to the RBCs during incubation, the passage of the coated RBCs through the protein G column is retarded. Uncoated RBCs will pellet to the bottom of the column (22).

Solid Phase Adherence Methods

Two forms of solid-phase adherence assays are available for RBC serologic testing. In direct tests, antibody is fixed to wells of a microplate, and RBCs are added (e.g., anti-A and anti-B for direct testing of donor/recipient ABO typing). After centrifugation, RBCs expressing the corresponding antigen will efface across the well; RBCs lacking the antigen will pellet to the bottom of the well.

In indirect tests (e.g., for detecting unexpected antibodies to RBC antigens), RBC membranes are affixed to microplate wells; test serum or plasma is added and the plates washed to remove unbound globulins. Indicator RBCs, which are coated with anti-IgG, are then added and the plates centrifuged. The indicator RBCs efface across the well in a positive test and pellet to the center of the well in a negative test (22). An overview of a solid phase adherence assay for antibody detection is depicted in Fig. 70.5.

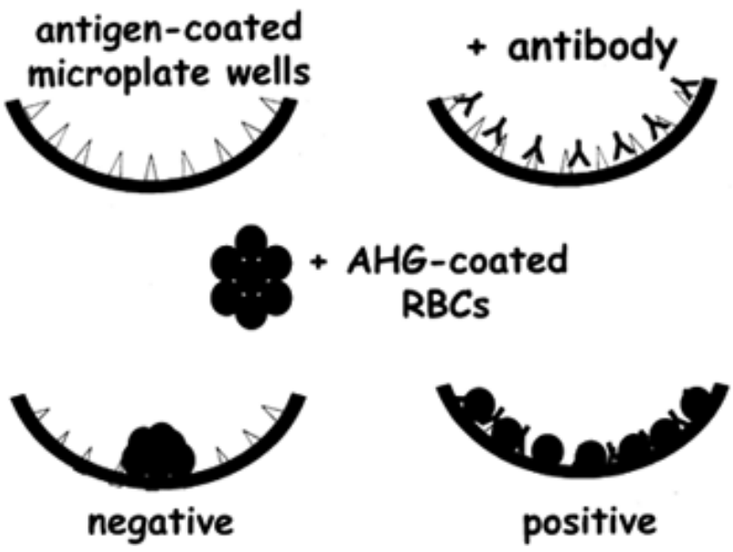


FIGURE 70.5. Solid phase adherence assay for detecting unexpected antibodies.

Reagent RBCs for Antibody Detection

The U.S. Food and Drug Administration (FDA) (23) currently mandates that sets of reagent RBC samples licensed for use in pretransfusion antibody detection tests carry expression of the C, c, D, E, e, Fy^a, Fy^b, Jk^a, Jk^b, K, k, Le^a, Le^b, P_1 M, N, S, and s. Such RBCs must not be pooled (1,13).

It is impossible to find a single donor with RBCs that carry all these antigens, for adults rarely, if ever, have strong expression of both Le^a and Le^b on their RBCs. Thus, reagent RBCs for antibody detection are available commercially as sets of either two or three samples. The Rh phenotypes of RBCs used in two-sample sets are R1R1 (D+C+c-E-e+) and R2R2 (D+C-c+E+e-). In three sample sets, an rr (D-C-c+-E-e+) sample is provided, in addition to R1R1 and R2R2 RBCs. Use of three RBC samples facilitates the inclusion of RBCs from individuals homozygous for particular blood group genes. Such RBCs tend to have a stronger expression of an antigen compared with RBCs from individuals heterozygous for the same gene; this phenomenon is known as dosage. It is easier to find double-dose expression of blood group antigens among three reagent RBC samples than among two samples; however, use of three samples increases the workload for antibody detection by 50% and rarely affords detection of significant alloantibodies that are not detected with two reagent RBC samples (24).

Further Options

Anti-IgG Versus Anti-IgG + C3

In addition to the selection of test methods and reagent RBCs for detection of unexpected antibodies, there are other options in pretransfusion testing. Table 70.7 shows the approximate number of transfusion services still performing tests that some workers

consider redundant because such testing is neither mandatory nor required by accrediting agencies (18). Use of polyspecific AHG reagents that contain anticomplement activity in addition to anti-IgG may facilitate detection of some antibodies that coat RBCs with complement components including, notably, anti-Jk^a and anti-Jk^b. Failure to detect such antibodies may lead to acute intravascular destruction of antigen-positive RBCs (5). However, use of polyspecific AHG also facilitates detection of IgM complement-binding autoantibodies that are of no clinical significance (19). Many of these antibodies are autoantibodies directed toward a self-antigen called I and are found naturally in virtually all normal adult human sera (4).

Room Temperature Incubation

The detection of some agglutinating alloantibodies, as well as the detection of IgM autoantibodies, is facilitated by the use of a room temperature incubation phase. However, antibodies that do not react at body temperatures rarely, if ever, cause significant destruction of transfused incompatible RBCs (4). Although room temperature incubation of tests was commonly utilized as recently as the early 1980s, many laboratories have since abandoned the practice (25).

Microscopic Examination of Tests

Examination for agglutination may be macroscopic, performed using an illuminated concave mirror, or microscopic. The latter is rarely necessary in routine practice; indeed, such critical examination of serologic tests can result in incorrectly recording negative tests as positive (4,14).

Autocontrol

Some transfusion services routinely perform an autocontrol as part of pretransfusion testing. The autocontrol consists of testing the patient's serum against their own RBCs under the same conditions as those to which screening tests for unexpected antibodies are subjected. This test, or a DAT, is performed to detect globulins bound to the patient's RBCs *in vivo*. Such *in vivo* coating of RBCs occurs in patients with autoimmune hemolytic anemia or hemolytic disease of the newborn and may also be seen after therapy with some drugs or transfusion with incompatible blood (5). Moreover, a positive DAT may be the earliest manifestation of an alloimmune response to a previous, recent transfusion (26).

Inclusion of the DAT or autocontrol as part of routine pretransfusion testing is no longer advocated; in the absence of detectable serum antibodies, the predictive value of a positive DAT is so low (0.29%) that routine testing is not cost-effective (26). However, the DAT/autocontrol is a good predictive test for immune-mediated hemolysis when performed on patients with clinical manifestations of hemolytic anemia (27).

Prior Records Check

As part of ongoing quality assurance and for compliance with the *Standards for Blood Banks and Transfusion Services* of the American Association of Blood Banks (13), the results of current ABO, Rh, antibody detection, and compatibility tests must be checked against records of previous tests, if performed. This must be done before blood is released for transfusion, preferably at the time that pretransfusion tests are completed. The specific records that must be checked are those for ABO and Rh typing performed within the previous 12 months, and any difficulties in typing, unexpected antibodies, or severe adverse reactions to transfusion recorded in the past 5 years. Any discrepancies between past and present ABO and Rh typing results must be thoroughly investigated; the most likely explanation is that the present sample is not from the same individual whose blood was tested previously. Further, even in the absence of detectable unexpected antibodies in the current sample, a record of such unexpected antibodies in previous samples must be taken into consideration when selecting and cross-matching blood for present and future transfusions.

DONOR-RECIPIENT ISSUES

Part of "70 - Pretransfusion Testing"

Selection of Blood for Transfusion

ABO and Rh

RBCs and whole blood selected for transfusion should be compatible with the serum of the intended recipient. To avoid the hemolytic and often fatal consequences of an ABO mismatched transfusion, RBCs carrying A and/or B antigens should not be transfused to a patient unless the patient's RBCs also carry those antigens. Group O individuals should receive group O RBCs, but AB individuals can receive RBCs of any ABO type. Rh-negative individuals, particularly women with childbearing potential, should receive Rh-negative blood. Rh-positive individuals may receive blood of either Rh type (5).

Unexpected Antibodies

When unexpected antibodies are present, as indicated by positive screening tests, they must be identified. At a minimum, this involves testing the patient's serum against a panel of fully phenotyped reagent RBCs samples and determining the antigen(s) that is(are) common to all reactive samples. A detailed discussion of the approaches to antibody identification is beyond the scope of this chapter, and the interested reader is referred elsewhere (5,14). When the identified antibodies are known to cause accelerated destruction of transfused incompatible RBCs, blood selected for transfusion should be shown to lack the corresponding antigen or antigens (5,13). This entails testing donor units with reagent antisera that are available commercially or prepared in-house from previously investigated samples. Examples of potentially significant antibodies include those directed toward RH, JK, KEL, and FY system antigens and the S and s antigens of the MNS system, as well as most other antibodies active at 37°C and/or by the IAT (2,4). When antibodies with specificities directed toward M, N, P_1, and LE antigens are present, particularly when the antibodies react best at or below room temperature, blood selected for transfusion need only be shown to be compatible by IAT after 37°C incubation; demonstrating that

compatible units lack the relevant antigen(s) is not required (28,29). In instances in which clinically insignificant autoantibodies such as anti-I are present, least incompatible units should be selected for transfusion when it has been established that the autoantibody is not masking a concomitant, clinically significant alloantibody (5).

Cross-match

Before whole blood or RBCs are administered and except in an emergency, a major cross-match must be performed. This usually entails tests between donor RBCs selected for transfusion and the prospective recipient's serum or plasma sample that was used for ABO, Rh, and antibody detection tests. The methods used should be capable of detecting ABO incompatibility and include the IAT. However, in the absence of unexpected antibodies (and absence of records of prior detection of such antibodies) in the intended recipient's serum, only testing to detect ABO incompatibility is required (12).

Antiglobulin Crossmatch

When clinically significant unexpected antibodies are present or a patient's records indicate that such antibodies have been detected previously, blood selected for transfusion must be tested with the patient's serum or plasma by an IAT. Any of the methods described earlier for antibody detection can be used (5,13). An antiglobulin cross-match can also be done routinely on patients with nonreactive screening tests for unexpected antibodies. This will detect ABO incompatibility and may detect unexpected antibodies that were missed in pretransfusion screening tests. Unexpected antibodies to low-incidence antigens and antibodies manifesting dosage may be detected in this manner, as may antibodies missed in screening tests owing to technical error. However, the predictive value of a positive IAT crossmatch after nonreactive screening tests for unexpected antibodies is sufficiently low that many large hospital transfusion services do not perform the IAT cross-match except as required when unexpected antibodies are present or there are records of such antibodies (18,25).

Serological Detection of ABO Incompatibility

Only a procedure for detecting ABO incompatibility is required when screening tests for unexpected antibodies are negative and there is no record of the patient having had such antibodies in the past. Detection of ABO incompatibility can be done serologically or, in some situations, electronically through the use of computers. An immediate-spin cross-match between the prospective recipient's serum or plasma and donor RBCs suspended in EDTA-saline to prevent false-negative tests owing to prozone by complement-fixing, high-titer anti-A and -B is an acceptable serologic method for the detection of ABO incompatibility (30). Alternatively, the ABO groups of both the donor units and a blood sample from the intended recipient can be confirmed immediately before the units are released for transfusion (13).

Electronic Crossmatch

The *Standards for Blood Banks and Transfusion Services* (13) also permit the use of computers to detect ABO incompatibility, provided that:

1. The computer system has been validated on site; the validation process must show that the computer will prevent release of ABO incompatible units.
2. Only detection of ABO incompatibility is required. An antiglobulin cross-match must be done if there are unexpected antibodies present currently or by history.
3. There have been at least two determinations of the intended recipient's ABO group. One of the determinations must be from a current sample; the other determination can be made by duplicate testing of the current sample, by testing a second current sample, or by comparison with previous records.
4. The information in the computer includes the unit number, product component name, ABO and Rh type, and blood group confirmatory test interpretation for the donor unit.
5. The information in the computer includes the name, identification number, and ABO and Rh type of the intended recipient.
6. There is a method to verify correct entry of data.
7. There is logic to alert the user to discrepancies between donor unit label and confirmatory test results and to ABO incompatibilities between the intended recipient and the donor unit.

Facilities wishing to implement an electronic cross-match must submit a request for variance to the Code of Federal Regulations (23). Currently, 30 facilities are currently approved by the FDA to perform an electronic crossmatch (G. Conley, personal communication, 1999). Experience has shown that existing transfusion service computer software and carefully developed SOP (Standard Operating Procedures) can be used together to provide a safe and efficient means of detecting donor-recipient incompatibility without performing a serologic cross-match (31).

Release of Blood for Transfusion

Transfusion Form

A form indicating the intended recipient's name, unique identification number, and ABO and Rh type must be completed for each unit of blood to be released (13). This form must also include the donor unit identification number, its ABO and Rh type, the interpretation of the cross-match, and the identity of the person performing the test. If, as in an emergency, pretransfusion testing is incomplete, the form must indicate the current status of serologic testing (5).

Label

Before blood is released, blood bank personnel must attach a label or tag to the unit containing essentially the same information that appears on the transfusion form, namely, the recipient's first and last names, unique identification number, the donor unit number, cross-match interpretation, and the name of the person performing the test (5).

Inspection

The unit must be inspected visually before release; if any abnormality in color or appearance is noted the unit should not be issued (13). A record should be made of this inspection. The expiration date must also be checked to avoid issuing an outdated unit (5).

Release Records

In addition, there should be a suitable means (e.g., a log book) to record the name of the individual who releases a unit of blood, the patient for whom the blood was released, the date and time of issue, and the person to whom blood was issued or destination of the unit (5).

Bedside Check

Before administering blood, the physician's written order should be reviewed to verify the request for transfusion (5). The transfusionist who administers the blood is responsible for this and for performing a final errors check. Verification of the following must be recorded on the transfusion form (13):

1. Recipient identification: The name and identification number on the patient's wristband must be identical to the name and number on the form attached to the unit.
2. Unit identification: The unit number on the blood container must match the unit number on the transfusion form.
3. ABO/Rh: The ABO and Rh type on the donor unit primary label must agree with that recorded on the transfusion form.
4. Expiration date: The expiration date of the unit should be checked and the unit verified as acceptable for transfusion.

CONCLUSION

Part of "70 - Pretransfusion Testing"

In summary, pretransfusion testing constitutes a quality assurance program designed to detect serologic incompatibility between donor unit and the intended recipient and to prevent both clerical and technical errors that may have serious if not fatal consequences. Assurance of quality requires proper performance of each task. There can be no substitute for proper patient identification, proper sample labeling, and proper performance of serologic tests.

REFERENCES

1. Shulman IA. Controversies in RBC compatibility testing. In: Nance SJ, ed. *Immune destruction of RBCs*. Arlington, VA: American Association of Blood Banks, 1989:171–199.

2. Issitt PD. *Applied blood group serology*, 4th ed. Durham, NC: Montgomery Scientific Publications, 1998.

3. Race RR, Sanger R. *Blood groups in man*, 6th ed. Oxford: Blackwell Scientific Publications, 1965.

4. Mollison PL. *Blood transfusion in clinical medicine*, 9th ed. Oxford: Blackwell Scientific Publications, 1993.

5. Vengelan-Tyler V, ed. *Technical manual*, 13th ed. Bethesda, MD: American Association of Blood Banks, 1999.

6. Marsh WL. Scoring of hemagglutination reactions. *Transfusion* 1972;12:352–353.

7. Daniels GI, Anstee DJ, Cartron JP, et al. Blood group terminology 1995. From the ISBT Working Party on Terminology for Red Cell Surface Antigens. *Vox Sang* 1995;69:265–279.

8. Daniels GI, Anstee DJ, Cartron JP, et al. ISBT Working Party on Terminology for Red Cell Surface Antigens: Oslo report. *Vox Sang* 1999;77:52–57.

9. Landsteiner K. Uber Agglutinationserscheinungen normalen menescglichen Blutes. *Klin Wochenschr* 1901;14:1132–1134.

10. Watkins WM, Greenwell P, Yates D. The genetic and enzymatic regulation of the synthesis of the A and B determinants in the ABO blood group system. *Immunol Commun* 1981;10:83–100.

11. Cartron JP, Agre P. Rh blood group antigens: protein and gene structure. *Semin Hematol* 1993;30:193–208.

12. Stratton F. A new Rh allelomorph. *Nature* 1946;158:25–26.

13. Menitove J, ed. *Standards for blood banks and transfusion services*, 19th ed. Bethesda, MD: American Association of Blood Banks, 1999.

14. Judd WJ. *Methods in immunohematology*, 2nd ed. Durham: Montgomery Scientific Publications, 1994.

15. Low B, Messeter L. Antiglobulin tests in low-ionic strength salt solution for rapid antibody screening and crossmatching. *Vox Sang* 1974;26:53–61.

16. Nance S, Garratty G. A new technique to enhance antibody reactions using polyethylene glycol. *Transfusion* 1985;25:475(abst).

17. Lalezari P, Jiang AF. The manual Polybrene® test: a simple and rapid procedure for detection of RBC antibodies. *Transfusion* 1980;20:206–211.

18. Oberman HA, Judd WJ. Cost containment in transfusion medicine: a view from the United States. In: Cash JD, ed. *Progress in transfusion medicine, volume 3*. Edinburgh: Churchill-Livingstone, 1988.

19. Garratty G. The role of complement in blood group serology. *CRC Crit Rev Clin Lab Sci* 1985;20:25–56.

20. Lapierre Y, Rigal D, Adam J, et al. The gel test: a new way to detect red cell antigen-antibody reactions. *Transfusion* 1990;30:109–113.

21. Judd WJ, Steiner EA, Knafl PC. The gel test: sensitivity and specificity for unexpected antibodies to blood group antigens. *Immunohematology* 1997;13:132–135.

22. Judd WJ. ‘New’ blood bank technologies. *Clin Lab Sci* 1998;11:106–113.

23. U.S. Food and Drug Administration. *Code of federal regulations (21-CFR), parts 606 and 660*. Washington, DC: U.S. Government Printing Office, 2000.

24. Judd WJ. Testing for unexpected red cell antibodies—two or three reagent red cell samples. *Immunohematology* 1997;13:90–92.

25. Maffei LM, Johnson ST, Shulman IA, et al. Survey on pretransfusion testing. *Transfusion* 1998;38:343–349.

26. Judd WJ, Barnes BA, Steiner EA, et al. The evaluation of a positive direct antiglobulin test (autocontrol) revisited. *Transfusion* 1986; 26:220–224.

27. Judd WJ. Investigation and management of immune hemolysis—autoantibodies and drugs. In: Wallace ME, Levitt J, eds. *Current applications and interpretation of the direct antiglobulin test*. Arlington, VA: American Association of Blood Banks, 1988:47–103.

28. Cronin CA, Pohl BA, Miller WV. Crossmatch-compatible blood for patients with anti-P1. *Transfusion* 1978;18:728–730.

29. Issitt PD. Antibodies reactive at 30 Centigrade, room temperature and below: a technical workshop. In: Butch SH, ed. *Clinically significant and insignificant antibodies: a technical workshop*. Washington, DC: American Association of Blood Banks, 1979:13–28.

30. Judd WJ, Steiner EA, O'Donnell DB, et al. Discrepancies in ABO typing due to prozone: how safe is the immediate-spin crossmatch? *Transfusion* 1988;28:334–338.

31. Butch SH, Judd WJ, Steiner EA, et al. Electronic verification of donor-recipient compatibility: the computer crossmatch. *Transfusion* 1994;34:105–109.

71
Blood Component Therapy

S. Breanndan Moore

BASIC RATIONALE

Part of "71 - Blood Component Therapy"

The basic rationale of component therapy is the idea that a particular patient's blood cell or plasma deficiencies can be identified and precisely what is needed to correct those deficiencies can be supplied in a timely manner. This approach is logical from the standpoints of the individual patient and of those whose responsibility it is to ensure an adequate supply of blood and components for the population at large. By processing each unit of donated blood into multiple subunits containing concentrates of various cellular and noncellular products, one can benefit multiple patients from a single donation. If each recipient is transfused with only the moiety that he or she requires, then the system is truly efficient and effective.

BACKGROUND

Part of "71 - Blood Component Therapy"

At the beginning of the 20th century, medical pioneers began to use Landsteiner's seminal discovery of the ABO blood group system when choosing blood donors for a patient. In those days, their concerns were primarily the patient's desperate clinical straits, the availability of a matchable donor, and the ability to obtain the desired blood and infuse it without causing harm to the donor or recipient. Because of these significant concerns and the lack of standardized instruments and practices, the transfusion of blood was reserved for cases in which all routinely accepted methods of therapy had been exhausted and the patient's deteriorating condition warranted extraordinary measures. The prospect of receiving blood was often feared as much as the operation that might have created the need.

The ability to isolate many fractions or components from donated blood has been a major breakthrough in recent years. This process has been greatly stimulated by the increasing sophistication of both diagnostic acumen and treatment modalities.

The enormous array of complex surgical procedures, as well as the refinements in the management of medical diseases, has resulted in the successful treatment of patients who might, even a decade ago, have died. These advances have been accompanied by an equivalent and concomitant increase in the quantity of blood and blood components used. Many of these surgical and medical therapeutic regimens are directly dependent on the availability of an adequate and safe supply of appropriate blood and components.

Because the demands for blood sometimes outstrip the available supply, it is important to be concerned about the ability to respond not only to current patient requirements but also to the inevitable future demands generated by medical and surgical technological advances. There is no reason to believe that the trends of increasing utilization of blood and, particularly, blood components will change in the future.

Technological Advances

When blood was first collected from donors, it was drawn into glass bottles through rubber tubing. Such bottles and tubing were still being used in this country in the 1950s. In fact, glass bottles were used in Russia until just a few years ago. These glass bottles and rubber tubings were sterilized and reused.

The only component preparation possible from blood drawn into glass bottles was plasma separation by allowing the unit to stand immobile. The advent of disposable plastic bags and tubing created a dramatic new capability, i.e., manipulation of the donated blood in such a manner as to allow the separation of platelet-rich plasma from the red cells. This, in turn, gave us platelet concentrate and fresh frozen plasma (FFP). Subsequently, Pool et al. (1,2) discovered the antihemophilic properties of the cold-precipitated proteins from FFP. Thus, cryoprecipitated plasma was born, which has revolutionized the treatment of hemophilia.

The next major advances were in the development of various anticoagulants and preservatives. These developments progressively increased the functional shelf life of blood and components by providing the necessary substances to maintain the metabolism of the preserved cells at higher levels than was previously possible. The use of newer plastics allowed better gas exchange across the bag, allowing the cells in the bag to breathe and thereby enhancing their function after storage.

The development of both continuous and semicontinuous apheresis machines permitted the selective separation of relatively large quantities of platelets or granulocytes (or both) from single donors. This led to the use of single-donor cellular components for patients who were septic (granulocytes) or alloimmunized (platelets). The use of HLA class I typing to match donor and recipient added another dimension to the efficacy of platelet therapy. This technique has been further modified to permit selective removal of specific plasma contents. This is accomplished by passage of the plasma over columns designed to

remove specific moieties, e.g., IgG antibodies and the pruritogenic factors in biliary cirrhosis. These techniques also have been modified to produce enriched, activated, autologous lymphoid cell preparations for use in the treatment of malignant disease.

Fractionation on a commercial scale has permitted the acquisition and heat treatment of albumin from huge pools of plasma. This product has had extensive use in various situations, as described later.

Donor and Donation Efficiency

When the demand for blood outstrips the supply, episodic blood shortages occur, necessitating cancellation or postponement of elective surgical procedures and dependence on other geographic regions for the needed blood and components. Unfortunately, many large urban populations that have large blood needs also have inadequate donation patterns. This fact alone makes it particularly important that we achieve the maximal use of the units that are donated. Obviously, the manufacture of components from as many units as possible is the best way to achieve such efficiency. Apart from the need to respond to the transfusion needs of patients, we have an obligation not to misuse or underutilize the blood donated.

Economic Factors

The costs of obtaining a unit of blood have increased over the past few years for several reasons. Among these are (a) the increased cost of donor recruitment because of personnel and advertising charges, (b) the cost of the ever-growing list of donor tests mandated by regulatory or accrediting agencies, and (c) the somewhat less clearly defined but real costs incurred as a result of blood banks' greater attention to transfusion safety. These costs include the significant time involved in the administration of autologous, directed, and minimal-exposure transfusion (MET) programs (3).

With the autologous, directed, and MET programs, a specific unit must be clearly identified (from the moment of acceptance by the blood bank) as coming from a specific donor and designated for a specific recipient. This degree of commitment requires considerably more administrative effort and computer support than does the older homologous system. These added costs are inevitably passed on to the patients, but we have an obligation to be as efficient as possible in the use of the donations to minimize these supplementary costs.

COMPONENT PREPARATION

Part of "71 - Blood Component Therapy"

Anticoagulants and Preservatives

Various solutions have been developed to prevent the clotting of donated blood and to preserve its functional elements. These solutions have gradually become more complex as changes have been introduced to prolong the effective shelf life of the component, to alter its flow characteristics, and to increase the yield of components from the donation. The characteristics of these various solutions are discussed in detail in Chapter 69. Suffice it to say that the improvements in these solutions have played a major role in the burgeoning field of component therapy in the past decade.

Basic Techniques

Centrifugation

When blood was collected in glass bottles, most of it was stored and then transfused as whole blood because large-scale processing of these units could not be accomplished easily. Simple gravity separation of plasma could be performed but only by breaking the seal on the bottle. Elaborate efforts were required to carry out this separation aseptically, and although the plasma could be subsequently frozen, the residual red blood cells were often discarded or used for purposes other than infusion into patients.

The development of plastic closed-system multiple bags revolutionized blood banking and the practice of transfusion medicine. The plastic bags were malleable enough to allow centrifugal separation of plasma from red blood cells and subsequent diversion of the supernatant plasma into satellite bags while still maintaining the closed system by the connected plastic tubing. This simple technique, still widely used, permitted the easy and aseptic separation of red blood cells and plasma so that each could be safely stored and subsequently transfused. This simple change provided the ready availability of components that permitted aggressive surgical and medical therapeutic procedures.

The availability of a simple, centrifugation-based separation technique also led to experimentation with the duration and force that could be applied in either the initial centrifugation step or in a second centrifugation of the plasma. These manipulations eventually resulted in the production of platelet concentrates. For example, a gentle spin of a bag of freshly donated whole blood at approximately 2,500 rpm for 3 minutes yields a supernatant platelet-rich plasma that can be separated into a satellite bag. A second strong spin of this latter bag, at approximately 4,100 rpm for 5 minutes, packs down the platelets, allowing separation of the platelet-poor plasma into another satellite bag.

The centrifuges used for blood separation need to be maintained carefully and tested regularly so that they produce the requisite centrifugal forces in a consistent manner. Likewise, they have to be loaded carefully so that the cups are balanced. Otherwise, the rotor will shake and produce turbulence in the bags. The motors of these centrifuges are also susceptible to damage from imbalanced cup holders.

Plasma Separation

Plasma expressors are simple mechanical devices that consist of two flat, rectangular plates (metallic or hard plastic) hinged together at one end. A bag of previously centrifuged whole blood is carefully placed between the plates. A simple spring device pushes the hinged plates into apposition so that the supernatant plasma in the bag is extruded through one of the ports at the top of the bag. This port leads into one of a series of plastic satellite bags connected to one another by flexible plastic tubing. This latter tubing is an integral part of the closed, sterile, multiple-bag system. When the requisite quantity of plasma has been expressed,

the interconnecting tubing can be clamped with two small metal clips placed close together. The primary bag can then be separated by applying a scissors to the short segment of the tubing between the metal clips.

Leukocyte-Reduced Red Blood Cells

Several techniques are available for the removal of the majority of the white blood cells from a bag of whole blood or red blood cells (3). These methods may involve simple centrifugation with subsequent separation of the buffy coat from the remaining cells or the use of various filters to accomplish the same goal. The use of machines to wash the red blood cells will also provide variable degrees of white blood cell removal. The indications and more detailed discussion of the effectiveness of these methods follow later in this chapter.

Cryoprecipitation Techniques

When the plasma is separated from the red blood cells and then frozen at -18°C or lower, some proteins precipitate. These proteins are rich in coagulation factors VIII and XIII and fibrinogen. If frozen plasma is thawed at 37°C, these precipitates go back into solution. However, Pool discovered that, if the thawing occurred between 1° and 6°C, these cold-precipitated proteins remained as a precipitate from which the rest of the plasma can be readily expressed after prompt centrifugation (1). This cryoprecipitated protein should be refrozen within 1 hour.

The variation of this technique used to enhance the concentration of fibrinogen in the product so that it can be used topically as a fibrin glue is discussed below.

Red Blood Cell Washing

Devices are available that use solutions compatible with red blood cells to suspend and centrifuge the mixture in repeated cycles to produce a suspension of red cells in the isotonic saline-based wash solution. Variable quantities of white blood cells, platelets, or plasma constituents may remain with the red blood cells, depending on the assiduousness of the procedure used, and the technique may be modified to decrease the level of these contaminants. Because these washing techniques involve the rupture of the seals on the closed system, the shelf life of the washed red blood cell product is only 24 hours after washing.

Freezing Techniques

To preserve its labile coagulation factors, plasma must be rapidly frozen within 6 hours after blood collection (4). This can be accomplished by the use of a dry ice-ethanol bath or by a mechanical freezer capable of maintaining a temperature of 30°C or lower. When the plasma is frozen, it may be maintained at -18°C for as long as 1 year. Similarly, the supernatant plasma from cryoprecipitation can be refrozen at -18°C or lower and maintained for 1 year. It is preferable to use temperatures of -30°C or lower for optimal factor activity in the product.

Freezing red blood cells is considerably more difficult because the structural integrity, viability, and function of the cells must be maintained. There are several methods described for this, including high glycerol, agglomeration, and low glycerol (4). All three methods utilize the ability of glycerol to preserve the integrity of the red blood cells during freeze/thaw processes. Glycerol, a clear, syrupy fluid, is a trihydric alcohol and is miscible with both water and alcohol. It is relatively inert pharmacologically and does not cause problems except that, if the red cells have been improperly deglycerolized before infusion, there may be shifts in intracellular fluids. Another cryoprotectant that is sometimes used is dimethyl sulfoxide (DMSO), a by-product of petroleum distillation. When given intravenously, it can cause nausea, vomiting, and a strong, pervasive, garliclike taste and odor (5).

The cell injury caused by freeze/thaw seems to be a result of the formation of intracellular ice crystals. Intracellular water freezes more slowly than extracellular water if the rate of freezing is less than 10°C/min. This causes an osmotic gradient that results in the flow of water out of the cells, and damage may result from cellular dehydration (6). At much more rapid rates of freezing, there is little time for this gradient to form; therefore, very little dehydration occurs. However, because the cells do not lose their water, ice crystals tend to form and destroy intracellular organelles. Cryoprotective agents alter the tonicity of cells and thus slow the rate of freezing but not enough that severe dehydration occurs. The high-glycerol method is the one most widely used because (a) it does not require a controlled rate of freezing (the less expensive freezers perform satisfactorily) and (b) it requires only -80°C for its initial freezing as opposed to -196°C for the low-glycerol method, which also requires liquid nitrogen storage.

Multiple variations of the freezing methods have been reported, but some facts apply to all of them. Red blood cells should be frozen within 6 days after collection. Rejuvenation can be carried out several days later than this but before glycerolization. This rejuvenation process is designed to restore red cell adenosine 5′-triphosphate (ATP) and 2,3-diphosphoglycerate (2,3-DPG) levels. Each method has limitations, but all of them yield a product of high red blood cell purity and function, minimal white blood cell and platelet contamination, and virtually no residual plasma proteins. This is important when the presence of such contaminants might cause adverse reactions. Frozen red cells are approximately twice as expensive as ordinary red cells because of the costs of freezing, storage, and thawing. In addition, the deglycerolization and washing, which are obligatory before transfusion, take approximately 45 minutes per unit. Therefore, the cells cannot be used for emergency resuscitation. The method is used to preserve rare or unusual units as well as for autologous blood stored before anticipated surgical use.

PATHOGEN INACTIVATION TECHNIQUES

The emergence of human immunodeficiency virus (HIV) as a major health problem and the demonstration of its transmission by blood components have led to significant improvements in donor selection and deferral processes and focused regulatory attention on blood banking in general. They also acted as a catalyst for the identification, characterization, and development of testing methods for several other transfusion transmissible

viruses [e.g., hepatitis C virus (HCV), human T-cell lymphotrophic viruses (HTLV) I and II]. Although one might speculate about the relative contribution of each of these advances, there is no doubt that transfusions are safer today than they have ever been.

Media and judicial and regulatory attention to transfusion risks have also helped to create a public expectation when one is dealing with a biologically derived product that cannot be rendered completely free of transmissible pathogens. Nevertheless, this expectation has stimulated the development of several pathogen inactivation methods that appear to be quite promising, particularly for the plasma-based components. It has long been recognized that various forms of heat treatment could be effectively used to prevent transmission of hepatitis B virus (HBV) or non-A non-B hepatitis (now known as HCV). Recent developments have utilized methods involving the addition to the blood of a variety of chemical agents usually with some external radiation source to activate the chemical additive.

Heat (Pasteurization) Treatment

Treatment of albumin for 10 hours at 60°C in the presence of stabilizers has been used for three decades to destroy both enveloped and nonenveloped viruses. However, early studies indicated that various coagulant proteins were adversely affected by this method and it could not be used for all plasma-based products. More recently, limited studies have again suggested a possible role for this method with only approximately 10% loss of coagulant factor activity (7). Confirmatory data and clinical studies would be desirable. Interestingly, manipulations of the heating methods have led to some studies of so-called high temperature, short-time heating applied to plasma that had been spiked with several viruses. Processing at 77°C for 0.006 seconds resulted in very high levels of viral killing, e.g., more than 4.4 $\log_{10}$ for HIV. Protein structure and biological activity of some coagulant proteins were well maintained, but factor V was the most labile with 60% loss (8).

Solvent/Detergent Treatment

The discovery by Horowitz et al. (9) that an organic solvent tri(n-butyl)phosphate plus a detergent disrupted lipid-enveloped virus while leaving protein structure and function largely intact paved the way for development of this approach so that it could be applied to plasma products. The same group also demonstrated that gamma irradiation could be effective for viral killing but only at doses likely to produce adverse cellular effects (10).

Initial studies of solvent/detergent (SD) treatment utilized ethyl ether as the solvent, but the highly explosive nature of ether was a significant drawback to its general application. Conversely, tri(*N*-butyl)phosphate is nonvolatile with a proven efficacy for viral inactivation. SD methods work in the following manner: viral lipid membranes absorb the detergent that then disrupts and disintegrates the membrane in the presence of the solvent. This membrane destruction allows release of the nucleic acid core, which is destroyed by various plasma enzymes, thereby rendering the virus noninfectious (11).

SD methodology is effective at rendering noninfectious the following transmissible viruses: HIV-1 and -2, HCV, HBV and HTLV-I and -II. It is not effective against nonlipid-enveloped viruses such as hepatitis A virus (HAV) or Parvovirus B 19 nor is it thought to be useful against the prion responsible for Creutzfeldt-Jakob disease.

Commercial application of the SD method involves pooling large numbers of units of FFP before treatment (approximately 2,500). This pooling of the product raises concerns about the possibility of (hitherto undiscovered) pathogens that may not be inactivated and the scope of whose transmissibility would be greatly magnified by the pooling process. In light of events in the past two decades, this fear cannot be lightly dismissed. Conversely, current estimates of the viral destructive effects of this method are such that it is likely that one vial of contemporary plasma-derived coagulant-factor concentrate prepared by the SD method has less than one chance in 10^{16}, 10^{13}, and 10^{6} of infectious HIV, HBV, and HCV, respectively (12).

U.S. Food and Drug Administration (FDA) licensure was obtained in 1985 for the manufacture of antihemophilic concentrate, and such licensure was granted in May 1998 for similarly treated FFP. Clinical studies of SD-FFP indicate that coagulant factor activity is approximately 10% to 15% less than that for untreated plasma (13), and it is likely that medical indications for SD plasma will be essentially the same as those for FFP (11). The slight reduction in active coagulant factors should not require any compensatory dosage increase in most patients. However, in those requiring more than approximately three units of FFP, some compensation for the 10% to 15% loss should be considered if SD-FFP is to be substituted for FFP.

Cost is likely to be a significant factor in decisions about the use of SD-treated plasma. Resources spent on the SD treatment have already raised questions about its cost-effectiveness (13). Questions are also likely to be stimulated by the possibility that only a single supplier will be producing SD-FFP in the United States for the next few years. Questions about equality of access to the supply are obvious as are fears about the national vulnerability to manufacturer recalls. For these and other reasons, it is unlikely that SD-treated plasma will become the sole U.S. source of plasma-based products in the near future, although this product has completely replaced FFP in Norway and Belgium, and it is widely used in many other European nations. After more than 11 million units of SD-treated plasma blood products have been transfused, there have been no reports of adverse effects owing to the treatment (12). SD-treated plasma has been shown to be as effective as FFP for the treatment of a wide variety of clinical situations including thrombotic thrombocytopenic purpura.

Photodecontamination

As long ago as 1955, Murray et al. (14) demonstrated that ultraviolet (UV) light could be used to inactivate hepatitis viruses, but it also caused significant protein denaturation. Prince and colleagues (15) later showed that the addition of β-propiolactone to UV treatment was associated with inactivation of approximately 10^{7} infectious doses of HBV, whereas approximately 40% of factor IX activity was retained. The use of light sources as activators has some obvious drawbacks because the degree of light penetrance necessary for activation can be extremely difficult to

achieve without undue heat build up, and clearly this method could not be readily applied to units of red blood cells.

Investigators have concentrated their efforts on the treatment of platelet units that could be modified to spread the content very thinly over a relatively large surface area to counteract the low penetrance of the light and the requirement that the treated product not become excessively warm during the treatment.

Experimental devices have been developed to deliver approximately 3 J of UVA (320 to 400 nm) light per cm^2 in approximately 4 minutes to the surface of the bag. Psoralens can be added to the contents of the bag. These are planar aromatic molecules that can bind irreversibly to single- and double-stranded nucleic acids by intercalation. On illumination by UVA light, intercalated psoralens form interstrand cross-links with RNA and DNA. Bacteria or viruses, thus modified, are inactivated because transcription and replication are prevented if sufficient psoralen/UVA-generated lesions are created to overwhelm natural repair mechanisms. In one excellent study (16), very high levels of HIV, HBV, *Staphylococcus epidermidis* and *Klebsiella pneumoniae* were inactivated in platelet concentrates, whereas platelet functions were adequately retained. Practical application of this methodology will require improved methods for delivery of the light without concomitant generation of damaging heat. It is also important to point out that studies of UV light as a method of reducing HLA immunogenicity of platelets have utilized UVB rather than UVA light. Another of the potential problems associated the use of psoralens in units of platelets is the mutagenicity of residual psoralens. It has already been shown that hydrophobic resins can be experimentally used to adsorb residual functional psoralens and thereby significantly reduce their mutagenicity (17). Use of UVA light and psoralens remains experimental and seems unlikely to become a standard and FDA-approved method without some significant technological breakthrough. However, because of the enormous potential for elimination of bacterial, as well as viral threats from platelets, this approach is the object of significant highly focused research.

Methylene Blue

Another chemical additive that has been used to sterilize plasma is methylene blue (MB). Because light must also be used to activate the MB, this method is essentially limited to plasma products. The MB acts similarly to psoralens by intercalating between nucleic acid bases and rendering the virions replication incompetent. As for psoralens, questions of mutagenicity arise because of residual MB in plasma. Methods of adsorption of residual MB have been developed (18). MB methodology also shares another drawback with psoralens in that it does not address the pooling of intracellular viruses. An advantage of the MB method is that it can be applied to single units of plasma, thereby eliminating the pooling that is required for SD treatment. To date, there are no FDA-approved lighting devices that would permit MB treatment of large batches of units of plasma. Any large donor collection center would need to be able to MB treat large batches of units (50 to 100) because they may collect several hundreds of donations each day. MB treatment appears to have minimal effect on coagulation factor activity in plasma and has been used extensively and safely for several years in Germany and Switzerland (19).

Other Inactivation Methods

Unsaturated fatty acids have been shown to inactivate lipid-enveloped viruses in culture media but have significantly reduced activity in this regard in plasma or cryoprecipitated plasma (20) and so are unlikely to generate much interest despite their success in retaining coagulant factor activity in SD-treated plasma. Similarly, gamma radiation has been evaluated for plasma deliberately spiked with HIV. The doses of radiation necessary to inactivate the virus were such that the biological properties of the plasma proteins were compromised (21). A novel approach using cross-linked starch iodine was evaluated for some model viruses (lipid and nonlipid enveloped). This method showed promise because it retained approximately 70% of the coagulation factor activity (22). Further evaluation using truly pathogenic human viruses will need to be carried out. Another interesting method was found with the discovery of a potent oxidant ozone to plasma. As much as 11 log inactivation of HIV-1 was achieved by this method with minimal effect on coagulation factor activity (23).

Although technological advances can catapult any of these (or other) methods into a more prominent position, SD treatment of plasma has a clear lead because of its FDA licensure.

AVAILABLE BLOOD PRODUCTS AND INDICATIONS FOR USE

Part of "71 - Blood Component Therapy"

Whole Blood

A unit of whole blood has a volume of approximately 500 mL (440 mL of blood + 63 mL of anticoagulant/preservative). The hematocrit value is approximately 35% (12 g of hemoglobin per deciliter). When the concept of dividing each donation of whole blood into its constituent components became widely accepted as a rational and efficient way of providing optimal transfusion care, the use of whole blood declined. In the hospital setting, the degree of sophistication of the overall medical and surgical practice was judged, at least partly, by the ratio of red blood cell use to whole-blood use. Some surgeons continued to maintain that, because the patient bled whole blood, it was logical to reinfuse whole blood. Apart from the obvious need to fractionate blood so that components can be prepared, the most cogent argument to be made against the proponents of routine use of whole blood is that whole blood is not whole.

The deterioration of labile coagulation factor activity (largely factors V and VIII) and the virtual absence of platelet or granulocyte function negate claims that whole blood is whole in the functional sense. To make it whole, one would have to collect it and transfuse it rapidly, i.e., within hours, which is a logistic impossibility. Cardiac surgeons sometimes have requested fresh whole blood for use during, and particularly after, cardiopulmonary bypass surgery. In a recent paper, Lavee et al. (24) reported that one unit of fresh whole blood (collected on the morning of surgery) increased the platelet count to the same extent as six units of platelet concentrate did and restored platelet aggregation and volume comparably with eight and 10 units of platelets, respectively. These data must be interpreted with caution because previous data (25) indicate that the use of fresh autologous blood did not decrease blood use after bypass.

Although the use of whole blood is often decried, its use does provide both oxygen delivery and volume in one product. That is precisely what is needed in the treatment of acute hemorrhagic, hypovolemic shock. However, oxygen delivery can be provided by red blood cells, and volume can be provided by giving crystalloid or colloid solution without increasing the number of donors to which a patient needs to be exposed. Blood flow rates are slower with the more viscous red blood cells than with whole blood, but this can be overcome by adding isotonic crystalloid or colloid solution directly to the bag of red blood cells before use. However, in acute or critical clinical situations, the latter manipulation is somewhat cumbersome and time-consuming; this has led to renewed requests for whole blood (26).

More recently, the use of storage additive solutions with anticoagulants such AS-1 or AS-3 has restored the whole-blood flow characteristics to red blood cells (27), thereby facilitating their use in the operating room or the emergency department. In the absence of availability of these types of anticoagulant-preserved blood, with additives, it is reasonable and probably wise to maintain a small stock of whole blood in the blood bank or in the emergency department in an appropriate, blood bank–controlled refrigerator. Parenthetically, if blood for transfusion is stored anywhere in the medical complex outside the blood bank, it is very important that it should be under the direct control and supervision of the blood bank personnel to ensure adequate quality control. It has been widely recognized that, if the blood bank abdicates this responsibility, storage probably will not be carried out properly.

If whole blood is used for emergency resuscitation, it should be used volume for volume relative to blood loss. Because it is nearly impossible to estimate acute blood loss volume in real time, the size of the loss is merely a guess and depends to some extent on the experience of those trying to make that guess. Because of this, as well as other factors, the use of whole blood in such clinical settings is not infrequently associated with volume overload. It is very difficult to judge fluid volume status in a rapidly fluctuating clinical situation unless one has the luxury of having continuous central intravascular pressure measurements. Even this approach is not necessarily foolproof.

Red Blood Cells

This component is essentially the red blood cells that remain after plasma and platelets have been separated from whole blood. The number of red blood cells in this component is the same as the number in the bag of whole blood from which it was derived. The concentration of red blood cells will vary according to the amount of residual plasma in which they are suspended or the volume of additive solution (e.g., AS-1, AS-3) added to the red blood cell bag. The volume of a bag of red blood cells is approximately 250 mL and that of a bag of red blood cells preserved with AS-1 is 330 mL. Other modifications, such as the removal of white blood cells or freeze/thaw/washing, may decrease the volume to as low as approximately 180 mL.

Unless platelets or white blood cells have been selectively removed from the donated whole blood, they will be contained in the final product. After 2 to 3 days of storage at 4°C, there is little, if any, remaining function in the platelets or granulocytes, and they progressively disintegrate with further storage. The degenerating granulocytes in particular contribute to the development of microaggregates in the bag.

The single indication for transfusion of red blood cells is the need for increased oxygen delivery to the tissues. In situations of acute blood loss with intravascular volume deficit, red blood cells provide the means of oxygen delivery, and crystalloid or colloid solution supplements provide the volume. When there is symptomatic chronic anemia unresponsive to various hematinics, the use of red blood cells may be reasonable and appropriate if the risks of withholding transfusion outweigh those of giving it. In other words, the mere existence of anemia or even of the symptoms attributable to it is not sufficient grounds for the transfusion of red blood cells.

The transfusion trigger has long been the subject of considerable controversy. For many years, perioperative transfusion practice was guided by the principle that a hemoglobin value of less than 10 g/dL was an indication for transfusion. Although this simple rule was almost universally accepted and applied, little hard evidence existed to support it other than calculations that indicated a potential for impaired oxygenation of tissues. These data usually did not take into account the appropriate calculation for cardiac output changes, oxygen extraction, or changes in the hemoglobin/oxygen affinity curve (28).

It now is generally accepted that most healthy humans do not experience a compensatory significant increase in cardiac output until the hemoglobin value is below 7 g/dL. Clearly, patients with conditions that compromise their cardiovascular or cerebral blood supply may require therapy before their hemoglobin value reaches this level. Therefore, each patient must be evaluated individually with regard to the effects of anemia before the decision to transfuse is made.

It is also important to note that there is no evidence that wound healing is impaired until a hematocrit value of 15% to 17% is reached (28,29). Minor postoperative anemia clearly is not an indication for transfusion on the basis of wound healing or, indeed, of infection (29). On the contrary, there is some evidence that perioperative transfusions may be associated with an increased incidence of infection (30).

Patients with nonsurgical chronic anemia, such as the anemia associated with chronic renal failure, adapt remarkably well to levels of hemoglobin in the range of 5 to 7 g/dL and can perform everyday duties reasonably well at those levels. It is standard practice to transfuse these patients when the symptoms of the anemia are beginning to incapacitate them or when new or more ominous symptoms appear, such as angina pectoris or central manifestations of anoxia. It is clear that a gradual decline or long-term decrease in hemoglobin level is tolerated by patients much better than a more acute onset.

When the cause of anemia is being investigated, the blood samples for the laboratory studies should be obtained before transfusing the patient. The increase one expects in the hemoglobin value after the transfusion of red blood cells depends on many factors, including the patient's weight and hemodynamic stability and the presence or absence of active bleeding. In a stable, nonbleeding, 70-kg patient, one would expect to see an

increase of approximately 1.5 g/dL or approximately 3% to 4% in hematocrit value per unit transfused. It is necessary to allow an hour or so for intravascular equilibration before the posttransfusion hemoglobin level is determined. Patients with chronic anemia often have an expanded plasma volume, which puts them at particular risk of volume overload, especially if whole blood is transfused rather than red blood cells.

Modified Red Blood Cells: Leukocyte-Reduced Red Blood Cells

Red blood cells can be rendered leukocyte poor by various techniques, including centrifugation to separate the buffy coat or filtration to remove leukocytes. Each method has its intrinsic advantages and disadvantages, and the method chosen may depend on the indication in a given patient. Meryman et al. (31) compared 13 methods of preparing this product, all based on centrifugation, automated red blood cell washing, or freeze/thaw techniques. These authors concluded that centrifugation methods removed 65% to 87% of the leukocytes, but to achieve the higher percentage, one had to use 6- to 10-day-old blood and remove 90 mL of cells, which decreased the hemoglobin content by approximately 30%. Automated washing or freeze/thaw techniques removed as much as 96% of leukocytes and between 5% and 10% of the hemoglobin.

The two main indications for the use of leukocyte-reduced red blood cells are (a) transfusion of patients with a history of febrile nonhemolytic transfusion reactions and (b) prevention of the development of antileukocyte antibodies and platelet refractoriness. More than 25 years ago, Perkins and colleagues (32) first clearly demonstrated that the granulocytes were primarily responsible for febrile nonhemolytic reactions and that this reaction was dependent on both the dose and the rate of administration of the granulocytes. In that classic study, the patients responded, with a minimal degree of temperature elevation, to as few as 0.25×10^9 leukocytes. It is generally believed (33) that the majority of patients with a history of febrile nonhemolytic reactions and antileukocyte antibodies can be transfused satisfactorily with red blood cell preparations containing 0.3×10^9 leukocytes. A much greater degree of removal of contaminating granulocytes can be achieved by various filtration techniques, which also have the advantage of removing fewer red blood cells from the unit.

For a red blood cell product to be designated leukocyte poor, 70% of the white blood cells must be removed, and only 70% of the original red blood cells need to remain (as much as 30% loss of red blood cells).

At the Mayo Clinic (Rochester, MN), the upright spin method yields approximately 80% red blood cell recovery and 80% removal of white blood cells; with nylon-wool filtration, the values are 90% and 95%, respectively. Looked at in another fashion, one unit of upright-spin, leukocyte-poor red blood cells exposes the patient to as few as 0.6 to 0.8×10^9 white blood cells, and the filtered product exposes the patient to as few as 0.06 to 0.08×10^9 such cells.

Most patients can receive upright-spin, leukocyte-poor red blood cells without subsequent reactions, but some break through at this dose and then require filtered red blood cells. After several years of experience with the transfusion of many hundreds of such patients, The Mayo Clinic has yet to see a highly sensitized patient have a febrile nonhemolytic reaction to filtered red blood cells. The patients who are most likely to have febrile nonhemolytic reactions are multiparous women or previously transfused patients (33).

Numerous studies examined the role of antigranulocyte antibodies in these reactions; although there is a correlation, the data are not clear-cut. This probably reflects the difficulty in making the clinical diagnosis with any degree of consistency and confidence because of the confounding factors in many patients. Underlying clinical conditions such as sepsis, autoimmune disease, and medications make this a diagnosis by exclusion. A second reason for lack of clear-cut correlation is the fact that many different tests have been devised for the detection of antibodies to leukocytes, and they probably do not measure antibodies to the same antigenic targets. A few of the tests that have been used are leukoagglutination (slide), leukoagglutination (microcapillary tube), lymphocytotoxin (complement-dependent dye exclusion), granulocyte membrane (fluorescence), and granulocyte membrane (radioimmunometric).

Another factor that plays a role in deciding what is the most reasonable approach is the shelf life of the product. If one uses either centrifugation or a spin/filter, closed-system method, the normal shelf life of the red blood cells is retained. Conversely, if one uses a system that requires breaching the integrity of the system (such as washing or non-in-line filters), the subsequent shelf life becomes 24 hours. Clearly, the latter component is useful only if transfused immediately after preparation. The need to maintain an inventory of leukocyte-poor products in the blood bank with a reasonable shelf life is obvious (34).

Filtration through a white blood cell filter that can be used at the bedside (or in the blood bank immediately before transfusion) has been advocated. The use of such a filter resulted in 87% red blood cell recovery and 6.1×10^6 residual white blood cells (35). Problems noted by those authors were prolonged transfusion time and poor red blood cell recovery if only one unit was infused through the filter. Another type of filter that can be used in this setting is the microaggregate filter, originally designed for the removal of microaggregates from blood before massive transfusion (an indication that is highly controversial). These filters have been used successfully in the blood bank to prepare leukocyte-poor red blood cells just before their infusion (36).

Another recent advance in the preparation of leukocyte-poor red blood cells is the development of a closed-system method that combines centrifugation and filtration steps and the removal of microaggregates (which enhances red blood cell storage) (37). This method preserves the shelf life of 42 days for the product and apparently yields 98% white blood cell removal, i.e., 10^8 cells remaining. This approach is significant because it seems to combine efficiency of white blood cell removal with ease of operation and logistic convenience.

The second indication for the use of leukocyte-poor red blood cells is the prevention of the platelet refractory state. In their landmark paper in 1957, Brittingham and Chaplin (38) recognized the role of leukocytes in febrile nonhemolytic transfusion reactions, and subsequent studies indicated that patients

with preexisting alloimmunization against HLA antigens had decreased graft survival (39). This led to a significant increase in the use of frozen/thawed red blood cells for transfusion in dialysis patients. In 1973, Opelz et al. (40) presented the apparently paradoxical data that transfused patients had better renal graft survival than untransfused patients. Opelz and Terasaki published (41) the definitive data in 1976 that indicated a previously unsuspected effect of transfusions, i.e., immunosuppression.

The fact that transfusions generally led to alloimmunization was not in dispute. Platelet refractoriness was assumed to be owing to exposure of the patients to repeated doses of platelets with their abundant endowment of HLA class I antigens (HLA A, B, and C antigens). However, data began to accumulate suggesting that the platelets themselves did not evoke the immune response, but rather the contaminating white blood cells in the platelet infusions did. Interestingly, these data were analogous to data implicating passenger donor white blood cells in solid organ grafts as the culprits in evoking transplant rejection phenomena (42).

Because of the enormous strides in the past decade in the therapy of patients with hematologic malignancies, these patients achieve more remissions and survive longer. Each attempt at inducing remission usually includes transfusion support with platelets and, sometimes, red blood cells. Clinicians are increasingly faced with the specter of platelet refractoriness. It is logical to try to prevent this rather than to merely respond to its presence by using expensive, HLA-matched apheresis platelets or, occasionally, cumbersome platelet cross-matching systems.

Prevention has been achieved by the drastic reduction of the white blood cell content of red blood cell and platelet preparations used for transfusions. Again, simple centrifugal removal of the buffy coat often suffices to prevent the occurrence of febrile reactions in previously sensitized patients. Prevention of sensitization seems to require that each transfusion deliver less than 5×10^6 leukocytes. Meryman (43) summarized the supportive data for this statement most impressively. In practical terms, to prevent alloimmunization, one must be prepared to process red blood cells (and also platelets) to remove 95% to 98% of the leukocytes to achieve the level of 5×10^6 residual cells. This requires either frozen/thawed red blood cells or the newer filtration methods, which can achieve this level of white blood cell removal with minimal red blood cell loss. Sniecinski et al. (44) showed the value of this approach in a thorough prospective study. In particular, the use of these filters permits patients to have more transfusions over a longer period before they become sensitized.

It is difficult to find hard data on the financial impact of this approach, but keep in mind that perhaps 50% of patients are destined to become refractory if one makes no special effort to prevent it. A policy of deliberately using only leukocyte-poor products for all patients with hematologic malignancies would be expensive (because of the cost of filters and labor). Conversely, however, the cost of providing platelet support to a highly refractory patient who is consuming vast quantities of platelets (even the expensive HLA-matched variety) is enormous.

Indirectly, the transfusion of many units of blood and components exposes patients to various other complications, such as hepatitis and cytomegalovirus (CMV) infection. Meryman (43) pointed out that “it is hard to argue that the prevention of HLA alloimmunization by filtration is an extravagance.”

Platelet Transfusions: Random-Donor Platelet Concentrates

As described above, platelets generally are separated from donated whole blood by a series of centrifugation steps designed first to separate platelet-rich plasma from red blood cells and then to concentrate the platelets in the platelet-rich plasma to decrease the volume of the product and to maximize plasma availability for other purposes. Platelets also can be obtained from the buffy coat, although this method is not widely used (45). The standards of the American Association of Blood Banks state that platelet preparations from whole blood must contain a minimum of 5.5×10^{10} platelets in at least 75% of units. Platelet preparations obtained by cytapheresis must contain 3×10^{11} platelets in at least 75% of units (46). The platelets (either variety) must be stored in sufficient plasma and agitated appropriately at a temperature that permits the pH to remain above 6.0 in all units.

Each bag of platelet concentrate contains 50 to 60 mL of platelets and plasma. Except in small children, platelet concentrates are generally transfused as pooled products from six to eight units of platelet concentrate. This number is often utilized because each unit of platelet concentrate will increase the circulating platelet count by approximately 7,500/μl in an average 70-kg person, and the target platelet count in many patients is 50,000/μL. Although ABH antigens can be detected on platelets in small amounts, in general, ABO matching seems to be clinically important only in alloimmunized patients who get better increments from ABO-matched or -compatible platelets.

The use of platelet transfusions has increased dramatically in the past decade, but not all this increase has been medically justifiable. The success in bone marrow and hepatic transplantation and in leukemia/lymphoma therapy has resulted in considerably more patients who truly need and benefit from platelets, but some experience has shown that the use of platelets in some clinical settings has tended to become arbitrary, with little or no scientific basis for the practice.

Unfortunately, the demands by some accreditation agencies for simplistic algorithms of treatment tend to exacerbate the problem. For instance, if the accrediting agency recommends that there be a transfusion guideline that indicates that a platelet count of fewer than 50,000/μL is one of the indications for platelet transfusion, this may be translated in the minds of some clinicians into a rigid rule that all patients with counts below this value need platelet transfusions. The medicolegal atmosphere may contribute to the problem. Common sense dictates that each patient should be assessed carefully before deciding to transfuse, and this caveat applies to all transfusion settings, not just those involving platelets.

The indications for platelet transfusion are significant thrombocytopenia or thrombopathy. The first question in this regard concerns the platelet count. There are excellent data to show a good general correlation between platelet count and the likelihood of spontaneous bleeding. In general, if the count exceeds 50,000/μL, the patient is unlikely to bleed even intraoperatively and should not be given a prophylactic platelet transfusion unless there is additional evidence of platelet dysfunction (47). Significant risk of spontaneous hemorrhage appears when the

platelet count is below 20,000/μL, and it is greatest when the count is 5,000 or lower (47).

The template bleeding time is sometimes used as an indicator of platelet function, but this test should be interpreted with considerable caution. Bleeding time tests can be performed in several ways, and the method used must be carefully standardized and normal ranges established in each center. Notwithstanding such precautions, there is a poor correlation between the bleeding time and clinical hemorrhage. In one study of patients who had normal clinical hemostasis after cardiopulmonary bypass, the bleeding time was 1.8 times the mean preoperative value (48). It is generally believed that the bleeding time correlates with clinical bleeding only when the value is more than twice the upper limit of normal (49). Even then, the correlation is not strong.

Many leukemic patients have platelet counts between 5,000 and 20,000/μL while undergoing chemotherapy or radiotherapy, and clinicians must carefully balance the dangers of spontaneous bleeding against the problems of alloimmunization (and perhaps disease transmission) generated by exposure to platelets. If a patient has a platelet count of 20,000/μL but has petechiae, gastrointestinal bleeding, or other bleeding manifestations, transfusion may be appropriate. Conversely, many experienced hematologists will not give platelets to a stable patient with levels as low as 5,000/μL when there are no signs of bleeding.

Platelets in Special Situations

General Surgery

Although there have been no controlled, prospective studies on the appropriate transfusion trigger for patients undergoing surgery, it is generally agreed that, for adequate surgical hemostasis, the platelet count should be greater than 50,000/μL in most patients. Judgment must be used in deciding when to transfuse because factors other than the platelet count may be important. For example, platelet dysfunction may exist alone or with thrombocytopenia. The following are associated with platelet dysfunction (50): uremia, fibrinogen/fibrin degradation products, ingestion of aspirin, and medications such as semisynthetic penicillins. In each of these situations, patients may bleed despite having a platelet count well above that normally associated with platelet-related bleeding problems. Therefore, in patients who require surgery, the presence of any of these situations is grounds for further coagulation investigation, including a bleeding time test. If the bleeding time is at least twice the upper limit of normal, prophylactic platelet transfusions might be reasonably considered if the underlying problem could not be corrected before surgery.

There are few objective data to demonstrate the effectiveness of platelets given to patients with diffuse intravascular coagulation and fibrinolysis. Some might even argue that it is tantamount to “adding fuel to the fire” to give platelets in this situation. Every effort should be made to correct the underlying cause rather than to attempt to replace the cells (or coagulation factors) that are being consumed.

Although there are no definitive data to support the practice, the preoperative transfusion of enough platelets to correct the bleeding time or to bring it to below twice normal is logical. One might also use the same reasoning in dealing with patients thought to be at risk of central nervous system bleeding, where even small hemorrhages can have permanent and devastating consequences (51).

Cardiac Surgery

Because cardiopulmonary bypass devices have long been recognized to cause physical damage to circulating platelets, patients undergoing this procedure have been studied extensively. Most have been found to have mild thrombocytopenia and platelet fragmentation with prolongation of the bleeding time. Thus, routine platelet transfusions might seem logical in these patients. However, in multiple retrospective and prospective studies of such patients, several points consistently emerge: (a) the platelet problems are usually self-correcting within 24 hours postoperatively, (b) postoperative bleeding is most often owing to inadequate surgical hemostasis and less often to thrombocytopenia or thrombopathy, and (c) prophylactic platelet transfusions do not seem to diminish blood loss. Therefore, there is no justification for routine prophylactic platelet transfusion after cardiopulmonary bypass (52).

There are clearly some situations in which such transfusions are probably reasonable. Repeat thoracotomy is often associated with oozing from the newly disrupted scar tissue, and this type of bleeding can be difficult to control with cautery or sutures. There are no good data on the effectiveness of platelet transfusions in these patients, but experienced cardiac surgeons seem convinced of their usefulness when surgically correctable bleeding does not seem to be present and the pattern of bleeding suggests capillary oozing.

At the Mayo Clinic, platelet use in cardiac bypass cases has increased from 33% of cases in 1979 to 38% in 1988. This may be owing to an increase in more complicated congenital anomaly repairs, requiring the use of porous artificial membranes for reconstruction. In such cases, it appears that the platelet transfusions play a role in rapidly plugging the pores of the material and halting oozing from its surface. It also may reflect greater use of aspirin-containing drugs or drugs such as streptokinase or tissue plasminogen activator.

Sepsis

It has long been recognized that sepsis seems to be associated with thrombocytopenia in some patients. Likewise, thrombocytopenic septic patients often respond suboptimally to platelet transfusions. Indeed, sepsis is one of the differential diagnoses in a patient who is unresponsive to platelets. The suboptimal transfusion response may be caused by disseminated intravascular coagulopathy, in which platelets are consumed, but it is also seen when sepsis exists but there is little or no evidence of consumptive coagulopathy. Elegant studies by Kelton et al. (53) and Tate et al. (54) demonstrated an increased platelet-associated IgG level in septic patients. Presumably, this leads to an increased rate of platelet destruction and removal by the macrophage/monocyte system. In addition, there is evidence that patients with septicemia have autoantibodies to platelet antigens, which accounts for the increase in platelet-associated IgG. The target antigens are postulated to be crypt-antigens exposed by the septicemia (55).

The logical approach to the thrombocytopenic patient with sepsis is aggressive treatment of the underlying sepsis rather than directing primary attention to the thrombocytopenia. However, there are clinical situations in which one must address both problems simultaneously, and it should be remembered that as long as sepsis is present, posttransfusion increments may be significantly suboptimal and HLA matching or other strategies will not correct the problem. Efforts to cross-match platelets in this situation generally are futile and may distract one from the chief problem—the underlying sepsis.

Uremia

Patients with uremia have varying degrees of platelet dysfunction, which prolongs bleeding time. Since the 1950s, the defect in the availability of platelet factor III in uremia has been known, and subsequently, a defect in platelet aggregation was noted. These measurable defects were often noted to be mild and out of proportion to the prolongation of bleeding time and the degree of clinical bleeding. When dialysis was found to reverse the functional abnormalities of the platelets, various substances, including urea, phenols, and guanidinosuccinic acid, were blamed for the platelet abnormalities, but each claim was somewhat controversial (56).

More recent studies concentrated on the role of coagulation factor VIII and, more particularly, its factor VIII:von Willebrand factor (VIII:VWF) moiety. Alteration of the platelet-membrane receptor for VIII:VWF seems to be a reasonable explanation for the prolongation in bleeding time and the platelet adhesiveness defects seen in uremia (57).

Several clinical studies demonstrated that the use of diamino-8-D-arginine vasopressin (DDAVP) can shorten the bleeding time in uremia (58). This seems to act by causing the secretion into the plasma of increased numbers of the large multimers of VIII:VWF, which may overcome the binding defect in uremic platelets. This property of DDAVP has led to its popularity in treating uremic patients before urgent surgical procedures when there may not be time for dialysis. The lack of serious side effects and the freedom from disease transmission problems have contributed to the popularity of DDAVP rather than the use of platelet transfusions for this purpose (58). The effects of DDAVP are relatively short-lived (hours), and repeated treatments can lead to tachyphylaxis.

Leukocyte-Reduced Platelets

The indications for this product are essentially the same as those for leukocyte-depleted red blood cells: (a) to ameliorate the symptoms of febrile nonhemolytic reactions in sensitized patients, (b) to prevent development of the refractory state, and (c) to decrease exposure to some viruses known to reside in white blood cells (e.g., CMV).

Various filtration methods have been described recently, and these have achieved a remarkable level of leukocyte removal (as much as 99%). This level of removal gives a final product of pooled platelet concentrates with approximately 5×10^6 residual white blood cells per pooling bag. This degree of filtration is sufficient to prevent febrile responses in most sensitized patients. The platelet loss with these filters has been approximately 10% in recent years. Platelet loss greater than this leads to the need for transfusion of additional units of platelets to compensate. Cotton-wool filters remove 100% of granulocytes, 95% of monocytes, 90% of B lymphocytes, and 85% of T lymphocytes (59). Cotton-wool filtration causes almost no activation of platelets and only minimal decreases in postinfusion survival that would be of little or no consequence in a clinical setting (60).

My own experience over several years is that cotton-wool filtration is very effective at preventing febrile reactions in patients with a history of repeated, debilitating reactions to unmodified red blood cell or platelet transfusions. This experience has been corroborated by others (61) who found such an effect when the white blood cell content was decreased to less than 10^8.

Polyester fiber filters remove almost as many white blood cells but are more expensive (62). The use of a special pooling bag for centrifugal removal of the buffy coat eliminates 93% of the white blood cells, so that febrile reactions are prevented (63). The value of this system is its ability to accomplish the white blood cell removal in a closed system shortly after the blood is obtained from the donor. This early removal should improve red blood cell and platelet storage pH while maintaining the original shelf life of all components. Whether this simple and attractive system removes enough white blood cells to prevent alloimmunization is still open to question. The immunizing dose of leukocytes is unknown, but some data suggest that 1×10^8 allogeneic leukocytes may be needed (64).

As yet, the data are insufficient to ascertain whether the use of filtered leukocyte-poor platelets will significantly diminish the occurrence of CMV or other leukocyte-borne viral infections in susceptible hosts. However, preliminary data strongly suggest that this is the case with filters capable of 3 to 4 log white cell reduction.

Other Platelet Preparations

Platelets can be washed to remove plasma proteins for patients with severe allergic reactions. A semiautomated method has been described that removes 96% of the plasma protein and yields 90% platelet recovery (65).

Platelets can be successfully frozen and then thawed for transfusion as long as 3 years later with approximately 25% loss of platelets in the process. This is usually accomplished by using 5% to 10% dimethyl sulfoxide (DMSO) and an ultralow mechanical freezer capable of maintaining temperatures of -80° to -90°C. If storage is at -120°C, platelets can be maintained for as long as 4 years. This is particularly useful for the storage of autologous platelets from alloimmunized patients who achieve leukemic remission with its attendant rebound thrombocytosis (66). It also has been used successfully to treat an infant with neonatal isoimmune thrombocytopenia with its mother's frozen platelets, which had been collected early in pregnancy (67).

Platelets in Massive Transfusion

Patients who undergo massive transfusions within a short period (e.g., one or more blood volumes in 12 to 24 hours) have well-documented coagulation defects resulting from the loss of blood,

the inadequacy of the physiologic response to the rapid replacement of the lost platelets and labile coagulation factors, and the transfusion of blood deficient in these same moieties by virtue of storage losses (68). Because the degree of deficiency is directly related to the duration of storage of the blood, such deficiencies vary among units of blood chosen randomly for transfusion. Although most blood banks generally try to transfuse their oldest units first, during massive transfusion situations, such an elective transfusion practice may be inappropriate.

There is also considerable variation in the degree of clinically significant coagulopathy seen in patients who receive massive transfusions. These factors make it difficult to establish firm correlations between the degree of coagulopathy and the quantities of blood infused. It should be recalled that the platelets may have been removed from the donated blood for processing into platelet concentrates, although this is not a realistic concern because platelets no longer function after 48 hours in stored whole blood at 4°C. Coagulation factors V and VIII are at approximately a 50% level after 1 week of storage. Therefore, massive infusion of stored blood is going to do little for the resultant dilutional thrombocytopenia or for depletion of labile plasma clotting factors.

Logically, one might expect that the amount of blood lost, the degree of thrombocytopenia observed, and the resultant microvascular clinical bleeding would be closely correlated. In several well-conducted studies, this was not the case. Also, prophylactic platelet infusions based on a predetermined replacement formula have not proved to be effective in altering clinical bleeding patterns or in reversing prolongation of bleeding time (69,70). The authors of these studies concluded that such indiscriminate platelet transfusion therapy is unjustified.

Single-Donor Platelets (Apheresis Platelets)

The American Association of Blood Banks Standards stipulate that a bag of platelets acquired by apheresis should contain a minimum of 3×10^{11} platelets in at least 75% of units tested. These platelets must be stored suspended in sufficient plasma so that the pH at the temperature of storage is 6.0 or greater at the end of the allowed storage time. In the newer plastic bags, this storage time is now up to 5 days. The collection of platelets by apheresis achieved popularity because of the recognition that platelet refractoriness was a complication of repeated exposures to random-donor platelet concentrates. This lack of responsiveness occurs in some recipients after only a few exposures; in others, it is delayed or does not appear at all.

In their classic study, Howard and Perkins (71) demonstrated that approximately 30% of patients never became refractory despite repetitive exposures. These so-called nonresponders often have malignancies, including those of the hematopoietic or lymphoid system, and their failure to become refractory raises intriguing questions about their underlying disease, their treatment regimens, or their immunogenetics (72).

Those who do become refractory display great individual variability in both the number of platelet exposures and the time taken to become refractory. The state of platelet refractoriness is said to exist when a patient has progressively smaller increases in platelet counts after platelet transfusions. Many factors may contribute to this state, and each must be considered in the approach to a clinical problem. The most common factors are listed in Table 71.1.

TABLE 71.1. FACTORS CONTRIBUTING TO PLATELET INCREMENT OR FUNCTION AFTER TRANSFUSION

Platelet Factors	Patient Factors
Method of collection	Size of patient
Method and duration of storage	Medications being taken
Presence of appropriate agitation	Sepsis
pH	Splenomegaly
Number of platelets infused	Disseminated intravascular cogulopathy
Type of anticoagulent	Bleeding
Donor taking antiplatelet medication	Circulating immune complexes
	Alloimmunization

A strong association has been found between the presence in the recipient of alloimmune lymphocytotoxic antibodies (often with HLA class I specificity) and failure to achieve the expected posttransfusion increment when other factors (such as those listed in Table 71.1) could be ruled out (71). There is abundant evidence that HLA class I antibodies are responsible for most alloimmune refractory states. This was first demonstrated by Yankee and colleagues (73), who showed that matching for what we now call HLA class I antigens was associated with an excellent response in previously refractory patients.

Although these studies clearly outlined a therapeutic approach for alloimmunized patients, they also brought to light some problems that still plague us. First, the diagnosis of the alloimmunized state is sometimes quite difficult because patients in whom it may be suspected often have one or more of the factors that mimic the alloimmunized state (Table 71.1). Second, when a patient is considered to be refractory and is placed on a regimen of HLA-matched platelets, it is virtually impossible to persuade clinicians to resume the use of unmatched platelets. Thus, the patient may be locked into a dependence on expensive, HLA-matched platelet support. The original diagnosis of alloimmunization may have been erroneous or the patient may have been truly alloimmunized but may have lost the antibodies with time.

There have been occasional reports of patients, particularly those with extensive malignant disease, losing their antibodies (74). This phenomenon has been seen in patients who were not given additional platelets after induction of the alloimmune state (71), but it also has been seen in patients who continued to receive only cryopreserved, autologous platelets or HLA-matched platelets (74). Strangely, it has even been seen occasionally in patients who continue to be transfused with random-donor platelets unmatched for HLA (75,76).

These observations suggest that although a patient may be known to have HLA antibodies concomitantly with a refractoriness to platelets, it should not be assumed that this patient always will be refractory to platelets that are not HLA matched. In a patient with relapse of a malignant disease, it is probably worth giving a trial transfusion of random platelets rather than automatically ordering HLA-matched, apheresis platelets and thus having the associated expense and logistic problems.

Another approach for the alloimmunized patient has been the use of repeated plasma exchanges to decrease the antibody titer. In one study, this approach was used during the preparatory chemotherapy and radiation regimen before bone marrow transplantation (77). It had mixed results, but there was a correlation between apparent success and the presence of preformed antibodies before the plasma exchanges. Those who responded had at least two exchanges before showing some response.

This approach cannot be applied to all alloimmunized patients because of the expense and time involved and the mixed success of the program. Because of the finding that some antigens such as HLA may be adsorbed to the platelet membranes, it has been suggested that these antigens or others may be eluted from the platelets during storage and the platelets might lose some of their antigenicity as a result. Although such elution can be shown in experimental situations, it has been demonstrated that routine platelet storage is associated with no clinically relevant diminution in the antigenicity of HLA, $Pl^{A\ 1}$, or ABO systems on platelets (78). Another study has shown that although some HLA antigens are taken up by platelet membranes in small amounts, the major portion of HLA class I antigens seems to be integral membrane constituents (79). In assessing the clinical response to platelet infusions, the 1-hour posttransfusion platelet count is a reliable predictor of the need for HLA-matched platelets. It is a useful and timely test that permits early decision making regarding therapy (80). These authors also found that the 10-minute increment also correlated very well with the 1-hour count.

To minimize the use of apheresis platelets, the factors associated with the ineffective response to pooled, random-donor platelets were examined in a recent study of 941 pooled platelet transfusions given to 133 patients with bone marrow failure (81). The major factors, by multiple linear regression analysis, that determined the increment at 1 hour after transfusion were previous splenectomy, marrow transplantation, disseminated intravascular coagulation and fibrinolysis, the use of amphotericin B, hypersplenism with splenomegaly, and HLA antibody grade as measured by the reactions of serum against a panel of four random-donor platelets. A linear relationship was found between increasing antibody grade and decreasing increment. Several other factors, such as concurrent antibacterial therapy, clinical grade of bleeding, and pyrexia, did not correlate as clearly (81).

Although hematologists and blood bankers are well aware of the problems involved in providing adequate and effective platelet support for refractory patients, not all heavily transfused patients become truly alloimmunized and require HLA-matched platelets. Approximately 40% to 50% of patients with leukemia and perhaps 10% of those with solid organ malignancy develop the antibodies in question. As mentioned previously, the underlying disease or its therapy may play some role in this apparent selectivity. Also, some intriguing data on the cellular and subcellular interactions involved in the development of the alloimmune state are emerging (82). It appears that, along with the early alloimmune response, there are signs of lymphoid cell activation, such as interleukin-2 (IL-2) and HLA-DR, on peripheral lymphoid cells and measurable increases in interferon gamma, neopterin, and β_2-microglobulin in plasma. However, such responses are not seen when purified thrombocytes are used, even if challenges are repetitive. Conversely, purified platelets will induce the elaboration of immunoglobulins that can block the Fc receptor and inhibit erythrocyte rosette formation. These immunoglobulins are immunosuppressive, and there is evidence that they are induced by stimuli that exhibit only HLA class I antigens (such as platelets).

If HLA-matched platelets are needed, the best possible responses should be from HLA class I perfect matches. Within a family, there is a 1:4 likelihood of siblings being perfect matches, a 2:4 likelihood of a single haplotype match, and a 1:4 chance of total mismatch. Therefore, when HLA-matched platelets are needed, one should try to use family members as donors because a random sibling has a 3:4 chance of being at least one haplotype matched and parents or children are obligatory one-haplotype matches. Thus, even if HLA typing on the patient or family members cannot be obtained, an intelligent guess can be made to make possible a quick response.

It is not uncommon for HLA laboratories to have difficulty in typing the lymphocytes of patients with leukemia even though there are techniques available for the separation of normal from abnormal lymphoid cells. The patient's peripheral lymphocyte population may simply be overwhelmingly abnormal cells. These abnormal cells often show aberrant and increased sensitivity to the complement generally used for the standard, complement-dependent cytotoxicity testing used for HLA class I typing.

If family members are unavailable or unsuitable for some reason, it usually is necessary to use a pool of blood bank donors who have been HLA-typed for this purpose. When one has to select from the population donors who are matched for a highly polymorphic system such as HLA, one can appreciate the need for a large pool of HLA-typed donors to increase the likelihood of finding a good match. One strives to find the closest match at all times. Fortunately, one of the features of the HLA system is that multiple groups of antigens at the same locus (*A*, *B*, or *C*) are highly cross-reactive with one another. Similarly, patients with these same cross-reactive groups of antigens rarely induce antibodies against one another's antigens. This means that alloimmunized patients will not recognize cross-reactive antigens as foreign and thus will be serologically compatible with these antigens. Clinically, the use of such cross-reactive antigen matches works virtually as well as precise antigen matching. The ability to utilize cross-reactive antigens automatically broadens the pool from which one can select clinically effective platelet donors for a specific patient. The donor pool needed to support a population of alloimmunized patients will vary according to the polymorphism of HLA in the particular ethnic group in question, the number of alternative donors one stipulates for each patient, and the degree of matching one stipulates for support.

There are various estimates of the donor pool needed in different populations. For instance, Takahashi et al. (83) calculated that a pool of approximately 5,000 typed Japanese donors is sufficient to provide five perfectly (class I) matched donors 80% of the time for Japanese patients, but as many as 25,000 American donors will be needed to provide similar support to the ethnically diverse American patient population.

Another factor in determining donor pool size is the tremendous variability in HLA class I antigens on platelets from person

to person; this applies to all class I antigens, but particularly to HLA B12 and its subtypes (84). This variability may account for the remarkable variation in incremental responses seen in patients transfused with platelets from donors with the same apparent degree of mismatching. This antigen may be selectively mismatched for patients with excellent results, and this further increases the available pool for alloimmunized patients because B12 is common in the American population (approximately 25%) (84).

The diallelic Bw4/Bw6 system within HLA class I can also be used effectively for matching because many patients who are alloimmunized in fact have antibodies to so-called skeletal antigens or cross-reactive groups (85). Matching for *C*-locus antigens is not generally useful, and these antigens seem to be more deeply embedded in the cytoplasmic membranes and more sparsely distributed than *A*- or *B*-locus antigens (86). Certain cross-reactive groupings such as A3, A1 and A11, or B17, B21, B5, and B15 do not seem to be effective clinically in terms of transfusions (87).

Because of the variable responses to HLA-matched platelets and the expenses and difficulties encountered in their acquisition, it is sometimes debated whether they should be used. It would be foolish to downplay these facts, but it would also be unwise to ignore the effect of continuing to transfuse incompatible or unmatched products. First, this type of transfusion to an alloimmunized patient is highly unlikely to result in clinically meaningful increments. Second, the infusion of unmatched platelets is likely to be followed by increasingly severe febrile nonhemolytic reactions, which are debilitating and frightening to the patient and may confuse the clinician who must try to differentiate such reactions from the onset of sepsis. Third, there are rare instances of severe, prolonged neutropenia after infusion of incompatible platelets (88). In these cases, the median granulocyte count 20 hours after transfusion was decreased 30% and remained decreased for as long as 4 days. These reported cases involved patients with aplastic anemia in whom the normally expected rebound granulocytosis did not occur.

Because of the prevalence of significant platelet refractoriness owing to alloimmunization, various strategies have been devised to maximize the likelihood of obtaining satisfactory incremental responses to platelet transfusion. The most effective strategy is HLA matching, but a considerable body of data supports the concept of platelet cross-matching either alone or in concert with HLA matching. Virtually every type of antiplatelet antibody test has been evaluated: indirect tests such as lymphocytotoxicity tests and direct tests such as immunofluorescence, enzyme-labeled immunosorbent assays, solid-phase red blood cell adherence, radiolabeled antiglobulin, and platelet aggregometry (89,90,91,92 and 93). Nearly all studies demonstrated correlations between HLA matching and increments, between negative tests and increments, and between the degree of matching and negative tests. The data have to be viewed with some circumspection because of the obvious difficulty in finding patients who lack any of the confounding factors such as sepsis.

The concept of an inexpensive, predictive cross-match is appealing because it can be used instead of the expensive pool of HLA-matched donors. This is feasible only if the test can be performed on platelet aliquots stored for long periods of time (92). It could even be used to select groups of random-donor platelet concentrates on an *ad hoc* basis, making use of each day's supply of regular homologous donors. The solid-phase red blood cell adherence method seems to lend itself to this approach, according to early results (94).

Although ABH antigens are present in small amounts on platelet surfaces, in general they do not significantly influence the posttransfusion increments. Therefore, it is not necessary that platelets always be ABO compatible. However, ABO antigens can interfere with sensitive cross-match testing and, in alloimmunized patients, mismatching for ABO may influence increments (95). Lee and Schiffer (95) suggested that one should routinely try ABO-compatible, random-donor platelets before assuming that a patient needs HLA-matched platelets.

It has been suggested that, even though only approximately 50% of leukemic patients become truly alloimmunized, one should begin therapy with HLA-matched apheresis platelets to prevent the development of the alloimmunized state. This question was discussed lucidly by Schiffer and Slichter (96), who concluded that this approach was not cost-effective.

A different approach to the prevention of alloimmunization is the exclusive use of leukocyte-depleted blood components, including platelets. These could be filtered, random-donor platelet concentrates or apheresis products. In one study, the use of leukocyte-poor platelets was associated with platelet refractoriness in only 9% of recipients (97). In another study, 15% of recipients of filtered, leukocyte-poor platelets were alloimmunized (44). The use of UV irradiation to pretreat platelets and other blood components in an effort to render them nonimmunogenic shows some early promise (98), and the development of practical and effective radiation delivery systems probably will revolutionize the transfusion of all patients in the future.

Granulocyte Transfusions

Patients who are neutropenic are susceptible to bacterial and fungal infections, and the degree of neutropenia correlates directly with the frequency and severity of such infections (99). Combating infection with granulocytes was first reported in 1934 by Strumia (100), who used a patient with chronic myelogenous leukemia as the donor. It was reasoned that such patients had a greatly expanded pool of functioning, albeit abnormal, granulocytes. However, it was not until modern cell separators were developed in the 1960s that granulocyte transfusions became part of the routine armamentarium of the hematologist. At that time, it became routine to utilize healthy normal blood donors as a source of the cells, and several techniques were used to boost the yield of granulocytes from them.

Intravenous or intramuscular steroid injections increased apheresis yields by altering the extent of margination of white blood cells, thereby increasing the pool of harvestable cells. With this technique, granulocyte units from normal donors contain an average of 1×10^{10} granulocytes. Because granulocytes are essentially tissue cells and spend only a small fraction of their life span in the circulation, it is impossible to assess their effectiveness after transfusion by evaluating posttransfusion increments. The average adult produces approximately 1×10^{11} granulocytes per day, but only approximately 2×10^{10}, or 20%, circulate.

During severe infections, particularly bacterial, the marrow can respond by increasing the production rate perhaps as much as 10-fold. This increased rate can be maintained for many days or even weeks. If one assumes that such numbers of granulocytes are produced because they are needed to combat microbes, the folly of trying to acquire similar numbers for transfusion from donors and to maintain tissue levels is evident.

This central dilemma, perhaps more than any other, has been a stumbling block to the widespread, effective use of granulocyte transfusions in infected adults. Similarly, the use of other end points for evaluation is plagued by problems. For instance, virtually all treated patients receive large doses of antimicrobial agents and may undergo surgical management of localized septic foci. If defervescence occurs or if local or systemic signs of infection or inflammation decrease, it is often difficult to attribute this to the granulocyte transfusions. Granulocyte transfusions have lost favor in recent years largely because of the availability of new families of effective antimicrobial agents and perhaps also because of improvements in the speed and accuracy of bacterial identification and antimicrobial sensitivity testing.

Because of these difficulties in evaluating efficacy, it seems reasonable to pose several practical questions and to try to provide answers. What is the ability of transfused granulocytes to migrate to sites of tissue infection? How adequate is their function? What are the indications for their use (type and duration of infection, organisms most responsive, most effective treatment regimens, matching requirements, and prophylactic use)? What collection and storage systems are best?

In nonalloimmunized patients, transfused granulocytes have normal circulation characteristics, phagocytose bacteria *in vivo* and migrate to sites of infection. There seems to be a general consensus that these transfusions are indicated for severely and persistently neutropenic patients (fewer than 250 granulocytes/μL) with infection caused by gram-negative bacteria unresponsive to adequate appropriate antimicrobial therapy for at least 48 hours (86). The data supporting these indications were presented convincingly by McCullough (101). There are no good data to support the use of granulocytes in viral or fungal infections or in patients who do not have documented gram-negative sepsis.

Earlier studies indicated that filtration techniques gave a higher yield of granulocytes but that the granulocytes had measurable defects in function. Transfusion of filtration granulocytes ceased when, in addition to widespread severe, recurrent febrile reactions, cases of priapism were reported. The following comments deal only with granulocytes acquired by apheresis and centrifugation. The use of granulocytes tagged with indium 111 permitted elegant studies of the migration and localization of transfused cells (102) and the impact of recipients' antibodies (103) (including ABO) on these phenomena (104). The transfusion of small amounts of indium-labeled autologous granulocytes also became a useful tool in localizing sites of inflammation (102).

Granulocyte transfusions should be ABO compatible because of the large number of contaminating red blood cells in the granulocyte bag. McCullough et al. (104) found that although it was assumed that ABH antigens were found on the granulocyte membrane by some techniques, the intravascular recovery, survival, and tissue localization of ABO-compatible and ABO-incompatible granulocytes were the same. There is no good evidence that HLA matching is advantageous. The presence of preformed leukocyte antibodies in the recipient of granulocyte transfusions influences the fate of these cells. The type of antibody test used may have some bearing on the correlation. Granulocyte agglutinating, but not granulocytotoxic or lymphocytotoxic, antibodies decrease granulocyte recovery and half-life of cells, increase their hepatic sequestration, and inhibit migration and localization of cells to sites of infection (103,105). Other studies compared the clinical outcomes of alloimmunized and nonsensitized patients who received granulocytes. The data from at least one such study (106) indicated that patients with antigranulocyte antibodies, as measured by indirect immunofluorescence, had a significantly worse final outcome when treated with granulocytes. The antibodies presumably played a role in inhibiting the desired effect of the granulocytes.

Whether granulocytes should be used prophylactically has not been satisfactorily determined, despite several studies. Some studies were terminated because of unacceptable numbers of patients with severe reactions. From one multicenter, controlled trial of prophylactic granulocyte transfusions in patients with acute myelogenous leukemia, it was concluded that the incidence of reactions (including pulmonary infiltrates in 57%) was too great to warrant the therapy, which had not reduced the incidence of infections or improved bone marrow recovery, remission rate and duration, or survival (107). One particularly lethal complication is diffuse pulmonary hemorrhage which occasionally occurs when granulocytes are transfused to patients receiving amphotericin B, although this is controversial (108,109).

Clift et al. (110) found substantial benefits in bone marrow transplant recipients. Others found no obvious benefit and significant hazards; therefore, routine use cannot be recommended. If granulocyte transfusions are to be used, it seems prudent to give at least 1×10^{10} cells at least once daily for at least 5 days to ensure even minimal success.

In septic neonates, the use of even one transfusion of granulocytes has been associated with dramatic responses in those with depleted marrow neutrophil reserve (111). This topic is discussed in more detail elsewhere in this text.

Adoptive Immunotherapy

Adoptive immunotherapy can be viewed as the attempt to reconstitute immune competence (usually antitumor) by supplying surrogate lymphoid or other immune system cells.

The role of the immune system in cancer surveillance has been speculative for many years as indirect or circumstantial evidence accumulated. Such data consisted of the increased incidence of various malignancies in patients with disorders that included clinically significant depression of the immune system and similar increases in malignant disease in patients with long-term iatrogenic immunosuppression such as with organ transplants. In addition, the complexity of the various normal cell types and subtypes began to unravel with the discovery of methods to produce monoclonal antibodies to various marker epitopes on lymphoid cell subpopulations. Finally, the identification, categorization, and eventual production (by recombinant DNA methods) of various cytokines permitted further elucidation

of the intricate interplay of cells and cytokines in various aspects of the immune response.

As basic knowledge of immune function grew, the prospects of replacing or augmenting specific steps in the response became more realistic. The development of automated blood cell separators gave clinicians and researchers a tool for collecting large numbers of lymphoid or other immune cells so that they could be manipulated *in vitro*. The transfusion of such cells with or without cytokines is adoptive immunotherapy.

Lymphokine-Activated Killer Cells

In addition to the natural killer (NK) cells that play some role in normal antitumor activity, incubation of mouse or human lymphoid cells with IL-2 generates a population of cells capable of lysing fresh tumor cells *in vitro* (112). These lymphokine-activated killer (LAK) cells seem to differ from both NK cells and cytotoxic T cells. Subsequent studies indicated that LAK cells had a broad range of malignant cell targets, but they spared normal cells and did not seem to be major histocompatibility complex–restricted in their targets. LAK cells are generated after short-term (24 hours) incubation with IL-2. Prior gamma radiation of the lymphoid cells inhibits LAK cell development after IL-2 incubation. This suggests that IL-2–induced proliferation is a prerequisite of LAK cell activity (113).

The origin of LAK cells is still debated, but they seem to be derived from NK cells or from NK-like cells, which are large, granular lymphocytes morphologically and produce NK cell–associated surface markers NKH1 and CD16. However, it should not be assumed that LAK activity is confined to one subtype of lymphocytes. Perhaps it should be thought of as a functional activity rather than a specific cell type.

In animal studies, LAK cells cause regression of established cancers and inhibit growth of pulmonary and hepatic metastatic lesions. These experimental studies yielded a wealth of information on the LAK phenomenon, as summarized succinctly by Klein and Leitman (114): (a) LAK cells alone produce little obvious benefit, (b) IL-2 injected alone has only a modest effect at high doses, (c) a mixture of LAK cells and IL-2 gives striking benefit, (d) for maximal effect, the cells must be cultured with IL-2 for 3 to 4 days, and (e) in humans, a dose of at least 10^{11} LAK cells probably would be needed. It is hypothesized that these cells migrate to sites of malignancy and, when there, expand *in situ* in response to endogenous IL-2. Presumably, they then destroy the adjacent malignant cells by some mechanism.

Early phase I/II studies involved patients with advanced malignancy unresponsive to conventional therapy (114). The regimen was an initial 5-day infusion of IL-2 (100,000 U/kg per 8 hours), followed by five daily leukapheresis procedures. The collected cells were incubated in IL-2 for 3 to 4 days and transfused back to the patient along with another bolus of IL-2. The results in the first group of patients gave grounds for cautious optimism; approximately 10% had a sustained, complete regression of tumor, and perhaps another 10% had a partial response. Renal cell carcinoma, melanoma, and non-Hodgkin's lymphoma responded best (114).

A limiting factor in these early studies was the terminal status of many of the patients. Another practical limiting factor is the enormous cost because of the high doses of recombinant IL-2 needed. Finally, the severe toxicity associated with the therapy is bound to limit its usefulness. At least two treatment-associated deaths have been reported, and virtually all patients have had chills and fever. Other side effects were related to massive fluid shifts resulting in severe hypotension, interstitial pulmonary edema, and massive fluid retention (114).

Another experimental study has demonstrated that LAK cells can be used to purge bone marrow of malignant cells without decreasing or harming its normal hematopoietic cell lines (115). The processing of LAK cells has been automated in a sterilized, closed system. This shows promise because the yield of cells is excellent, and processing time and personnel requirements are decreased (116).

Tumor-Infiltrating Lymphocytes

It is generally true that tumors showing infiltration by lymphocytes have a more favorable prognosis than those without such infiltration. It is reasonable to assume that the lymphocytes seen infiltrating the tumor are in some way responsible for the improved prognosis, and these lymphocytes may well be enriched by cells that have antitumor activity. Recently, lymphocytes that were isolated from murine tumors, expanded under the influence of IL-2, and transfused into mice showed activity against hepatic and pulmonary metastatic lesions that were unresponsive to LAK/IL-2 therapy (117). In the murine model, tumor-infiltrating lymphocytes (TILs) seem to be 50 to 100 times more potent than LAK cells (114). TILs from the tumor are prepared as a single-cell suspension and then incubated for as long as 6 weeks with IL-2. TILs can be seen aggregating at tumor sites for as long as 9 days after transfusion (unlike LAK cells). The long-term effects of the use of TILs are still to be determined in human subjects.

Photopheresis

This is a technique by which a patient's lymphoid cells are removed by apheresis and photoactivated extracorporeally with UVA light. The combination of methoxsalen and UVA light causes loss of viability of nearly 90% of the lymphoid cells. These cells are reinfused. Orally administered methoxsalen transiently intercalates DNA, but when photoactivated by UVA light, it is converted to a form that reacts with the DNA to create permanent covalent cross-links. The photodestruction of these lymphoid cells over the ensuing few days and their recognition by the immune system is thought to lead to a specific immunologic reaction with the malignant clones.

This experimental therapy has been evaluated in patients with cutaneous T-cell lymphoma. In the initial study, 27 of 37 patients showed a mean decrease in cutaneous involvement of 64% after 22 weeks of therapy (118). The mechanism of the beneficial result is unknown, but it is thought to be related to the fact that the therapy is directed extracorporeally to a clone or clones of malignantly transformed cells. The prolonged response in some patients previously shown to have relentlessly progressive disease suggests that this type of T-cell cutaneous lymphoma is at least partly under the control of a specific immunologic response (118).

Early experience with this technique at the Mayo Clinic revealed a 3% patient reaction rate, consisting largely of nausea or vomiting, slight and transient hypotension, or paresthesia. The procedure was generally well tolerated and consistently produced more than 10^9 white blood cells, with minimal red blood cell or platelet contamination, for reinfusion (119).

Noncellular Components

Fresh Frozen Plasma

When plasma is separated from the red blood cells of a donor within 6 hours after donation and is then held at -18°C or lower, it is FFP. If these requirements are not met, it is called liquid plasma. FFP can be stored at -18°C for as long as 1 year. If possible, it is preferable to store FFP at -30°C or lower temperatures to maintain optimal coagulation factor VIII level. Plasma stored at -30° to -40°C retains approximately 90% of such activity, whereas only approximately 60% is retained at approximately -20°C (120). FFP is largely used for its content of the labile coagulation factors V and VIII, but it also contains the so-called stable factors. If properly prepared and stored, each unit contains approximately 200 to 250 mL of plasma and, because of some processing loss, approximately 180 to 200 U of the various coagulation factors.

FFP can be thawed routinely in a water bath at 30° to 37°C with slight agitation of the bag. Rapid microwave warming can be equally effective (121). When FFP is thawed, its labile coagulation factors (V and VIII) begin to destabilize, although FFP is not usually transfused as a source of factor VIII. The unit should be transfused within 6 hours (FDA regulation) or 24 hours (American Association of Blood Banks regulation). Some studies detected significant alterations in plasma factors associated with microwave thawing. Luff and co-workers (122) reported decreases in factors IX, X, and XI and fibrinogen and similar changes in albumin and total protein. They noted that microwave technology does not provide uniform heating, and they were not able to distinguish between simple denaturation of protein and the presence of a nonspecific coagulative property of the microwave. There can be local superheating of thawed plasma immediately adjacent to still-frozen plasma, giving an unpredictable and uneven thawing cycle.

FFP corrects bleeding abnormalities in some patients with congenital (storage pool disease) or acquired (uremia) platelet abnormalities. The reasons for these responses are probably complex, but one factor may be the platelet microparticles found to contaminate most FFP and especially cryoprecipitate (123).

After the introduction and widespread acceptance of the principle of blood component separation and transfusion, the use of FFP increased dramatically. For example, the national use of this component increased from 180,000 units in 1971 to 1.54 million units in 1980 (124). Several studies indicated that FFP is being used as a volume expander or to reconstitute whole blood and that objective criteria for its use are rarely documented in patient charts. In one such study, only 27% of the transfusions were judged to be appropriate (125). These same studies emphasized the need for the education of clinicians regarding the appropriate use of FFP, and one study (126) demonstrated the dramatic positive response to a daily survey and education program in one hospital.

There are few specific indications for FFP. They are limited to the treatment of deficiencies of coagulation factors for which specific factor concentrates are unavailable or inappropriate. The National Institutes of Health Consensus Conference Summary also stated that FFP was a reasonable therapy for patients with uncontrolled bleeding and multiple coagulation protein deficiencies, even though there are few data to support the efficacy of FFP in these circumstances (127). Its use can also be justified for patients with significant hepatic dysfunction and coagulopathy and as a replacement fluid in therapeutic plasma exchange for thrombotic thrombocytopenic purpura.

For patients with specific deficiencies of factors II, V, VII, IX, X, or XI, the use of FFP may be justified, although the mere presence of such demonstrable defects *per se* is insufficient grounds for transfusion. The levels of the factors that need to be attained for hemostasis vary considerably. For factor X, approximately 10% of normal is needed and can be easily achieved with FFP; patients with severe deficiency of factor IX can rarely be brought to hemostatic levels with FFP alone without incurring serious risk of volume overload (127).

Patients with severe liver disease pose a considerable challenge because of the multiplicity of their coagulation factor deficiencies. The factors differ physiologically, and coupled with the low intravascular recovery of some of these factors after transfusion, judging transfusion needs can be difficult. In a 70-kg person, one unit of FFP will raise the circulating levels of factors VII, IX, and V by 1% to 2%. With ascites or anasarca, the extravascular distribution may be increased, and intravascular recovery correspondingly decreased (126). Bleeding in these patients often can be controlled by as little as 10 to 12 mL of FFP per kilogram of body weight.

Warfarinlike anticoagulants act by causing the release of dysfunctional vitamin K–dependent factors II, VII, IX, and X as well as proteins S and C. This situation is best reversed by the use of vitamin K or its analogs, but this may require 24 hours. For more rapid reversal, FFP can be used (in addition to vitamin K) in doses of approximately 10 mL/kg (128). Because of the different half-lives of these factors, their rates of disappearance and reemergence vary. After vitamin K administration, factor VII increases at the fastest rate, achieving 30% of normal in approximately 24 hours, but factors IX and X lag. Medical or surgical emergencies would dictate the concomitant use of FFP to achieve more rapid reversal.

FFP is indicated in the reversal of antithrombin III deficiency in patients who are undergoing surgery or who need heparin for the treatment of antithrombin III deficiency–related thrombosis.

Numerous studies of FFP use in massive transfusion settings failed to document a significant role for this component in that setting and even failed to demonstrate reversal of documented coagulopathy in these situations. Thus, its empiric or prophylactic use during massive transfusion episodes is not supported.

Single-Donor Plasma

If a unit of FFP has not been used after a year of storage at -18°C or lower, it may be relabeled single-donor plasma and, as such, may be stored for 4 more years at -18°C or lower. Likewise,

a unit of FFP that has been thawed at 30° to 37°C and then not transfused within 24 hours (storage at 1° to 6°C) can be relabeled single-donor plasma. This product contains all the constituents of FFP, except for the labile coagulation factors V and VIII, at levels reasonably comparable with those in FFP. Therefore, it can be used in many of the previously discussed situations. Each unit also carries the disease transmission risks associated with any single-donor product.

Cryoprecipitate-Depleted Plasma

If cryoprecipitate has been removed from a unit of whole blood or FFP, the label must so state because such plasma is deficient in factors VIII and XIII, fibrinogen, and fibronectin. This component can be used for volume repletion, but it carries the risks of disease transmission common to FFP and single-donor plasma.

Plasma Protein Fraction

This product is processed commercially to provide an isosmotic, protein-containing product; the protein is largely albumin (83% to 88%). In addition, it contains no more than 1% gamma globulin, sodium at 130 to 160 mEq/L, and potassium at less than 2 mEq/L. This solution is stabilized by 4 mmol/L sodium acetyl tryptophanate and 4 mmol/L sodium caprylate. The final product is heated at 60°C for 10 hours for sterilization. The shelf life is 5 years at 5°C or 3 years if stored at between 5° and 30°C. Plasma protein fraction can be used as an emergency volume expander during treatment of hypovolemic shock owing to burns, trauma, or even hemorrhage. In young children, it can be used as an initial therapy for shock owing to dehydration or infection. The resulting increase in blood volume lasts as long as 48 hours (129). Like other isosmotic protein resuscitation fluids, it must be used with caution in patients with congestive heart failure, severe anemia, high cardiac output states such as hyperthyroidism, or any other condition associated with increased plasma or blood volumes.

In patients undergoing cardiopulmonary bypass surgery, a profound hypotensive reaction has been documented when plasma protein fraction was administered rapidly (faster than 10 mL/min) (130). This was subsequently shown to be caused by Hageman factor fragments, which have a potent vasodilation effect owing to their generation of bradykinin in recipients (131). These reactions have been accompanied by flushing, urticaria, nausea, headache, and backache in conscious patients. There also have been reports of anaphylactic or anaphylactoid reactions to infusions of various plasma protein solutions in atopic patients. Such severe reactions were estimated by Ring and Messmer (132) to occur rarely (0.003%).

Plasma protein fraction and other protein-containing solutions are sometimes used to treat hypoalbuminemia or hypoproteinemia. However, such therapy is effective only in the short term because the protein must be broken down into its constituent amino acids before it can be incorporated into new protein. This breakdown and reincorporation requires several weeks to complete, and it results in only approximately 45% of the infused protein becoming incorporated into the body protein pool (133).

Albumin

Human serum albumin products are widely used in North America because of their relative purity, their standardization of content, and most of all, their freedom from viral contamination (as a result of heat treatment at 60°C for 10 hours). Albumin is available as a 5% solution or a 25% solution that can be diluted at the bedside. Albumin has a molecular weight of 67,000, contains approximately 600 amino acids, and is conspicuously deficient in tryptophan. The viscosity of albumin solutions is remarkably low, and even 25% albumin is almost isoviscous with blood. The molecule carries a net negative charge of -19 at blood pH, a property that plays a role in the binding of albumin to a large number of compounds such as bilirubin, fatty acids, hormones, and medications. Albumin is synthesized mainly in the liver, and because the liver carries only small reserves, serious hepatic dysfunction quickly leads to hypoalbuminemia and its sequelae. Normally, approximately 15 g is synthesized daily in the steady state, but malnutrition or chronic hepatic disease will decrease this rate significantly.

As with other plasma proteins, albumin penetrates capillary walls more slowly than water and electrolytes, and this property confers the molecule's colloid osmotic or oncotic properties. Albumin contributes approximately 80% of the colloid oncotic pressure of plasma, which is approximately 27 mm Hg. When the total serum protein level drops to approximately 5.2 g/dL (the so-called control valve), the colloid oncotic pressure decreases to approximately 20 mm Hg, and further decreases lead to increases in the interstitial fluid compartment, resulting in ascites, edema, or anasarca.

The most common causes of hypoalbuminemia are burns, chronic hepatic disease, chronic protein-losing enteropathy, and acute nephroses. Extensive burns cause the largest acute losses of albumin from the blood. During the first 90 hours after a third- or fourth-degree burn of 50% of the body surface, albumin losses can be the equivalent of approximately two whole normal pools of circulating albumin. A daily loss of as much as 30 g of albumin can be seen for several weeks after a major burn (134). The albumin infusions have to be carefully monitored in terms of the plasma protein levels to ensure adequate oncotic levels. The same caveat applies, to a somewhat lesser degree, to the use of albumin infusions in the other disease states mentioned previously.

Currently, one of the most frequent uses of albumin infusion is as a replacement fluid in plasma-exchange procedures. These procedures are widely used for the treatment of various neurologic syndromes and as therapy for the pruritus of primary biliary cirrhosis. The use of albumin infusions as a long-term treatment for malnutrition is totally inappropriate, but it may be reasonable as a short-term therapy immediately preoperatively in seriously malnourished persons or in those with severe chronic hepatic dysfunction.

The use of albumin infusions rather than crystalloid solutions in the treatment of hypovolemic shock has been controversial for many years. The generally accepted rationale for albumin rather than crystalloid has been that the albumin enhanced fluid reabsorption for the interstitial spaces and, thus, fluids remained longer in the vascular space. In recent years, however, studies have indicated that, except in the gut and in the renal circulation,

there is not sustained reabsorption of fluid at the venous end of capillaries. There appears to be a transient reabsorption that only lasts for a few minutes (135). It was also pointed out that in many disease states, the tissue damage is associated with increased permeability of the capillary walls. This allows direct drainage of fluids out into the interstitial spaces (135). The leakage of the albumin diminishes its capacity for exerting the oncotic pressure one expects if the capillary walls are intact. Indeed, a recent review of randomized controlled trials of the use of albumin in the resuscitation of critically ill patients suggests that the use of albumin is associated with a death rate 6% higher than that found in patients treated without albumin (136).

Cryoprecipitated Antihemophilic Factor

Pool first described a method of concentrating coagulation factor VIII by freezing freshly obtained donor plasma and then slowly thawing the plasma at 1° to 6°C. This precipitates the cold-insoluble proteins that contain approximately 50% of the original factor VIII. This precipitate can then be separated aseptically from the remainder of the plasma, refrozen within 1 hour at -18°C (or preferably at -30°C), and stored for as long as 1 year. This product revolutionized the therapy of classic hemophilia A.

The original source plasma from which FFP is made must be frozen within 6 hours after collection from the donor, and the cryoprecipitate must not be thawed at any time during storage except for immediate infusion. These restrictions apply because of the extreme sensitivity of factor VIII to heat. Federal regulations require that an average bag of cryoprecipitate contain 80 IU of factor VIII, and American Association of Blood Banks standards require that at least 75% of the bags contain this amount of coagulation factor VIII. A cryoprecipitate bag from a single whole-blood donation may contain 80 to 120 IU of factor VIII, 40% to 70% of VWF and 20% to 30% of factor XIII in the original unit of plasma, 150 to 250 mg of fibrinogen, and approximately 55 mg of fibronectin (α_2-opsonic globulin) in a volume of less than 15 mL. Cryoprecipitate bags are often pooled after final thawing at 37°C, just before infusion, to give an adequate dose of factor VIII.

In patients with classic hemophilia, serious bleeding episodes generally are associated with factor VIII levels below 3% of normal. In these patients, the minimal hemostatic level of factor VIII is approximately 30%; if levels of approximately 50% are attained and maintained for approximately 2 weeks, major surgery can be carried out safely. It is not necessary to achieve levels of 100%, and to attempt to do so would expose patients needlessly to a great many more donors. There are several empiric schemata for the dosage of cryoprecipitate in classic hemophilia, but a reasonable starting dose is one bag for each 6 kg of body weight (137). After the initial therapy, repeat doses will be needed approximately every 12 hours, and the level of such doses will need to be monitored by factor VIII assays.

The use of cryoprecipitate is hampered by the fact that it must be maintained at -18° to -30°C, and its pooling and infusion generally requires direct professional involvement. These factors limit its applicability to home therapy programs, an approach that has freed thousands of hemophiliacs from dependence on medical supervision. Unfortunately, these programs depended heavily on the use of pooled factor concentrates, which became contaminated by HIV and hepatitis viruses several years ago. This led to the disastrous prevalence of HIV positivity in hemophiliacs. Because of this, some hematologists are once again relying on cryoprecipitate for the treatment of young hemophiliacs who have not yet been exposed to HIV in factor concentrates. Monoclonally produced factor concentrates and those that are heat treated are now essentially free of infectious HIV or hepatitis viruses. The availability of factor VIII produced from cloned DNA eventually will eliminate the risk of virus transmission altogether, but it will greatly increase the cost of therapy for hemophilic patients.

There have been numerous modifications made to the standard method of making cryoprecipitate, and readers are referred to the reports in the literature (138,139 and 140).

Fibrin Glue (Cryoprecipitated Fibrinogen)

Because of its high content of fibrinogen, cryoprecipitate has been employed as a source of fibrin for topical use in some types of trauma or surgical situations. For this purpose, one can use a standard bag of cryoprecipitate to which thrombin is added in the operating room to form a coagulum. This has been utilized particularly for oozing on raw surfaces such as a lacerated liver or in areas where adhesions had to be dissected and no discrete, identifiable small vessels could be easily ligated. Cardiac surgeons have found this latter application to be of particular use when repeat thoracotomy must be performed (141). If sufficient time is given, the blood bank can make autologous fibrin glue and eliminate viral transmission risks (142).

Artificial Blood Substitutes

When this term is used, it generally refers to artificial substitutes for red blood cells. Because of the problems encountered with the use of red blood cell transfusions, such as disease transmission, red blood cell alloimmunization, and blood shortages, numerous attempts have been made to develop a substitute for red blood cells to deliver oxygen to tissues. The most promising of these lines of research have involved the use of (a) solubilized hemoglobin, (b) perfluorocarbon solutions, and (c) encapsulated hemoglobins. Each of these approaches has had problems that have prevented its introduction on a broad scale. Hemoglobin that has been rendered free of stroma will carry oxygen; however, its normal tetramer composition is quickly modified to a dimer. These dimers are rapidly excreted in the urine. The hemoglobin in a red blood cell is associated with 2,3-DPG, which interacts with it to allow oxygen dissociation to the tissues. Free hemoglobin has a very high affinity for oxygen and does not release it readily. The oncotic pressure of free hemoglobin is double the normal plasma oncotic pressure. Therefore, the maximal allowable hemoglobin concentration in the circulation is approximately 7 g/dL. Even stroma-free hemoglobin has some nephrotoxicity. The major activities in the development of hemoglobin have centered on attempts to cross-link the hemoglobin and on efforts toward microencapsulation (143).

Fluosol-DA is an emulsion of two perfluorocarbons, perfluorodecalin

and perfluorotripropylamine, in a balanced salt solution with hydroxyethyl starch (143). This product, which must be frozen before use, has an intravascular half-life of approximately 7 hours and transports gases in proportion to the partial pressure of the gas with which it is in equilibrium (Henry's law). If it is used to substitute for red blood cells, the patient must be inhaling a very high concentration of oxygen. The solubility of oxygen at 37°C and a partial pressure of 760 mm Hg is approximately 45 mL/100 mL of perfluorocarbon. Red blood cells have a similar oxygen content at approximately 100 mm Hg. As a result, they work while the patients breathe room air (144). Several toxic reactions resulting from the use of perfluorocarbons have occurred, and this has led to withdrawal of their approval by the FDA. These reactions were variable, but several involved pulmonary problems apparently secondary to complement activation (145).

Encapsulating hemoglobin in liposomes is an alternative method of preparing red blood cell substitutes (146). These liposomes with hemoglobin have a half-life of only 5 hours and tend to be retained in the liver. However, they do seem to function adequately to deliver oxygen, and they have an oxygen dissociation curve similar to that of whole blood. Much work remains to be done before we can transfuse patients with an oxygen-carrying substitute that is nearly as efficient, as desirable, and as effective as the red blood cell.

REFERENCES

1. Pool JG, Hershgold EJ, Pappenhagen AR. High-potency antihaemophilic factor concentrate prepared from cryoglobulin precipitate. *Nature* 1964;203:312.

2. Pool JG, Shannon AE. Production of high-potency concentrates of antihemophilic globulin in a closed-bag system. *N Engl J Med* 1965;273:1443–1447.

3. Widmann FK, ed. Blood and blood components: preparation, storage and shipment. In: *Technical manual of the American Association of Blood Banks*, 9th ed. Arlington, VA: American Association of Blood Banks, 1985:35–58.

4. Widmann FK, ed. Cryopreservation of blood and tissue. In: *Technical manual of the American Association of Blood Banks*, 9th ed. Arlington, VA: American Association of Blood Banks, 1985:59–70.

5. Meryman HT. Cryoprotective agents. *Cryobiology* 1971;8:173–183.

6. Meryman HT. Cryopreservation of blood cells and tissues. In: Petz LD, Swisher SN, eds. *Clinical practice of transfusion medicine*. New York: Churchill-Livingstone, 1981:297–315.

7. Williamson LM, Allain JP. Virally Inactivated fresh frozen plasma. *Vox Sang* 1992;69:159–165.

8. Charm SE, Landau S, Williams B, et al. High-temperature short-time heat inactivation of HIV and other viruses in human blood plasma. *Vox Sag* 1992;62:12–20

9. Horowitz B, Wiebe ME, Lippin A, et al. Inactivation of viruses in labile blood derivatives: 1. Disruption of lipid-enveloped viruses by tri (N-butyl) phosphate detergent combinations. *Transfusion* 1985;25:516–522.

10. Horowitz B, Wiebe ME, Lippin A, et al. Inactivation of viruses in labile blood derivatives: II physical methods. *Transfusion* 1985;25:523–527.

11. Gerhard G. Solvent/detergent-treated plasma: indications, risks, and benefits. In: Fridey JL, Simpson MB, eds. *Component selection: polemics and politics*. Bethesda, MD: American Association of Blood Banks, 1996:45.

12. Klein HG, Dodd RY, Dzik WH, et al. Current status of solvent/detergent-treated frozen plasma. *Transfusion* 1998;38:102–104.

13. AuBuchon JP, Birkmeyer JD. Safety and cost effectiveness of solvent-detergent treated plasma: In search of a zero risk blood supply. *JAMA* 1994;272:1210–1214.

14. Murray R, Oliphant JW, Tripp JT, et al. Effect of ultraviolet radiation on the infectivity of icterogenic plasma. *JAMA* 1955;157:8–14.

15. Prince AM, Stephan W, Brotman B. B-propiolactone/ultraviolet irradiation: a review of its effectiveness for inactivation of viruses in blood derivatives. *Rev Infect Dis* 1983;5:92–107.

16. Lin L, Cook DN, Weisehahn GP, et al. Photochemical inactivation of viruses and bacteria in platelet concentrates by use of a novel psoralen and long wavelength ultraviolet light. *Transfusion* 1997;37:423–435.

17. Margolis-Nunno H, Robinson R, Ben-Hur E, et al. Elimination of potential mutagenicity in platelet concentrates that are virally inactivated with psoralens and ultraviolet A light. *Transfusion* 1995;35:855–862.

18. AuBuchon JP, Pickard C, Herschel L, et al. Removal of methylene blue from plasma via an adsorbent filter. *Vox Sang* 1998;74:1–6.

19. Zeiler T, Riess H, Wittman G, et al. The effect of methylene blue phototreatment on plasma proteins and in vitro coagulation capability of single donor fresh-frozen plasma. *Transfusion* 1994;34:685–689.

20. Horowitz B, Piet MPJ, Prince AM, et al. Inactivation of lipid-enveloped viruses in labile blood derivatives by unsaturated fatty acids. *Vox Sang* 1988;54:14–20.

21. Hiemstra H, Tersmette M, Vos AHV, et al. Inactivation of human immunodeficiency virus by gamma radiation and its effect on plasma and coagulation factors. *Transfusion* 1991;31:32–39.

22. Highsmith FA, Xue H, Caple M, et al. Inactivation of lipid-enveloped and non-lipid enveloped model viruses in normal human plasma by crosslinked starch-iodine. *Transfusion* 1994;34:322–327.

23. Wells KH, Latino J, Gavalchin J, et al. Inactivation of human immunodeficiency virus type I by ozone *in vitro*. *Blood* 1991;78;1882–1890.

24. Lavee J, Martinowitz U, Mohr R, et al. The effect of transfusion of fresh whole blood versus platelet concentrates after cardiac operations. A scanning electron microscope study of platelet aggregation on extracellular matrix. *J Thorac Cardiovasc Surg* 1989;97:204–212.

25. Pliam MB, McGoon DC, Tarhan S. Failure of transfusion of autologous whole blood to reduce banked-blood requirements in open-heart surgical patients. *J Thorac Cardiovasc Surg* 1975;70:338–343.

26. Head DR. Whole blood versus packed red cells [Letter]. *Transfusion* 1985;25:591.

27. Pineda AA, Viggiano NR, Clare D, et al. Flow rate of AS-1 preserved red blood cells and whole blood—a comparison. *Transfusion* 1985;25:449(abst).

28. National Institutes of Health Consensus Development Conference. Perioperative red cell transfusion. *Transfus Med Rev* 1989;3:63–68.

29. Hunt TK, Rabkin J, von Smitten K. Effects of edema and anemia on wound healing and infection. *Curr Stud Hematol Blood Transfus* 1986;53:101–111.

30. Blumberg N, Heal JM. Transfusion and host defenses against cancer recurrence and infection. *Transfusion* 1989;29:236–245.

31. Meryman HT, Bross J, Lebovitz R. The preparation of leukocyte-poor red blood cells: a comparative study. *Transfusion* 1980;20:285–292.

32. Perkins HA, Payne R, Ferguson J, et al. Nonhemolytic febrile transfusion reactions. Quantitative effects of blood components with emphasis on isoantigenic incompatibility of leukocytes. *Vox Sang* 1966;11:578–600.

33. Menitove JE, McElligott MC, Aster RH. Febrile transfusion reaction: what blood component should be given next? *Vox Sang* 1982;42:318–321.

34. Meryman HT, Hornblower M. The preparation of red cells depleted of leukocytes: review and evaluation. *Transfusion* 1986;26:101–106.

35. Sirchia G, Rebulla P, Parravicini A, et al. Leukocyte depletion of red cell units at the bedside by transfusion through a new filter. *Transfusion* 1987;27:402–405.

36. Snyder EL, Bookbinder M. Role of microaggregate blood filtration in clinical medicine. *Transfusion* 1983;23:460–470.

37. Davey RJ, Carmen RA, Simon TL, et al. Preparation of white cell-depleted red cells for 42-day storage using an integral in-line filter. *Transfusion* 1989;29:496–499.

38. Brittingham TE, Chaplin H Jr. Febrile transfusion reactions caused by sensitivity to donor leukocytes and platelets. *JAMA* 1957; 165:819–825.

39. Terasaki PI, Kreisler M, Mickey RM. Presensitization and kidney transplant failures. *Postgrad Med J* 1971;47:89–100.

40. Opelz G, Sengar DPS, Mickey MR, et al. Effect of blood transfusions on subsequent kidney transplants. *Transplant Proc* 1973;5:253–259.

41. Opelz G, Terasaki PI. Prolongation effect of blood transfusions on kidney graft survival. *Transplantation* 1976;22:380–383.

42. Unanue ER, Allen PM. The basis for the immunoregulatory role of macrophages and other accessory cells. *Science* 1987;236:551–557.

43. Meryman HT. Transfusion-induced alloimmunization and immunosuppression and the effects of leukocyte depletion. *Transfus Med Rev* 1989;3:180–193.

44. Sniecinski I, O'Donnell MR, Nowicki B, et al. Prevention of refractoriness and HLA-alloimmunization using filtered blood products. *Blood* 1988;71:1402–1407.

45. Bertolini F, Rebulla P, Riccardi D, et al. Evaluation of platelet concentrates prepared from buffy coats and stored in a glucose-free crystalloid medium. *Transfusion* 1989;29:605–609.

46. Holland PV, Schmidt PJ, eds. *Standards for blood banks and transfusion services*, 12th ed. Arlington, VA: American Association of Blood Banks, 1987.

47. Slichter SJ, Harker LA. Thrombocytopenia: mechanisms and management of defects in platelet production. *Clin Haematol* 1978;7:523–539.

48. Harker LA, Malpass TW, Branson HE, et al. Mechanism of abnormal bleeding in patients undergoing cardiopulmonary bypass: acquired transient platelet dysfunction associated with selective α-granule release. *Blood* 1980;56:824–834.

49. Consensus Conference. Platelet transfusion therapy. *JAMA* 1987;257:1777–1780.

50. Slichter SJ. Indications for platelet transfusions. *Plasma Ther Transfus Technol* 1982;3:259–264.

51. National Institutes of Health Consensus Conference. Platelet transfusion therapy. *Transfus Med Rev* 1987;1:195–200.

52. Simon TL, Akl BF, Murphy W. Controlled trial of routine administration of platelet concentrates in cardiopulmonary bypass surgery. *Ann Thorac Surg* 1984;37:359–364.

53. Kelton JG, Neame PB, Gauldie J, et al. Elevated platelet-associated IgG in the thrombocytopenia of septicemia. *N Engl J Med* 1979;300:760–764.

54. Tate DY, Carlton GT, Johnson D, et al. Immune thrombocytopenia in severe neonatal infections. *J Pediatr* 1981;98:449–453.

55. van Der Lelie J, van der Plas-Van Dalen CM, von Dem Borne AEGKR. Platelet autoantibodies in septicaemia. *Br J Haematol* 1984;58:755–760.

56. Stewart JH, Castaldi PA. Uraemic bleeding: a reversible platelet defect corrected by dialysis. *Q J Med* 1967;36:409–423.

57. Carvalho ACA. Bleeding in uremia-a clinical challenge. *N Engl J Med* 1983;308:38–39.

58. Mannucci PM, Remuzzi G, Pusineri F, et al. Deamino-8-D-arginine vasopressin shortens the bleeding time in uremia. *N Engl J Med* 1983;308:8–12.

59. Sniecinski I, St. Jean J, Nowicki B. Preparation of leukocyte-poor platelets by filtration. *J Clin Apheresis* 1989;5:7–11.

60. Holme S, Ross D, Heaton WA. *In vitro* and *in vivo* evaluation of platelet concentrates after cotton wool filtration. *Vox Sang* 1989;57:112–115.

61. Stec N, Kickler TS, Ness PM, et al. Effectiveness of leukocyte (WBC) depleted platelets in preventing febrile reactions in multi-transfused oncology patients. *Transfusion* 1986;26:569(abst).

62. Kickler TS, Bell W, Ness PM, et al. Depletion of white cells from platelet concentrates with a new adsorption filter. *Transfusion* 1989;29:411–414.

63. Schiffer CA, Patten E, Reilly J, et al. Effective leukocyte removal from platelet preparations by centrifugation in a new pooling bag. *Transfusion* 1987;27:162–164.

64. Ferrara GB, Tosi RM, Azzalina G, et al. The production of anti HLA-A cytotoxic antisera through planned immunization by intravenous injection of small aliquots of whole blood. I. Immunization of unrelated recipients. *Tissue Antigens* 1972;2:359–373.

65. Vesilind GW, Simpson MB, Shifman MA, et al. Evaluation of a centrifugal blood cell processor for washing platelet concentrates. *Transfusion* 1988;28:46–51.

66. Schiffer CA, Aisner J, Dutcher JP, et al. A clinical program of platelet cryopreservation. *Prog Clin Biol Res* 1982;88:165–180.

67. McGill M, Mayhaus C, Hoff R, et al. Frozen maternal platelets for neonatal thrombocytopenia. *Transfusion* 1987;27:347–349.

68. Moore SB. Management of transfusion in the massively bleeding patient. *Hum Pathol* 1983;14:267–270.

69. Mannucci PM, Federici AB, Sirchia G. Hemostasis testing during massive blood replacement: a study of 172 cases. *Vox Sang* 1982;n.s.42:113–123.

70. Reed RL II, Ciavarella D, Heimbach DM, et al. Prophylactic platelet administration during massive transfusion. *Ann Surg* 1986; 203:40–48.

71. Howard JE, Perkins HA. The natural history of alloimmunization to platelets. *Transfusion* 1978;18:496–503.

72. Holohan TV, Terasaki PI, Deisseroth AB. Suppression of transfusion-related alloimmunization in intensively treated cancer patients. *Blood* 1981;58:122–128.

73. Yankee RA, Grumet FC, Rogentine GN. Platelet transfusion therapy. The selection of compatible platelet donors for refractory patients by lymphocyte HL-A typing. *N Engl J Med* 1969;281:1208–1212.

74. Lee EJ, Schiffer CA. Serial measurement of lymphocytotoxic antibody and response to nonmatched platelet transfusions in alloimmunized patients. *Blood* 1987;70:1727–1729.

75. McGrath K, Wolf M, Bishop J, et al. Transient platelet and HLA antibody formation in multitransfused patients with malignancy. *Br J Haematol* 1988;68:345–350.

76. Murphy MF, Metcalfe P, Ord J, et al. Disappearance of HLA and platelet-specific antibodies in acute leukaemia patients alloimmunized by multiple transfusions. *Br J Haematol* 1987;67:255–260.

77. Bensinger WI, Buckner CD, Clift RA, et al. Plasma exchange for platelet alloimmunization. *Transplantation* 1986;41:602–605.

78. Dunstan RA, Simpson MB. Stability of platelet surface antigens during storage. *Transfusion* 1985;25:563–566.

79. Santoso S, Mueller-Eckhardt G, Santoso S, et al. HLA antigens on platelet membranes: *in vitro* and *in vivo* studies. *Vox Sang* 1986;51:327–333.

80. Daly PA, Schiffer CA, Aisner J, et al. Platelet transfusion therapy. One-hour posttransfusion increments are valuable in predicting the need for HLA-matched preparations. *JAMA* 1980;243:435–438.

81. Bishop JF, McGrath K, Wolf MM, et al. Clinical factors influencing the efficacy of pooled platelet transfusions. *Blood* 1988;71:383–387.

82. Kotlán B, Gyódi É, Szabó T, et al. Comparative study of alloimmune reactions induced by leukocyte and platelet transfusions in humans: characteristic changes of activation markers, gamma interferon and FcR blocking antibody production. *Hum Immunol* 1988;22:19–29.

83. Takahashi K, Juji T, Miyazaki H. Determination of an appropriate size of unrelated donor pool to be registered for HLA-matched platelet transfusion. *Transfusion* 1987;27:394–398.

84. Schiffer CA, O'Connell B, Lee EJ. Platelet transfusion therapy for alloimmunized patients: selective mismatching for HLA B12, an antigen with variable expression on platelets. *Blood* 1989;74:1172–1176.

85. McElligott MC, Menitove JE, Duquesnoy RJ, et al. Effect of HLA Bw4/Bw6 compatibility on platelet transfusion responses of refractory thrombocytopenic patients. *Blood* 1982;59:971–975.

86. Moore SB. HLA and blood component therapy. In: Hackel E, Mallory D, eds. *Theoretical aspects of HLA*. Arlington, VA: American Association of Blood Banks, 1982:81–103.

87. Dahlke MB, Weiss KL. Platelet transfusion from donors mismatched for crossreactive HLA antigens. *Transfusion* 1984;24:299–302.

88. Herzig RH, Poplack DG, Yankee RA. Prolonged granulocytopenia from incompatible platelet transfusions. *N Engl J Med* 1974;290:1220–1223.

89. Kieckbusch ME, Moore SB, Koenig VA, et al. Platelet crossmatch evaluation in refractory hematologic patients. *Mayo Clin Proc* 1987;62:595–600.

90. Freedman J, Garvey MB, de Friedberg ZS, et al. Random donor platelet crossmatching: comparison of four platelet antibody detection methods. *Am J Hematol* 1988;28:1–7.

91. Heal JM, Blumberg N, Masel D. An evaluation of crossmatching, HLA and ABO matching for platelet transfusions to refractory patients. *Blood* 1987;70:23–30.

92. Petz LD. Platelet crossmatching [Letter]. *Am J Clin Pathol* 1988;90:114–115.

93. Welch HG, Larson EB, Slichter SJ. Providing platelets for refractory patients: prudent strategies. *Transfusion* 1989;29:193–195.

94. Rachel JM, Summers TC, Sinor LT, et al. Use of a solid phase red blood cell adherence method for pretransfusion platelet compatibility testing. *Am J Clin Pathol* 1988;90:63–68.

95. Lee EJ, Schiffer CA. ABO compatibility can influence the results of platelet transfusion: results of a randomized trial. *Transfusion* 1989;29:384–389.

96. Schiffer CA, Slichter SJ. Sounding board: platelet transfusions from single donors. *N Engl J Med* 1982;307:245–248.

97. Brand A, Claas FHJ, Voogt PJ, et al. Alloimmunization after leukocyte-depleted multiple random donor platelet transfusions. *Vox Sang* 1988;54:160–166.

98. Slichter SJ, Deeg HJ, Kennedy MS. Prevention of platelet alloimmunization in dogs with systemic cyclosporine and by UV-irradiation or cyclosporine-loading of donor platelets. *Blood* 1987;69:414–418.

99. Bodey GP, Buckley M, Sathe YS, et al. Quantitative relationships between circulating leukocytes and infection in patients with acute leukemia. *Ann Intern Med* 1966;64:328–340.

100. Strumia MM. The effect of leukocytic cream injections in the treatment of the neutropenias. *Am J Med Sci* 1934;187:527–544.

101. McCullough J. Granulocyte transfusion. In: Petz LD, Swisher SN, eds. *Clinical practice of transfusion medicine*, 2nd ed. New York: Churchill-Livingstone, 1989:469–484.

102. Dutcher JP, Schiffer CA, Johnston GS. Rapid migration of ^{111}indium-labelled granulocytes to sites of infection. *N Engl J Med* 1981;304:586–589.

103. McCullough J, Weiblen BJ, Clay ME, et al. Effect of leukocyte antibodies on the fate *in vivo* of indium-111-labelled granulocytes. *Blood* 1981;58:164–170.

104. McCullough J, Clay M, Loken M, et al. Effect of ABO incompatibility on the fate *in vivo* of ^{111}indium granulocytes. *Transfusion* 1988;28:358–361.

105. Dutcher JP, Schiffer CA, Johnston GS, et al. Alloimmunization prevents the migration of transfused indium-111-labelled granulocytes to sites of infection. *Blood* 1983;62:354–360.

106. Dahlke MB, Keashen M, Alavi JB, et al. Granulocyte transfusions and outcome of alloimmunized patients with gram-negative sepsis. *Transfusion* 1982;22:374–378.

107. Strauss RG, Connett JE, Gale RP, et al. A controlled trial of prophylactic granulocyte transfusions during initial induction chemotherapy for acute myelogenous leukemia. *N Engl J Med* 1981;305:597–603.

108. Wright DG, Robichaud KJ, Pizzo PA, et al. Lethal pulmonary reactions associated with the combined use of amphotericin B and leukocyte transfusions. *N Engl J Med* 1981;304:1185–1189.

109. Forman SJ, Robinson GV, Wolf JL, et al. Pulmonary reactions associated with amphotericin B and leukocyte transfusions [Letter]. *N Engl J Med* 1981;305:584–585.

110. Clift RA, Sanders JE, Thomas ED, et al. Granulocyte transfusions for the prevention of infection in patients receiving bone-marrow transplants. *N Engl J Med* 1978;298:1052–1057.

111. Christensen RD, Anstall H, Rothstein G. Neutrophil transfusion in septic neutropenic neonates. *Transfusion* 1982;22:151–153.

112. Grimm EA, Mazumder A, Zhang HZ, et al. Lymphokine-activated killer cell phenomenon. Lysis of natural killer-resistant fresh solid tumor cells by interleukin 2-activated autologous human peripheral blood lymphocytes. *J Exp Med* 1982;155:1823–1841.

113. Grimm EA, Ramsey KM, Mazumder A, et al. Lymphokine-activated killer cell phenomenon. II. Precursor phenotype is serologically distinct from peripheral T lymphocytes, memory cytotoxic thymus-derived lymphocytes and natural killer cells. *J Exp Med* 1983;157:884–897.

114. Klein HG, Leitman SF. Adoptive immunotherapy in the treatment of malignant disease. *Transfusion* 1989;29:170–178.

115. van den Brink MRM, Voogt PJ, Marijt WAF, et al. Lymphokine-activated killer cells selectively kill tumor cells in bone marrow without compromising bone marrow stem cell function *in vitro*. *Blood* 1989;74:354–360.

116. Ryan T, Atkins MB, Mier JW, et al. Lymphokine-activated killer cell generation and recovery. Comparison of an automated cell processing device and a manual procedure. *Transfusion* 1989;29:491–495.

117. Rosenberg SA, Spiess P, Lafreniere R. A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. *Science* 1986;233:1318–1321.

118. Edelson R, Berger C, Gasparro F, et al. Treatment of cutaneous T-cell lymphoma by extracorporeal photochemotherapy. Preliminary results. *N Engl J Med* 1987;316:297–303.

119. Burgstaler EA, Pineda AA. Photopheresis: blood cell recovery and reaction incidence of 247 procedures. *Transfusion* 1989;29(Suppl):44S.

120. Rock G. Factor VIII concentrates. In: Cash JD, ed. *Progress in transfusion medicine. Volume 2*. New York: Churchill-Livingstone, 1987:127.

121. Rock G, Tackaberry ES, Dunn JG, et al. Rapid controlled thawing of fresh-frozen plasma in a modified microwave oven. *Transfusion* 1984;24:60–65.

122. Luff RD, Kessler CM, Bell WR. Microwave technology for the rapid thawing of frozen blood components. *Am J Clin Pathol* 1985;83:59–64.

123. George JN, Pickett EB, Heinz R. Platelet membrane microparticles in blood bank fresh frozen plasma and cryoprecipitate. *Blood* 1986;68:307–309.

124. Snyder AJ, Gottschall JL, Menitove JE. Why is fresh-frozen plasma transfused? *Transfusion* 1986;26:107–112.

125. Blumberg N, Laczin J, McMican A, et al. A critical survey of fresh-frozen plasma use. *Transfusion* 1986;26:511–513.

126. Shanberge JN. Reduction of fresh-frozen plasma use through a daily survey and education program. *Transfusion* 1987;27:226–227.

127. National Institutes of Health Consensus Conference. Fresh frozen plasma: indications and risks. *Transfus Med Rev* 1987;1:201–204.

128. Counts RB. Acquired bleeding disorders. In: Menitove JE, McCarthy LJ, eds. *Hemostatic disorders and the blood bank*. Arlington VA: American Association of Blood Banks, 1984:41–47.

129. Bertrand JJ, Feichtmeir TV, Kolomeyer NK, et al. Clinical investigations with a heat-treated plasma protein fraction-Plasmanate®. *Vox Sang* 1959;4:385–402.

130. Bland JHL, Laver MB, Lowenstein E. Vasodilator effect of commercial 5% plasma protein fraction solutions. *JAMA* 1973; 224:1721–1724.

131. Alving BM, Hojima Y, Pisano JJ, et al. Hypotension associated with prekallikrein activator (Hageman-factor fragments) in plasma protein fraction. *N Engl J Med* 1978;299:66–70.

132. Ring J, Messmer K. Incidence and severity of anaphylactoid reactions to colloid volume substitutes. *Lancet* 1977;1:466–469.

133. Waterhouse C, Bassett SH, Holler JW. Metabolic studies on protein-depleted patients receiving a large part of their nitrogen intake from human serum albumin administered intravenously. *J Clin Invest* 1949;28:245–264.

134. O'Riordan JP, Aebischer M, Darnborough J, et al. The indications for the use of albumin, plasma protein solutions and plasma substitutes. In: *Council of Europe, European Public Health Committee Report*. Strasbourg, 1978:39–58.

135. Berger A. Why albumin may not work. *BMJ* 1998;317:240.

136. Cochrane Injuries Group Albumin Reviewers. Human albumin administration in critically ill patients: systematic review of randomized controlled trials. *BMJ* 1998;317:235–240.

137. Lusher JM. Replacement therapy for congenital and acquired disorders of blood coagulation-available products. In: Garratty G, ed. *Current concepts in transfusion therapy*. Arlington, VA: American Association of Blood Banks, 1985:27–49.

138. Rock G, Smiley RK, Tittley P, et al. *In vivo* effectiveness of a high-yield factor VIII concentrate prepared in a blood bank. *N Engl J Med* 1984;311:310–313.

139. Mason EC. Thaw-siphon technique for production of cryoprecipitate concentrate of factor VIII. *Lancet* 1978;2:15–17.

140. Bass H, Trenchard PM, Mustow MJ. Microwave-thawed plasma for cryoprecipitate production. *Vox Sang* 1985;48:65–71.

141. Dresdale A, Rose EA, Jeevanandam V, et al. Preparation of fibrin glue from single-donor fresh-frozen plasma. *Surgery* 1985; 97:750–754.

142. Silberstein LE, Williams LJ, Hughlett MA, et al. An autologous fibrinogen-based adhesive for use in otologic surgery. *Transfusion* 1988; 28:319–321.

143. Chang TMS. The use of modified hemoglobin as an oxygen carrying blood substitute. *Transfus Med Rev* 1989;3:213–218.

144. Bowman RJ. Red blood cell substitutes and artificial blood. *Hum Pathol* 1983;14:218–220.

145. Vercellotti GM, Hammerschmidt DE, Craddock PR, et al. Activation of plasma complement by perfluorocarbon artificial blood: probable mechanism of adverse pulmonary reactions in treated patients and rationale for corticosteroid prophylaxis. *Blood* 1982;59:1299–1304.

146. Kahn RA, Allen RW, Baldassare J. Alternate sources and substitutes for therapeutic blood components. *Blood* 1985;66:1–12.

72
Transfusion Therapy

Gary Stack

Steven Mechanic

Transfusion therapy is the most commonly practiced form of tissue transplantation. It is an essential component of the overall management of a wide range of medical disorders. Transfusion is most often an adjunct to the major treatment modality, as in the perioperative setting or in patients undergoing bone marrow ablation as a consequence of cancer treatment. Without transfusion therapy, many surgical procedures and cancer treatment regimens would not be possible or would have intolerable risk. In other situations, transfusion therapy is itself the definitive treatment modality, as in thrombotic thrombocytopenic purpura (TTP). This chapter summarizes the role of transfusion therapy in various clinical settings. In some clinical settings, pharmacologic alternatives to transfusion are mentioned.

Transfusion therapy generally follows a semiquantitative dosing schedule. Dosages often have not firmly been established, and the composition of blood components often is not well defined or tightly controlled, at least not to an extent that meets today's pharmacologic practices. Nonetheless, where possible and applicable, this chapter presents dosing recommendations for transfusion. However, in most cases, transfusion therapy continues to demand attentive clinical judgment based on the unique clinical situation of each patient.

This chapter also covers the major issues involved in the administration of blood components to transfusion recipients both in hospital and out-of-hospital settings. Rigorous systems must be in place to assure the proper identification and matching of the transfusion recipient and the blood component selected for that patient. This is vital because transfusion therapy presents hazards of immunologic incompatibility and because it is a common and high-volume activity that presents ample opportunity for error. Proper attention also must be paid to the infusion procedure itself to minimize the adverse effects of transfusion. Component modifications, such as gamma irradiation, leukoreduction filtration, and cytomegalovirus (CMV) risk reduction, should be carried out to reduce some risks of transfusion for some recipients. However, attention should be given to avoiding their inappropriate and costly overuse. When carried out properly and for the appropriate indications, transfusion therapy is a highly effective and generally well-tolerated treatment of anemia, acute blood loss, and a wide variety of other blood-related disorders.

GENERAL INDICATIONS FOR TRANSFUSION OF BLOOD COMPONENTS

Part of "72 - Transfusion Therapy"

Red Blood Cells

The transfusion of red blood cells (RBCs) is indicated in situations in which the loss of the oxygen-carrying capacity of the blood seriously threatens the well-being of the patient and there is inadequate time or possibility for increasing the RBC mass through erythropoiesis. Clinically apparent signs and symptoms of anemia useful in evaluating inadequate oxygen delivery include tachycardia, dyspnea, tachypnea, pallor, weakness, fatigue, light-headedness, syncope, and angina. The most serious health risks of reduced oxygen delivery include organ failure, such as myocardial infarction and stroke.

Guidelines for RBC transfusion are largely consensus driven, although to some extent evidence based. In general, RBC transfusion is considered appropriate for any patient with a hemoglobin level less than 7 g/dL (1). The transfusion trigger may even be set lower, i.e., approximately 6 g/dL, in patients who can better tolerate lower oxygen delivery, such as younger or healthier individuals and those who have adapted to a chronic anemia (2,3). Support for this lower threshold comes from data indicating that an acute decrease in hemoglobin concentration to 5 g/dL in otherwise healthy individuals at rest does not decrease oxygen consumption or increase blood lactate levels (4). For hemoglobin levels between 6 to 7 and 10 g/dL, RBC transfusion is indicated for patients who are symptomatic from their anemia and/or whose medical conditions predispose them to poor tissue perfusion, such as patients with cardiac or other circulatory disorders (1,2 and 3). For patients who are anemic due to vitamin B_{12} or folate deficiency, the first priority is pharmacologic replacement therapy; RBC transfusion is generally not indicated because of the patient's adaptation to the slowly developing anemia. RBC transfusion appears unnecessary for most patients with hemoglobin levels greater than 10 g/dL (1,2,3,4 and 5).

Some transfusion specialists caution that rigid adherence to guidelines can lead to the undertransfusion of some patients. Some evidence exists that higher hematocrits in the range of 30% to 35% give improved oxygen delivery and enhanced hemostasis that may be advantageous for patients who are at high risk for myocardial infarction or perioperative, nonsurgical bleeding (6,7). Transfusion thresholds higher than a 30% hematocrit for adults remain controversial. Neonates in respiratory

distress also may require hematocrits greater than 30% to 35% (3). Outcome studies are needed to resolve the debates over transfusion triggers. Because many variables modify the necessity and urgency of RBC transfusion, including the degree, duration, and etiology of the anemia, as well as the patient's medical condition, the decision to transfuse remains to a large extent a matter of clinical judgment.

Because one unit of RBCs will raise the hemoglobin by approximately 1.0 g/dL and the hematocrit by approximately 3% in a 70-kg adult, the dose of RBC units can be approximated by dividing the desired increase in % hematocrit by 3. For adults and children, the dose may also be calculated as follows:

$$\text{Volume of RBCs (mL)} = \text{EBV (mL)} \times \frac{\text{Desired Hct} - \text{Actual Hct}}{\text{Hct of RBCs}}$$

The blood volume and cardiac status of the recipient should be taken into account in determining the RBC dose and the transfusion rate. Patients with congestive heart failure or those who are otherwise at risk of a hypervolemic reaction may not tolerate the added volume of RBCs transfused. For these patients, transfusion must proceed slowly, and diuretic use should be considered. In some cases, it may be desirable for the transfusion service to split a unit in half so that each half may be transfused over the maximum allowable time period of 4 hours.

Platelets

Patients with platelet counts less than 5,000 to 10,000/μL are at significant risk of spontaneous hemorrhage (8). Fortunately, for patients with intact vasculature, most thrombocytopenic bleeding is not immediately life threatening and is ostensibly limited to skin manifestations, such as purpura, petechiae and hematomas, and minor mucocutaneous blood loss. Nevertheless, because of the risk of intracerebral hemorrhage, it is common practice for prophylactic platelet transfusion to be administered to patients with platelet counts less than 10,000/μL to provide a hemostatic margin of safety (2,9,10 and 11). Increasingly, evidence is accumulating to support the use of a platelet count of 10,000/μL as a safe threshold for prophylactic platelet transfusion in stable, thrombocytopenic patients (12,13,14,15,16 and 17). Some transfusion specialists advocate a higher threshold of 20,000/μL for patients with complications, such as fever, sepsis, coagulopathy, or falling platelet count (16,18).

For surgical procedures and significant bleeding, platelets are typically transfused to maintain a blood platelet count greater than 50,000/μL (2,9,11). Similarly, platelet transfusions sufficient to produce a platelet count increment of approximately 50,000/μL are indicated for patients undergoing surgical procedures or who have a significant hemorrhage if there exists thrombocytopathy related to medication or disease. Fortunately, should microvascular bleeding related to thrombocytopenia or thrombocytopathy develop in the operating room, it is usually readily apparent and readily treatable with platelet transfusion. Transfusion may be indicated even for patients with platelet counts of 50,000 to 100,000/μL in a setting of massive hemorrhage and extensive vascular injury, particularly when hemorrhage is ongoing and platelet counts are rapidly falling (9,11). The decision to transfuse in that clinical setting depends on clinical judgment related to the risk of continued blood loss.

A traditional dose of platelet concentrate is one unit per 10 kg of patient body weight or four units of platelet concentrate per square meter of the patient's body surface area. For infants and children, the dose also may be calculated as 10 mL/kg body mass. However, this represents an inexact approach because the platelet dose will depend on the platelet content per unit, which is variable. Assuming splenic sequestration of 33% in a 70-kg adult and an average platelet count per unit of 8×10^{10}, the transfusion of one unit should produce a theoretical platelet count increment of 10,000/μL. In fact, the rule of thumb is to expect an increment of 5,000 to 10,000/μL per unit of platelet concentrate transfused. For infants and children, a dose of 10 mL/kg should increase the platelet count by approximately 50,000/μL. In reality, the increment varies with numerous factors, including (a) the number of platelets infused, (b) the patient's clinical condition, including the rate of platelet consumption or destruction, (c) the storage time of the platelets, and (d) the immunologic compatibility of the donor platelets with the recipient.

In recent years, improvements in platelet preparation techniques have led to higher platelet counts per platelet concentrate. Because of this, quality control data regarding platelet counts per unit should be used in arriving at an appropriate pool size of platelet concentrates to transfuse. Whereas a decade or more ago, pool sizes were typically eight to10 units, today many centers have reduced pool sizes to four to six units. In contrast to platelet concentrates, platelets obtained by apheresis do not need to be pooled from multiple donors to achieve an effective dose. The number of platelets obtained in the collection of apheresis platelets from a single donor is typically equivalent to that in a pool of four to six units of platelet concentrate. The U.S. Food and Drug Administration (FDA) has set minimum standards for platelet content per unit of platelet concentrate (0.55×10^{11}) and apheresis platelet (3.0×10^{11}). Most blood collection centers exceed those minimum standards. Total platelet dosages per transfusion vary to some extent with local practice but generally fall in the range of 3 to 5×10^{11} for adults. The target platelet increment varies with the patient's pretransfusion platelet count and clinical situation. A target increment of 40,000/μL is practical for many clinical settings (19). In chronically thrombocytopenic patients, the therapeutic goal is to keep the platelet count above a trough level of 10,000 to 20,000/μL between transfusions to prevent spontaneous hemorrhage. Some evidence indicates that transfusing platelets to maintain an even higher trough level of platelets in chronically thrombocytopenic patients would actually require fewer transfused platelets overall owing to the lower clearance rate of platelets at higher concentrations (20,21). The relative advantages of fewer transfusions of higher platelet dose versus more frequent transfusions of lower dose is the subject of ongoing debate (21,22).

A useful formula for the evaluation of the response to platelet transfusion is the corrected count increment (CCI): CCI = [(Posttransfusion platelet count/μL - Pretransfusion platelet count/μL) × (Body surface area in m^2)] ÷ [Number of platelets transfused × 10^{-11}]. Because the CCI corrects for differences in

intravascular volume, which is proportional to the body surface area, and for the actual number of platelets transfused, it provides a better basis for comparison of transfusion responses than platelet count increments alone. Most platelet transfusion recipients should have CCIs greater than 7,500. A patient with CCIs repeatedly less than 5,000 is considered refractory to platelets (23). Patients without obvious risk factors for platelet refractoriness, such as HLA antibodies, platelet-specific antibodies, sepsis, disseminated intravascular coagulation (DIC), hemorrhage, immune thrombocytopenic purpura (ITP), splenomegaly, or some medications (e.g., amphotericin B), have an average CCI of 14,000 to 15,000 (24,25 and 26).

Plasma

The transfusion of fresh frozen plasma (FFP) is indicated for bleeding or prophylaxis of bleeding in the setting of a documented coagulopathy (2,10,27). FFP should not be transfused solely as a volume expander. FFP contains approximately the same levels of all coagulation factors that are present *in vivo*. By definition, 1 mL of plasma from a normal individual contains one unit of coagulation factor activity. Thus, a unit of FFP with an approximate volume of 200 mL contains approximately 200 units of each coagulation factor, as well as approximately 300 to 700 mg/dL of fibrinogen. FFP is indicated for the replacement of multiple coagulation factor deficiencies and for congenital deficiencies of some specific factors for which concentrates are not yet available or licensed, e.g., factors II, V, VII, X, XI, and XIII. It should be noted that cryoprecipitate is not a concentrated form of FFP but rather a concentrated form of a few factors only (see below). Because of this, FFP, not cryoprecipitate alone, should be used for multiple factor deficiencies.

FFP is not the treatment of choice for hemophilia A (factor VIII deficiency), hemophilia B (factor IX deficiency), von Willebrand's disease (vWD), or antithrombin deficiency. The preferred treatments are commercial concentrates, which are viral attenuated as a result of solvent/detergent (SD) treatment, heating, or monoclonal antibody purification or because they are produced using recombinant DNA techniques. Moreover, purified or partially purified concentrates deliver the required dose in a smaller volume, thereby avoiding risks of volume overload in some recipients.

A newer plasma preparation, SD-treated plasma (SD-plasma) is available, which is a pooled product that is treated with the solvent tri(n-butyl)phosphate and the detergent Triton X-100 to inactivate lipid-enveloped viruses (28,29). Although no viral inactivation technique can guarantee complete safety, SD-plasma has a greatly reduced risk of transmission of hepatitis B virus (HBV), hepatitis C virus (HCV), human immunodeficiency virus (HIV), and presumably other lipid-encapsulated viruses (28,29). It also is a uniform product with a standard 200-mL volume, as opposed to the more variable volume and content of FFP. However, SD-plasma compared with FFP increases donor exposures because it is derived from pooled plasma from as many as 2,500 donors. Moreover, the SD treatment used for SD-plasma is not effective against nonenveloped viruses, such as hepatitis A and Parvovirus B 19. Concern exists over the theoretical emergence of a new pathogenic, nonenveloped viruses whose risk of transmission would be increased by SD-plasma because it is a pooled product from many donors. For these reasons, as well as added cost and an exclusive distribution network, SD-plasma has not received immediate, universal acceptance in the United States. SD-plasma does contain significant levels of antibodies to hepatitis A virus (HAV) and Parvovirus B 19, which potentially have neutralizing activity for those particular viruses. In addition, the screening of plasma for HIV and Parvovirus B 19 by nucleic acid testing before SD treatment and the discarding of positive units may address specific concerns regarding these two viruses. SD-plasma has proven effective for most plasma indications (29). Although most coagulation proteins are present in SD-plasma at approximately the same levels as in FFP, some losses do occur. In particular, protein C, protein S, and antiplasmin are moderately reduced, and the high molecular weight multimers of von Willebrand factor (vWF), which have the highest vWF activity, are substantially diminished (28,29). The impact of these reductions is not expected to be significant for most plasma indications, although their impact on the effectiveness of SD-plasma in DIC has not been established.

Another more recently developed product is donor-retested (or delayed-release) FFP (DR-FFP) (29). DR-FFP is FFP that is not released until the donor tests negative for transfusion-transmitted infections both at the time of the initial donation and then again in repeat testing 16 weeks after the donation (29). The repeat testing eliminates most of the window period before the development of detectable antibodies directed against the viruses that are screened for during donor testing. As a result, DR-FFP is safer in theory than standard FFP.

Another plasma alternative to FFP is cryoprecipitate-reduced plasma. This is the plasma/supernatant portion that remains after cryoprecipitate is removed from FFP. It has reduced levels of the cryoprecipitable coagulation proteins (see section on cryoprecipitate). Because of reduced levels of high molecular weight multimers of vWF, it has been used to treat patients with thrombic thrombocytopenic purpura (TTP), especially those refractory to FFP (30,31). Table 72.1 summarizes the properties of various plasma preparations.

TABLE 72.1. PLASMA PREPARATIONS USED IN TRANSFUSION THERAPY[a]

Name	Content	Preparation	Use	Comments
FFP	All plasma proteins present	Frozen within 8 hr of collection	General	Single-donor product; multiple units usually transfused; some variability in volume and coagulation factor content
Solvent/detergent-treated plasma	Reduced content of largest vWF multimers, protein C, S and antiplasmin	Plasma treated with tri-*n*-butyl phosphate and Triton X-100 to inactivate liquid-enveloped viruses	General	Pooled product from as many as 2,500 donors, nonlipid enveloped viruses not inactivated; uniform 200 ml volume
Donor retested FFP (delayed-release FFP)	All plasma proteins present	FFP from donors who have been retested ≥112 d after donation and found negative for infectious disease markers	General	Not widely available; decreased risk only for infectious agents tested
Cryoprecipitate-reduced plasma	Reduced content of vWF, fibrinogen, factors VIII and XIII	Plasma remaining after removal of cryoprecipitate from FFP	TTP	Reduced large vWF multimers

[a]Other products, such as plasma, liquid plasma, recovered plasma, and source plasma are generally not used directly in transfusion therapy.
FFP, fresh frozen plasma; vWF, von Willebrand factor; TTP, thrombotic thrombocytopenic purpura.

The level of many coagulation factors required for hemostasis is approximately 20% to 30% of normal (32,33 and 34). For adult patients with a prolonged prothrombin time (PT) or activated partial thromboplastin time (aPTT) (1.5- to 1.8-fold increase relative to control) who demonstrate abnormal bleeding that is not immediately life threatening, two units of FFP (6 to 7 mL/kg) are a typical starting dose. Further therapy is based on the results of repeat coagulation testing. For significant hemorrhage or immediately life-threatening bleeding in adult patients with a prolonged PT or aPTT, transfusion of four or five units of FFP (12 to 15 mL/kg) may be indicated, followed by repeat coagulation testing. Four to five units of FFP provide approximately 800 to 1,000 mL of plasma, which is equivalent to approximately 30% of the plasma volume of a 70-kg recipient and, therefore, approximately 30% of normal levels of coagulation factors. Intravascular factor recovery will not be 100% for all factors, but because starting levels may be greater than zero, four or five units are a reasonable dose in adults to urgently obtain a minimum necessary hemostatic level of all factors. Dosages should be adjusted based on kilograms of body mass for pediatric patients or adults who are significantly larger or smaller than 70 kg. Because

multiple-unit plasma transfusions carry a risk of fluid overload, physicians must assess before transfusion whether the patient can tolerate the added intravascular fluid volume.

Cryoprecipitate

Cryoprecipitate, formally named cryoprecipitated antihemophilic factor, contains fibrinogen, vWF, factors XIII and VIII, and fibronectin along with some other cryoprecipitable plasma proteins. Cryoprecipitate is not a concentrated form of FFP because it does not contain clinically significant amounts of the other coagulation factors. Cryoprecipitate originally was predominantly intended as a source of factor VIII for the treatment of hemophilia A. However, commercial factor VIII concentrates, including recombinant factor VIII, which have a reduced or zero risk of hepatitis and HIV transmission, have been developed. Therefore, cryoprecipitate transfusion is no longer indicated for factor VIII replacement. A virus-attenuated factor VIII/vWF concentrate also is now FDA licensed and replaces cryoprecipitate as the preferred source of vWF. Cryoprecipitate is more important as a source of fibrinogen. A fibrinogen level greater than 50 mg/dL is necessary for efficient hemostasis (34). Fibrinogen repletion is generally recommended when fibrinogen levels are less than approximately 80 to 100 mg/dL in the setting of hemorrhage or in preparation for surgery (2,10). Similarly, for patients with dysfibrinogenemia, infusion of cryoprecipitate sufficient to give an increase in fibrinogen level of 80 to 100 mg/dL is indicated for the control of hemorrhage or in anticipation of an invasive procedure.

The content of fibrinogen in one unit of cryoprecipitate ranges from 100 to 350 mg and is often in the range of 200 to 250 mg. It is important to check with your blood supplier for quality control data regarding the fibrinogen content of cryoprecipitate because it is useful in dosage calculations. In a 70-kg adult with a 3-L blood volume, one unit of cryoprecipitate will theoretically raise the fibrinogen level by approximately 6 to 7 mg/dL, assuming 100% recovery of transfused fibrinogen in the intravascular space and an average content of 200 mg fibrinogen/unit. An estimate of the number of units of cryoprecipitate to transfuse can be calculated by dividing the desired increment in fibrinogen (in mg/dL) by 6. The dosage calculation can also be approximated using the following formula, assuming a desired level of 100 mg/dL of fibrinogen *in vivo* and a 200-mg average content of fibrinogen/unit:

$$\#\ \text{units} = \frac{(\text{plasma volume in dL})\ (100 - \text{X in mg/dL})}{200\ \text{mg}}$$

where X is the recipient pretransfusion fibrinogen level.

This is an inexact calculation because intravascular recovery may not be 100% and because consumption in many clinical settings may be ongoing. Other dosing approaches include transfusing one unit per 10 kg of body mass, which increases fibrinogen concentration by approximately 50 mg/dL, or to empirically transfuse a pool of 10 units. No matter which dosing approach is used, the patient's fibrinogen levels should be remeasured soon after transfusion to ensure that an appropriate level has been obtained.

Cryoprecipitate also is used in the preparation of fibrin glue or sealant, which is applied topically in surgical procedures rather than being infused intravenously (35,36 and 37). To make the fibrin glue, bovine thrombin is added along with calcium to cryoprecipitate. The thrombin converts fibrinogen to fibrin, which forms the glue. Topical fibrin glue has found multiple uses in various surgical procedures because it forms a flexible seal over suture lines. Autologous cryoprecipitate can be used to make fibrin sealant to address concerns about infectious risks. A commercial fibrin sealant that is treated to reduce the risk of viral transmission is available.

Granulocytes

Patients with severe neutropenia or neutrophil dysfunction are at risk of fatal bacterial, yeast, and fungal infections. Severe neutropenia

is seen in patients undergoing myeloablative therapy during various cancer treatments. Neutrophil dysfunction is seen in patients with chronic granulomatous disease. Neonates may have both severe neutropenia and relative neutrophil dysfunction. Transfusions of granulocyte concentrates have been employed to treat such patients when they have severe infections that are refractory to antibiotic or other therapy (38,39). Unfortunately, granulocyte transfusions administered to treat infections often have met with mixed success at best. For that reason granulocyte transfusions are uncommon at the present time. Therapeutic failures are generally attributed to inadequate doses of granulocytes in the transfused products. The availability of recombinant preparations of cytokines that stimulate endogenous granulocyte production, such as granulocyte-colony stimulating factor (G-CSF), has provided a pharmacologic means to shortening periods of neutropenia in some patients. This has reduced the need for granulocyte transfusions.

Granulocyte concentrates are typically obtained by apheresis from single donors, but can be prepared from buffy coats of single units of whole blood. The granulocytes are labile during storage and should be transfused as soon as possible and no later than 24 hours after collection. Regulatory standards require at least 1×10^{10} granulocytes per apheresis unit of granulocyte concentrate in at least 75% of units tested (40). Although the required minimum dose for therapeutic efficacy has not clearly been determined, transfusions of fewer than 1×10^{10} granulocytes are unlikely to be of benefit in adults. Adults typically receive one unit of granulocyte concentrate per day and neonates approximately 0.5 to 1.0×10^{9} granulocytes per kilogram per day (40).

At present, prophylactic granulocyte transfusions are not indicated because of their unlikely or uncertain efficacy. Granulocyte transfusions should be limited to the treatment of patients who are severely neutropenic (often defined as fewer than 3,000 neutrophils/μL in neonates less than 2 weeks old and fewer than 500 neutrophils/μL for other age groups) or who suffer chronic granulomatous disease and who have severe infections that are unresponsive to antibiotic therapy (38,39). Because one transfusion is unlikely to be adequate, a commitment to a minimum course of 4 to 7 days of daily granulocyte transfusions is recommended. Granulocyte transfusions generally should be continued until the marrow recovers (greater than 500 neutrophils/μL in the peripheral blood for adults) or the infection resolves.

Recent attempts to stimulate donor granulocyte levels by pretreating donors with G-CSF plus corticosteroids have resulted in significantly higher yields of granulocytes per collection (41,42). Because G-CSF does have some side effects, however, it is unclear at this time whether its use in this setting in healthy, volunteer donors will become widely accepted. Although this collection technique appears promising, experience is too limited at this time to know whether transfusion of greater numbers of granulocytes obtained with this new preparative regimen will improve the therapeutic usefulness of granulocytes.

Whole Blood

Whole blood is now almost exclusively used in the setting of autologous transfusion. Whole-blood transfusion is not an effective means of supplying patients with blood elements other than RBCs. The blood bank storage conditions of whole blood are not optimal for the preservation of platelets and coagulation factors. Platelets and factors V and VIII are particularly labile (43). Therefore, the production, storage, and transfusion of whole blood is generally neither an efficient nor a cost-effective approach to transfusion therapy. For patients requiring the transfusion of RBCs only, units of whole blood in some cases may even be undesirable because of their larger volume than RBC units, i.e. approximately 450 mL for whole blood compared with approximately 320 mL for additive-solution RBCs.

Fresh whole blood within several hours of collection still contains viable platelets and normal levels of all coagulation factors. If fresh whole blood were readily available, it would be the component of choice for a small number of patients who require the simultaneous transfusion of RBCs, platelets, and plasma, such as some patients undergoing massive transfusion. The use of fresh whole blood in that situation would limit the number of donor exposures for the transfusion recipient, which could decrease the risk of immunologic or infectious complications of transfusion. However, fresh whole blood from allogeneic donors is an impractical concept today due to the necessity for infectious disease testing of donor blood before transfusion. Moreover, the limited number of patients who would benefit from fresh whole blood is too unreliable a population on which to plan a blood inventory.

Conversely, autologous blood, particularly when collected by a hospital transfusion service, as opposed to a blood center, is often not fractionated into other components. Many transfusion services are not equipped for blood component production. Moreover, because most patients donating autologous blood have no predicted need for blood components other than RBCs, the time and expense of a fractionation effort are not justified. Autologous blood collected immediately preoperatively in the operating room during acute normovolemic hemodilution is fresh whole blood. As such, it contains functional platelets and coagulation factors if it is transfused operatively or immediately postoperatively. Various types of autologous blood collection are described in more detail below (section on autologous transfusion).

TRANSFUSION IN SPECIFIC CLINICAL SETTINGS

Part of "72 - Transfusion Therapy"

Hemorrhage, Surgical Blood Loss, and Anemia

Acute Blood Loss

The American College of Surgeons has classified acute blood loss based on the volume of blood lost and the corresponding clinical manifestations (44). The classification is shown in Table 72.2, along with recommendations for replacement fluids for these varying degrees of blood loss. A loss of as much as 15% of blood volume is classified as a class I hemorrhage. In an otherwise healthy individual, this relatively low volume of blood loss results in minimal symptoms, e.g., perhaps mild tachycardia, and should not require replacement therapy (45,46). The acute blood loss of 15% to 30% of blood volume is a class II hemorrhage,

which produces tachycardia, decreased pulse pressure, anxiety, and restlessness. When a class II hemorrhage reaches a blood loss of 20% to 30%, the patient's extracellular space should be maintained with the infusion of crystalloid solutions. In hemorrhage that results in 30% to 40% blood loss (class III hemorrhage), the patient shows signs of shock, including tachycardia, tachypnea, systolic hypotension, and altered mental status. Otherwise healthy, nonanemic patients still may receive crystalloid alone for class III hemorrhage. However, previously anemic, elderly patients, or patients at increased risk for organ ischemia may require RBC replacement for this or smaller blood losses depending on the degree of preexisting anemia or susceptibility to ischemia. Loss of more than 40% of blood volume (class IV hemorrhage) is life threatening and is characterized by severe shock. The blood volume should be restored rapidly with the infusion of crystalloid solution. Moreover, the restoration of the oxygen-carrying capacity of the blood with RBC transfusion in class IV hemorrhage may be indicated even in otherwise healthy individuals.

TABLE 72.2. CLASSES OF HEMORRHAGE

Class	% Blood Loss	Approx. Blood Lost (mL)	Approx. RBC Unit Equivalents	Replacement Fluid
I	<15	<750	<1.5	None
II	15–30	750–1,500	1.5–3.0	Crystalloid first; RBCs rarely needed if hematocrit previously normal
III	30–40	1,500–2,000	3.0–4.0	Crystalloid first; then possibly RBCs
IV	>40	>2,000	>4.0	Crystalloid; then RBCs

RBCs, red blood cells.

Crystalloid solutions, such as lactated Ringer's solution, as opposed to colloid solutions or RBCs, are the preferred initial choice for volume expansion (44). Oxygen transport is a function of both the hemoglobin content of blood and blood flow, which is volume dependent. In patients with normal hematocrits before the acute blood loss, the loss of blood volume, rather than red cell mass, is usually the most critical concern. Accordingly, volume loss may require immediate correction. For example, in patients with normal hematocrits, a loss of 50% of their RBC mass may decrease the hematocrit to 20% to 25%. This is sufficient to provide adequate oxygen transport for an otherwise healthy adult as long as the blood volume is sufficient for tissue perfusion. However, the loss of 50% of the blood volume leads to shock and death in most patients. Thus, volume replacement, not restoration of the RBC mass, is usually the first priority in acute blood loss. In fact, the transfusion of RBCs during the initial treatment of acute blood loss may even be undesirable owing to their increased viscosity compared with crystalloid solutions. The lower viscosity solutions may more rapidly restore perfusion in smaller blood vessels and in the extracellular space.

When urgent RBC transfusion is indicated for large-volume acute blood loss, ABO- and Rh-identical or -compatible RBCs, and preferably cross match–compatible, RBCs should be used if at all possible. Accordingly, it is important for the patient care team to send a blood specimen from an acute trauma patient to the laboratory immediately so that as much time as possible is available to find compatible blood. The initial correction of the patient's blood volume loss with crystalloid solutions should in most cases sufficiently stabilize the patient hemodynamically to allow the laboratory time to ABO and Rh type the patient. This requires approximately 10 minutes after receipt of the specimen. If the clinical situation permits a 30- to 45-minute delay in transfusion, an antibody screen and a major cross match with the donor unit(s) can and should be performed (40,47). An objective clinical assessment by the patient care team followed by clear communication with the laboratory is required to optimize pretransfusion testing in the setting of acute blood loss. The use of ABO- and Rh-identical units permits better use of RBC inventories. In addition, antibody screening and cross matching lessens the possibility, already small for most patients, that the transfused RBCs will produce a hemolytic transfusion reaction. With the use of ABO-compatible uncross matched units, the primary concern is the possibility that the recipient may have a non-ABO alloantibody. An estimated 0.04% of individuals will have a non-ABO alloantibody if they have not been previously transfused or pregnant. Approximately 0.3% of previously transfused, multiparous women and approximately 5% to 30% of chronically transfused patients, such as those with sickle cell anemia or thalassemia, have non-ABO RBC alloantibodies (48,49 and 50). However, even if the recipient has a non-ABO alloantibody and receives RBCs incompatible with that antibody, it is unlikely that a life-threatening intravascular hemolytic transfusion reaction will occur. Most, although not all, clinically significant non-ABO RBC alloantibodies are IgG molecules that do not fix complement to C9. Thus, they usually produce extravascular, not intravascular, hemolysis. Should a recipient of an emergency transfusion of noncross matched RBCs be found in subsequent testing to possess a RBC alloantibody, antigen testing should be performed retrospectively on any units transfused so that a potential hemolytic reaction can be anticipated.

If the short delay in the provision of RBCs due to the need to perform ABO and Rh typing of the recipient is life threatening, then group O, preferably Rh-negative, uncross matched RBCs should be transfused. For this reason, group O, Rh-negative RBCs are available in monitored refrigerators in emergency departments and trauma centers at some medical centers. These units can be given to those patients for whom the potential risk of incompatibility is small when compared with the risk of delaying the transfusion (51,52 and 53). In general, group O, Rh-negative, uncross matched RBCs can be used for emergency RBC transfusion with little risk for most patients. The infusion of a small amount of plasma (approximately 40 mL in units of additive-solution RBCs) containing anti-A or anti-B antibodies into an ABO-incompatible recipient usually does not pose a significant risk. The recipient may develop a positive direct antiglobulin test, but usually without significant hemolysis or a clinical reaction. Nevertheless, as with ABO/Rh-identical uncross matched RBCs, there

is the small possibility that the recipient may have a non-ABO alloantibody, particularly if the patient has been previously transfused or pregnant. Despite the relative safety of an emergency transfusion of group O, Rh-negative RBCs, clear guidelines for their use should be established to avoid overusage in acute trauma settings because they are a valuable and relatively scarce resource that constitutes only approximately 6% of the available RBCs. Table 72.3 summarizes the various possibilities for abbreviating compatibility testing for emergency transfusions and the corresponding estimated time required for testing in the blood bank.

TABLE 72.3. COMPATIBILITY TESTING FOR EMERGENCY RBC TRANSFUSION

Transfusion Urgency	ABO, Rh(d) Type	Type	Antibody Screen	Cross Match	Testing[a,b] Time
Lowest	ABO, Rh(d) identical or compatible	Yes	Yes	Yes	30 min
	ABO, Rh(d) identical or compatible	Yes	No	Yes	15 min
↓	ABO, Rh(d) identical or compatible	Yes	No	No	10 min
	ABO, Rh(d) identical or compatible	Yes	Yes[c]	No	0 min
Highest	O, Rh(d)–negative	Yes	No	No	0 min

[a]Testing times assume that the cross match is done at room temperature using an immediate spin only and no previous presence of a RBC alloantibody.

[b]Testing time does not include transportation times (variable) or all administrative steps (requiring several additional minutes).

[c]ABO, Rh(d) typing and antibody screen performed earlier on current valid serum specimen (within 3 days of transfusion).

If RBCs are transfused before Rh(d) typing of the recipient, it is preferable to use Rh(d)-negative units, if available. Approximately 15% of recipients will be Rh(d) negative and will be at significant risk of having or developing anti-D. The D antigen is the most immunogenic non-ABO red cell determinant and will result in antibody formation in approximately 85% of Rh(d)-negative individuals receiving 200 mL or more of Rh(d)-positive RBCs (54). The most significant risk is to Rh(d)-negative women of childbearing potential. If they develop an anti-D, any future Rh(d)-positive children may be at risk for Rh hemolytic disease of the newborn. All females of childbearing potential should be assumed to be Rh(d) negative until they are Rh(d) typed and should be transfused with Rh(d)-negative RBCs if at all possible. If Rh(d)-negative females of childbearing age receive Rh(d)-positive RBCs because of the lack of availability of Rh(d)-negative units, they can be given Rh immune globulin (RhIG) for the prevention of Rh alloimmunization.

An intravenous formulation of RhIG has been licensed in the United States and is now the preferred prophylaxis of anti-D formation for the transfusion of one unit or more of Rh(d)-positive RBCs because many injections of the intramuscular formulation would be required in that setting. According to manufacturer's recommendations, the IV RhIG can be administered as long as no more than 20% of the patient's RBCs are Rh(d) positive. Above 20%, the effects of an extravascular hemolytic reaction may become significant. Because anti-D is a relatively common antibody, the use of Rh(d)-negative RBCs for urgent transfusion in both males and females also eliminates the risk of an anti-D-mediated delayed extravascular hemolytic reaction in already alloimmunized patients.

Massive Transfusion

Massive transfusion is the replacement of the patient's total blood volume by transfusion within a 24-hour period. A practical definition of massive transfusion in adults is the transfusion of 10 or more units of RBCs within a 24-hour period because that approximates the normal total adult RBC mass. Although infusion of a large volume of RBCs in a short time period may be lifesaving, it also poses hazards to the transfusion recipient that are not seen after the transfusion of smaller volumes. It is important for the clinical care team to be aware of the complications of massive transfusion so that measures can be taken to prevent, monitor, or treat their development. Some of the major adverse effects of massive and rapid replacement of lost blood volume include dilutional coagulopathy and thrombocytopenia, citrate toxicity, acid-base and electrolyte disturbances, and hypothermia. Another consideration is the possibility of depletion of the hospital transfusion service/blood bank's supply of ABO, Rh(d)-identical blood components for a patient undergoing massive transfusion. This may necessitate switching ABO types for RBCs and plasma.

Dilutional Coagulopathy and Thrombocytopenia

With massive hemorrhage and subsequent large-volume transfusion, platelet and coagulation factor levels may be depleted owing to their consumption and/or dilution (55,56,57,58 and 59). Consumption is due to activation of primary hemostasis and the coagulation cascade secondary to trauma or injury. Dilution is owing to replacement of blood with crystalloid or colloid solutions or units of stored RBCs, all of which are deficient in functional platelets and coagulation factors. Massive transfusion, however, does not always result in platelet or factor deficiencies sufficient to warrant platelet or FFP transfusion. Theoretical models predict that massive transfusion with replacement of one blood volume will leave behind 37% of the original blood volume (60). If starting levels of platelets and coagulation factors are normal, dilution to 37% of original levels should leave hemostatically effective levels of platelets and coagulation factors, assuming no further production or consumption. In fact, the liver may continue to produce coagulation factors, and platelets may enter the circulation from extravascular sites, thus blunting the development of dilutional coagulopathy and thrombocytopenia.

Because of the likelihood of concurrent consumption of coagulation factors or platelets owing to hemorrhage and/or the development of disseminated intravascular coagulation, the extent of platelet or coagulation factor deficiency cannot be predicted reliably from the volume of blood lost or the number of units of components transfused. Accordingly, transfusion of platelets or FFP should be based on observed clinical bleeding as well as on the laboratory measurement of platelet counts and coagulation testing. Transfusion should not be based on predetermined formulas

linked to the number of RBC units transfused (61). Replacement formulas may lead to unnecessary transfusion, particularly if less than 1.0 to 1.5 blood volumes have been replaced (62,63). In addition, component transfusion guided by replacement formulas may be insufficient if a consumptive process such as disseminated intravascular coagulation develops.

A key to appropriate replacement therapy in massive transfusion, is the timely measurement of platelets and coagulation status. Empirical platelet and FFP transfusion should be considered only if the delays inherent in testing and reporting of laboratory results could be life threatening due to continued and uncontrolled hemorrhage. Specimens should be sent to the laboratory as soon as possible and at appropriate intervals thereafter so that further transfusion therapy can be guided by objective laboratory evidence. Plasma transfusion is indicated in the getting of abnormal bleeding with PT or aPTT greater than 1.5 times the normal control (2,10,34). Cryoprecipitate transfusion for fibrinogen (and possibly vWF) supplementation is indicated for fibrinogen levels less than 80 to 100 mg/dL in the setting of hemorrhage. Platelet transfusion is indicated for ongoing hemorrhage if the platelet count is less than 80,000 to 100,000/µL. A platelet threshold of 50,000/µL may be appropriate if hemorrhage is less severe.

Citrate Toxicity

Citrate is used as the anticoagulant in preservative solutions because of its ability to chelate calcium, which is required for the coagulation cascade. Citrate usually is rapidly metabolized by the liver to generate bicarbonate and generally does not result in systemic anticoagulation of the transfusion recipient. However, when large volumes of citrate-containing plasma are rapidly infused, the body's capacity to metabolize citrate may be transiently exceeded. When this happens, ionized calcium levels may be lowered owing to chelation by the unmetabolized citrate. The threshold for exceeding the body's capacity to metabolize citrate depends on liver function. However, even patients with normal liver function who are transfused with whole blood (or presumably its blood component equivalent) at a rate of greater than or equal to 1 L/10 min may develop signs of hypocalcemia (64,65,66 and 67). The effects of hypocalcemia range from minor circumoral paresthesias and muscle tremors to more severe complications such as tetany and cardiac effects. The cardiac effects include prolonged QT intervals and depression of T waves, which could potentially lead to ventricular fibrillation and cardiac arrest. Hypomagnesemia, presumably owing to chelation of magnesium by citrate, has also been reported (68). However, clinical complications have not been well documented.

Citrate toxicity is less likely to result from RBC transfusion today than from whole-blood transfusion in the past. This is because additive-solution RBC units have substantially less plasma and thus less anticoagulant per unit compared with whole blood. Citrate toxicity remains a concern in patients with liver failure who receive rapid transfusions of large amounts of plasma-containing components such as FFP and platelets. Ionized calcium levels should be measured in these patients, and they should be treated with supplemental calcium if the levels are sufficiently low or if prolonged QT intervals or signs of tetany are seen. Early mild reactions may be treated by slowing the rate of transfusion, if permitted by the clinical situation (67). Calcium infusion itself is not without risk and has been associated with the development of ventricular arrhythmias and cardiac arrest (69). Calcium must not be added to a unit of RBCs or infused simultaneously through the same line because it will recalcify the plasma and cause clots to form.

Acid-Base and Electrolyte Disturbances

Stored blood and components have an acidic pH owing to citrate/citric acid buffering and may become more acidic owing to the accumulation of lactic acid with increasing storage time (70). However, this acid load is not likely to significantly worsen the usually already existing lactic acidosis that results from periods of tissue hypoperfusion after massive and acute blood loss. On the contrary, owing to the metabolism of transfused citrate, which results in the generation of bicarbonate, massively transfused patients may suffer a delayed metabolic alkalosis (71). Extracellular potassium levels increase in stored blood at a rate of approximately 1 mEq/day during the first 3 weeks of storage (70). The total extracellular potassium load of stored transfused RBCs may approach 20 mEq per unit. However, even massively transfused patients rarely develop more than transient hyperkalemia (72). Hypokalemia has more often been reported in massive transfusions owing to the development of a metabolic alkalosis from citrate metabolism (see above) (73). The alkalosis provokes a compensatory transmembrane shift of hydrogen ions into the intravascular space. This hydrogen ion shift is electrically balanced by an intracellular movement of potassium ions, which, if of sufficient magnitude, can result in hypokalemia.

Hypothermia

The rapid transfusion of large volumes of cold (1° to 6°C) RBCs at a rate greater than 100 mL/min may lower the patient's core temperature. A drop in sinoatrial node temperature to below 30°C has been associated with the development of ventricular fibrillation (74,75,76 and 77). To avoid this potential complication, RBCs may be administered with a high throughput blood warmer. If the infusion rate imposed by the blood warmer is too slow or if a blood warmer is not available, rapid warming of RBCs by admixture with heated saline is also possible (78,79).

Switching ABO Blood Type

The massive blood transfusion of one or more patients may deplete or threaten to deplete the hospital transfusion service/blood bank of its supply of an ABO, Rh(d) type of RBCs or plasma. In that situation, it may become necessary to switch the ABO type of transfused blood components of the massively transfused patient or other patients. The permissible switches between ABO groups are shown in Table 72.4. In some cases, it also may be necessary to transfuse Rh(d)-positive RBC or platelet units into Rh(D)-negative recipients. This is acceptable in male patients and older female patients beyond childbearing age. Conversely, Rh(d)-positive RBCs and platelets should be avoided in women of childbearing potential owing to the possibility of Rh alloimmunization, which creates a risk of hemolytic disease of the newborn in any future pregnancies. If transfusion of Rh(d)-positive RBCs or platelets is unavoidable in such women, prophylaxis

against Rh alloimmunization with Rh IG should be carried out.

TABLE 72.4. COMPATIBLE ABO TYPES

Patient Type	Donor RBC Type	Donor Plasma Type
O	O	O, A, B, AB
A	A, O	A, AB
B	B, O	B, AB
AB	AB, A, B, O	AB

RBC, red blood cell.

Chronic RBC Transfusion

The two major illnesses that result in life-long requirements for RBC transfusion are sickle cell disease and ß thalassemia major. In sickle cell disease, RBC transfusions are indicated to treat intermittent aplastic and vasoocclusive crises (80,81 and 82). Despite having chronically low hemoglobin levels, patients with sickle cell disease typically have adequate cardiorespiratory compensation at their baseline hemoglobin and hematocrit levels. RBC transfusion is required generally only when these patients experience a decrease in hematocrit from baseline sufficient to induce symptoms of anemia. The goal of transfusion for such crises is to increase the hematocrit to approximately the patient's usual baseline, making sure to maintain the hematocrit below 35% to avoid hyperviscosity. The goal also is to keep the percentage of hemoglobin S-containing cells within the range of 20% to 40% to reduce the likelihood of intravascular sickling, which causes vasoocclusion (3,80). RBC exchange transfusions may be indicated for sickle cell patients with acute chest syndrome, stroke, refractory priapism, or retinal artery occlusion. Chronic RBC transfusions have been administered to some patients who have severe, recurrent vasoocclusive crises such as recurrent cerebrovascular accidents. In ß thalassemia, the major goal of RBC transfusion is to maintain hemoglobin levels in the range of 9.5 to 11 g/dL (3,81). RBC transfusions not only prevent symptoms of anemia but also suppress endogenous hematopoiesis, which in turn prevents bone marrow hyperplasia and skeletal deformities. A chronic transfusion program of RBC transfusions at regular intervals sufficient to suppress endogenous erythropoiesis, known as hypertransfusion, is recommended for ß thalassemia major beginning at a young age.

Patients who receive chronic RBC transfusions are at risk for several adverse effects of transfusion. Owing to frequent transfusion, these patients are at increased risk of febrile, nonhemolytic transfusion reactions. For this reason, the use of leukoreduced RBC units is recommended. In addition, approximately 30% of sickle cell patients and approximately 5% to 10% of ß thalassemia major patients become alloimmunized to one or more RBC antigens (83,84 and 85). Depending on the antibody specificity, this may make finding compatible RBCs difficult. Alloimmunization can be prevented in chronically transfused patients by transfusing units of RBCs that are phenotypically matched not only for ABO and Rh(d) but also for other clinically significant RBC antigens of the recipient. Because patients will make alloantibodies only to antigens that they lack, this approach prevents alloimmunization. However, because most patients do not become multiply alloimmunized, the cost-effectiveness of this approach has been debated (86,87). Some centers follow a compromise approach of providing RBCs that match the recipient's extended RBC phenotype only after formation of the first RBC alloantibody. Even if transfused RBCs are not matched for additional antigens, extended phenotyping of the recipient's RBCs is recommended to help identify alloantibodies should the patient subsequently become multiply alloimmunized. Antibodies to the patient's own RBC antigens can be ruled out.

Patients chronically transfused with RBCs are at risk of developing iron overload or transfusion hemosiderosis. Every unit of RBCs contains approximately 200 mg of iron (approximately 1 mg of Fe^{2+}/mL RBCs). Because the body has no mechanism for excreting iron, the excess iron from chronic transfusions accumulates in organs such as the liver, heart, and endocrine glands. Organ damage can eventually lead to cirrhosis, congestive heart failure, or endocrine dysfunction (88). To decrease or delay the toxic effects of iron overload, iron chelation therapy with deferoxamine is necessary for patients receiving chronic RBC transfusion therapy (88,89). Deferoxamine mobilizes and reduces iron stores via urinary excretion, although the rate of iron removal is slow. With judicious use of RBC transfusion and iron chelation therapy, the risk of iron overload can be reduced.

Autologous Transfusion

Autologous transfusion uses the recipient's own blood or blood components for subsequent reinfusion (90,91 and 92). The goal of autologous transfusion is to eliminate the risks of allogeneic transfusion, such as transfusion-transmitted viral infections and a variety of immune-mediated adverse effects. Autologous blood can be collected preoperatively weeks in advance of surgery or just before surgery in the operating room. Autologous blood also can be salvaged from blood shed in the intraoperative or postoperative periods. The use of autologous blood, however, does not eliminate all risks of transfusion (93,94). Bacterial contamination of the collected blood, which can lead to septic transfusion reactions, can still occur. Also, the possibility of hemolytic reactions resulting from transfusing the wrong patient owing to patient misidentification is not eliminated. Autologous blood, just like allogeneic blood, also may contribute to volume overload in at-risk patients. Collecting or salvaging autologous blood adds expense to medical care compared with the use of allogeneic blood (95,96). The added cost of autologous blood is exacerbated by the waste of approximately half of all autologous units collected preoperatively (97). Moreover, as the ability of infectious-disease testing of allogeneic blood to detect donor viral infections improves through the implementation of such innovations as nucleic acid testing, the relative advantage of transfusing autologous blood decreases. Thus, the cost-to-benefit ratio of autologous blood collection deserves ongoing reappraisal.

Preoperative Autologous Blood Donation

Preoperative autologous blood donations can be stored in liquid form for as long as 35 days as whole blood or for as long as 42 days if additive-solution RBC units are prepared from the autologous blood (40). For a patient to be eligible for preoperative

autologous blood donation, sufficient time must be available before surgery to permit the collection of the desired number of units. In addition, the planned surgical procedure should have a reasonable expectation of requiring the number of units requested by the patient's surgeon (91). A schedule of expected transfusion requirements for specific surgical procedures, commonly known as a maximum or standard blood order schedule, should be prepared jointly by the transfusion service and surgery department. Such a schedule is necessary for determining the number of RBC units the transfusion service should type and screen or cross match for specific surgical procedures. The schedule should call for the availability of cross matched units of RBC if the procedure typically requires transfusion for 10% or more of cases. The number of cross matched units assigned per procedure should be adequate to meet the transfusion needs for 90% of patients who undergo that procedure. This blood order schedule also can be used as a guide for determining patients' eligibility for autologous transfusion and for determining a reasonable number of autologous units to collect. Preoperative autologous blood collection should be considered only if the planned surgical procedure calls for the availability of cross matched RBCs. A typical schedule provides for the donation of one unit per week, with the last unit being donated at least 3 days before surgery. More aggressive collection schedules can be used as long as the patient's hematocrit remains higher than 33% at the time of each donation and provided the patient's intravascular volume is sufficiently maintained (40). Unless intravascular volume replacement is given, the donations should be spaced at least 3 days apart to allow for natural volume repletion. No age limits are established for autologous donation.

The amount of blood collected, 450 mL ± 10%, should be less than 15% of the patient's blood volume. Smaller volumes can be collected from small donors. However, the amount of anticoagulant/preservative solution must be reduced proportionately if the volume collected is less than 300 mL (40). A criticism of preoperative autologous blood donation is that decreasing the patient's hematocrit preoperatively increases the need for operative or postoperative transfusion (90,98,99). Most collection schedules, in which one unit is collected per week, minimally stimulate erythropoiesis, resulting in little increase in overall RBC mass (circulating plus donated) available to the patient at the time of surgery (100,101). This is owing to two limitations: (a) not enough preoperative time to replenish the RBCs lost and (b) failure to stimulate a significant erythropoietic response. Nevertheless, the hemodilution effect may be of some benefit. Patients with lowered hematocrits at surgery will lose less RBC mass for a given volume of blood loss and will benefit from receiving back their predeposited blood, which has the higher preoperative hematocrit, at a time of greater need. This hemodilution effect will be more significant in surgical cases with high blood loss.

Supplemental erythropoietin administration to autologous donors has been shown to improve the yield of autologous collections for patients with hematocrits lower than 40% (102,103 and 104). Although this adds additional cost to autologous collection, it may be a benefit for patients who are at particular risk from allogeneic transfusion, such as patients with multiple RBC alloantibodies for whom compatible RBCs are rare. Erythropoietin also may be administered to patients during the preoperative period for the purpose of increasing the patient's circulating, as opposed to collected, RBC mass, thereby lessening the chance and need for transfusion during surgery. In addition, the stimulation of erythropoiesis preoperatively increases the ability of the bone marrow to replace operative blood loss more rapidly during the postoperative period.

Acute Normovolemic Hemodilution

Acute normovolemic hemodilution (ANH) refers to the collection of units of whole blood in the operating room immediately before surgery after anesthesia is administered. Patients receive simultaneous volume replacement using crystalloid or colloid solutions (105,106 and 107). The blood typically stays in the operating room and may be kept at room temperature for as long as 8 hours until transfused. If not transfused during this time, the blood may be stored for up to 24-hours if placed at 1° to 6°C within 8 hours of collection (40). Normovolemic hemodilution has several benefits. First, it decreases the operative loss of red cell mass because a lower hematocrit causes a smaller loss of red cell mass for a given volume of operative blood loss. Moreover, the blood collected is available for transfusion during or after surgery, when the need for the RBCs may be greater than preoperatively. Furthermore, hemodilution decreases the viscosity of blood and therefore may increase blood flow rate through the microvasculature. Because oxygen transport is a function not only of the oxygen-carrying capacity but also of blood flow, this effect may tend to compensate for the loss in hemoglobin content of blood resulting from the blood donation (108). Hemodilution can be tolerated in many patients because oxygen delivery is normally in excess of that actually required by the tissues.

Blood collected immediately preoperatively and transfused within several hours contains functional coagulation factors and platelets, which may be useful if the patient experiences hemostatic problems postoperatively. For example, this could be a benefit for patients who develop a hemorrhagic syndrome during or after cardiopulmonary bypass (CPB) owing to an acquired platelet defect. Whole blood collected preoperatively before induction of the platelet defect will provide functional platelets and possibly obviate a need for allogeneic platelet transfusion after CPB.

ANH has several advantages over preoperative autologous donation. ANH is more cost-effective because the blood does not have to be tested or stored if it is transfused in the operating room. Also, the patient does not have to make potentially inconvenient trips to a collection facility multiple times preoperatively. ANH has been criticized for yielding limited savings in RBC mass except where large blood loss occurs (109,110). The combined use of ANH with preoperative supplemental erythropoietin (111) or with the intraoperative administration of a hemoglobin-based blood substitute is being evaluated for possible improved benefit. Patients who can potentially benefit the most from ANH are those with a high initial hematocrit and a large expected blood loss who can tolerate the dilutional anemia (112).

Intra- and Postoperative Blood Salvage

Blood salvage refers to the collection of the patient's own blood shed intraoperatively or postoperatively from surgical drainage

sites (113,114,115,116 and 117). After collection, the blood can be reinfused immediately without further processing, or it can be concentrated by centrifugation and saline washed. In all cases, it must be filtered; microaggregate filters are often used. Intraoperative blood salvage is used for patients having surgical procedures with a large anticipated blood loss (greater than two to three units). Intraoperative cell salvage machines or cell washers are available and have been employed for cardiac surgery, hip and knee replacements, as well as major gynecologic, prostate and vascular surgery. These devices recover shed blood from the operative area, centrifuge it, and wash it with isotonic saline. Intraoperative blood salvage has the advantage of being available for many patients who cannot donate autologous blood preoperatively. It is also an adjunct to preoperative autologous blood donation for those patients who have not donated sufficient blood to cover their total transfusion requirements. Blood salvage should not be attempted from any operative site that is potentially contaminated with bacteria. Blood salvage also should not be performed at surgical sites involving malignancies, owing to the theoretical risk of promoting metastases by the intravascular infusion of malignant cells (118,119). The salvage of blood shed postoperatively has been applied most commonly after cardiac and orthopedic surgery. Blood collected postoperatively, except in cases of brisk hemorrhage, usually has undergone extensive coagulation factor activation and defibrination and is substantially hemodiluted (120,121). It is generally partially hemolyzed and contains proinflammatory mediators and tissue debris. Some adverse effects have been reported (121,122 and 123). For this reason, some facilities place volume restrictions on the reinfusion of unprocessed, postoperatively shed blood.

Open-Heart Surgery

Owing to improvements in surgical technique, the use of blood-sparing pharmacologic agents, and acute normovolemic hemodilution, blood component usage has decreased in open-heart surgery. Nonetheless, open-heart surgery accounts for approximately 20% of all RBC units transfused in the United States (124). Significant variability in transfusion practice in open-heart surgery between different medical institutions has been reported (125,126). The variability in transfusion practice is related to variable approaches to the prophylaxis, recognition, and treatment of excessive bleeding that some patients develop after CPB. The etiology of the hemorrhagic syndrome after CPB is probably multifactorial and may be related to platelet dysfunction, mild thrombocytopenia, excessive fibrinolysis, decreased levels of coagulation factors, reduction or alteration in vWF, inadequate heparin neutralization, heparin rebound, and/or protamine excess (127).

The risk factors for excessive bleeding include repeat or redo of coronary artery bypass graft (CABG), greater than 3 to 4 hours of cardiopulmonary resuscitation (CPR) pump/oxygenator time, and cardiac valve replacement surgery (as opposed to CABG) (127). The preoperative bleeding time of patients, many of whom have recently been on aspirin, has not proven to be a reliable predictor of postoperative bleeding (128). Aspirin only modestly prolongs the preoperative bleeding time in most patients (i.e., 2 to 3 minutes) and does not usually lead to excessive bleeding. It is important for the surgeon to distinguish surgically correctable bleeding from microvascular bleeding that is amenable to transfusion. In the immediate postoperative period, excessive bleeding can be recognized as mediastinal drainage of more than 100 mL/h. Very high volumes of chest tube output, defined in one study as more than 300 mL in the first hour, more than 250 mL in the second hour and 150 mL or more thereafter, likely indicates a surgically correctable source of bleeding (129). Surgically correctable bleeding is localized to the operative site and will be seen postoperatively in the chest tube output only. Bright red, pulsatile chest tube bleeding is an obvious indicator of surgical bleeding. Microvascular bleeding, conversely, is generalized. Bleeding or oozing may be seen throughout the surgical field, including all surgical sites (e.g., sternotomy wound, graft harvest sites), as well as intravenous sites. Petechiae and/or purpura may be seen. Surgical bleeding usually requires returning the patient to the operating room for surgical resternotomy and repair of the bleeding source. Transfusion therapy alone would be futile for such cases. Conversely, transfusion therapy for excessive microvascular bleeding is appropriate.

Platelets

During surgery requiring mechanical CPB, platelets exposed to the foreign synthetic surfaces of the extracorporeal blood circuit undergo partial activation and degranulation. As a result, varying degrees of thrombocytopathy can develop (127,130,131 and 132). In anticipation of this thrombocytopathy, some surgical and transfusion specialists recommend that patients have a minimum preoperative platelet count of 100,000/µL. Platelet transfusions are not indicated while patients are still undergoing CPB because the transfused platelets will develop a functional defect. Routine prophylactic platelet transfusions administered after the patient has been removed from CPB are not indicated because the majority of patients do not suffer excessive postoperative bleeding. A controlled trial of routine administration of platelet concentrates failed to show any advantage (133). The risk factors for excessive postoperative bleeding, such as the performance of a redo CABG and prolonged time of CPB (more than 3 hours) and preoperative aspirin use, are not strong enough predictors to warrant prophylactic platelet transfusions. The decision to transfuse platelets must be made largely on clinical grounds. Mild to moderate thrombocytopenia is common after CPB because platelets are consumed during extracorporeal circulation. Thrombocytopenia after CPB is unlikely, in itself, to contribute to excessive bleeding in most cases because platelet counts rarely fall below 50,000/µL. However, a seemingly adequate circulating platelet count can be misleading because those platelets may be dysfunctional owing to the thrombocytopathy associated with CPB. Thus, after CPB, platelets should be transfused when bleeding is excessive and of a diffuse, microvascular nature (e.g., generalized oozing from surgical and intravenous sites), regardless of the patient's platelet count.

Plasma

Patients undergoing open-heart surgery are anticoagulated with heparin during CPB to prevent clotting in the extracorporeal circuit. Heparin inhibits coagulation by activating the natural anticoagulant molecule antithrombin. As a result, during CPB, the patient has a prolongation of the PT, aPTT, activated clotting

time, and thrombin time. At the end of CPB, the heparin effect is reversed by administering protamine, which binds to and inactivates the polyanionic heparin molecule. A residual or rebound heparin effect should be considered as a possible cause of excessive bleeding and abnormal coagulation tests after CPB. A heparin effect should be ruled out before considering coagulation factor replacement with FFP. A heparin-prolonged PT and aPTT can be detected and corrected *in vitro* by the addition of hexadimethrine bromide (Polybrene) or heparinase. Similarly, a prolonged thrombin time owing to heparin can be reversed *in vitro* by the addition of protamine or toluidine blue. Heparin characteristically will prolong the thrombin time, but not the reptilase time, whereas true coagulation factor deficiencies prolong both assays. Abnormal coagulation test results owing to residual or rebound heparin should be treated *in vivo* with the appropriate dose of protamine. FFP transfusion in this setting is ineffective because plasma will not inactivate heparin. A mixing study, in which the plasma used in the coagulation assay is a 50:50 mixture of normal control plasma and the patient's plasma, can help make a diagnosis of a coagulation factor deficiency. Abnormal coagulation tests that are not correctable with heparin neutralization *in vitro* but are correctable by mixing study should be considered the result of an actual reduction in clotting factor levels.

Although a moderate reduction in coagulation factor levels and corresponding prolongations in the PT and aPTT are common after CPB, these changes are rarely associated with excessive bleeding (134). Prophylactic plasma transfusions after CPB are not indicated or useful for most patients. FFP transfusion should be reserved for cases of excessive bleeding, when the PT- or aPTT-to-control ratios are greater than 1.5 to 1.8 and not attributable to heparin. This is based on the observation that PT or aPTT ratios greater than or equal to 1.8 are correlated with an 80% to 85% chance of developing diffuse microvascular bleeding and are associated with a reduction in clotting factor levels to less than 20% of normal (34). Plasma should not be transfused based on laboratory results alone, however, because of the poor correlation between abnormal coagulation tests and excess postoperative chest tube drainage (134).

One approach to treating excessive bleeding after CPB that appears to be diffuse and microvascular is first to transfuse platelets for a presumed platelet dysfunction. While platelets are being administered, laboratory testing can proceed to detect any significant coagulopathy and, if present, to determine whether it is heparin-related. A transfusion of platelets will also provide some plasma. For example, a pool of four to five units of platelet concentrate provides approximately 200 to 250 mL of plasma, a volume equivalent to approximately one unit of FFP. The plasma from platelet concentrates stored as long as to 5 days contains nearly the same levels of clotting factors as FFP, with the exception of the labile coagulation factors V and VIII, which decrease by approximately 50% and 30%, respectively, from initial levels (135).

Cryoprecipitate

Fibrinogen should be measured in settings of excessive postoperative bleeding to determine whether supplemental cryoprecipitate transfusion for fibrinogen replacement is indicated. Cryoprecipitate is transfused for laboratory-documented hypofibrinogenemia (fibrinogen less than 80 to 100 mg/dL) in the setting of excessive, diffuse microvascular bleeding after CPB. A fibrinogen level of 50 mg/dL or less is a sensitive predictor of microvascular bleeding (34). Typical doses of cryoprecipitate are one unit per 10 kg body weight or a pool of 10 bags per adult patient. An alternative guideline is to replenish levels to approximately 100 to 150 mg/dL using a calculated number of units as described earlier in this chapter. FFP transfusions also will provide fibrinogen, although not in a concentrated form. Because normal levels of fibrinogen are approximately 150 to 350 mg/dL, one unit of FFP could be expected to provide from 300 to 700 mg of fibrinogen, equivalent to approximately two to three units of cryoprecipitate. The larger volume of FFP compared with cryoprecipitate for a given level of fibrinogen makes FFP a less effective choice. Because cryoprecipitate does not provide significant amounts of coagulation factors other than factor VIII, factor XIII and fibrinogen, cryoprecipitate alone is not indicated for treating reductions in multiple coagulation factor levels.

Autoimmune Hemolytic Anemia

Warm Reactive Autoantibodies

Autoimmune hemolytic anemias are classified according to the thermal amplitude of the RBC autoantibodies (136,137,138 and 139). Approximately 70% of patients with autoimmune hemolytic anemia have warm autoimmune hemolytic anemia (WAIHA), characterized by the presence in the blood of IgG autoantibodies that are reactive with the patient's own RBCs at 37°C. Because of ongoing uptake and destruction of IgG-coated RBCs by the reticuloendothelial system, patients with WAIHA present with symptomatic anemia, i.e., fatigue, pallor, tachycardia, and dyspnea. This can be a life-threatening disorder, leading to cardiac failure and stroke, although for most patients hemolysis can be controlled with appropriate treatment. Therapy is aimed initially at decreasing accelerated RBC destruction with glucocorticoids, which appear to act at least in part by interfering with Fc receptor function and phagocytosis by splenic macrophages (139). For patients who inadequately respond to glucocorticoids, splenectomy is an option. High-dose intravenous IVIG is also a possible short-term option. Immunosuppressive agents, such as cyclophosphamide, azathioprine, and cyclosporine, are used to treat patients refractory to glucocorticoids and splenectomy. RBC transfusions are generally avoided, if possible. Transfusion is usually only of transient benefit because the autoantibodies typically react against all RBCs, causing transfused RBCs to be subject to the same immune destruction as the patient's own RBCs. Transfusion should not be withheld for patients with severe symptomatic anemia who have not yet responded to other treatments (136,139).

The autoantibodies present a laboratory challenge because they may mask the presence of clinically significant blood group alloantibodies in antibody screens and panels. If possible, the patient's RBC should have an extended phenotype determined before the first transfusion, which includes typing for antigens that stimulate common, clinically significant RBC alloantibodies (e.g., Rh, Kell, Kidd). This information is useful for selecting

compatible RBCs in the future should the patient develop RBC alloantibodies. All RBCs are typically incompatible on cross match with serum from a patient with WAIHA. Special RBC adsorption studies to remove the autoantibodies from the patient's plasma are necessary to determine whether an alloantibody is present (47,140,141). Units of RBCs that are least incompatible are often chosen for transfusion if there is variability in compatibility of different units. However, it is not clear whether the risk of hemolysis is lessened by selecting such units. Sometimes warm autoantibodies show some antigenic specificity, such as reactivity to the e antigen of the Rh system. In some cases, the transfusion of RBCs that are negative for the antigenic specificity of the autoantibody has yielded improved outcomes of the transfusion. Leukoreduced RBCs are preferred to lessen the chance of febrile, nonhemolytic transfusion reactions (FNHTRs). Because fever is a symptom of both FNHTRs and hemolytic reactions, the FNHTR will mimic an early hemolytic reaction and necessitate the discontinuation of a needed transfusion. Thus, FNHTRs should be avoided so as not to waste what might be a least incompatible unit. The transfusion of a smaller volume of RBCs (partial or split unit) at one time is generally advisable in patients with AIHA. This lessens the risk of volume overload and clinically significant hemolysis, which these patients are at increased risk of developing.

Cold Reactive Autoantibodies

Most individuals have clinically insignificant, cold-reactive autoantibodies to RBC surface antigens present at low titer (less than 1:32). These cold autoantibodies react more strongly at 0° to 5°C than at higher temperatures (137,138). The autoantibodies commonly have specificity for the I/i blood group antigens and typically are IgM. Rarely, cold reactive autoantibodies will have a broad enough thermal range to react at physiologic temperature (more than 30°C). Cold agglutinin disease is characterized by hemolysis and/or RBC agglutination and vascular occlusion mediated by pathologic cold-reactive autoantibodies when they have a broad enough thermal amplitude and are present at high titer (usually more than 1:1,000) (137,138 and 139). Serum autoantibody titers in cold agglutinin disease can reach 1:10,000 or more. Anemia is usually mild, although a small number of patients may develop a chronic hemolytic anemia. Hemolysis usually occurs via the uptake of C3b-coated RBCs by hepatic macrophages (138). Less frequently, the entire complement cascade is activated causing intravascular hemolysis.

If transfusion is necessary owing to symptomatic anemia or to replace blood loss in patients with clinically significant cold agglutinins, RBCs may be transfused through an FDA-approved blood-warming device. It is usually impractical or unnecessary to attempt to select RBCs that lack the RBC antigen to which the autoantibody is directed. If the cold autoantibody interferes with RBC antibody screens or cross matches, then cold autoabsorption and/or prewarming techniques may be employed (47,138). The patient's extremities, e.g., fingers, toes, ear lobes, and nose, where cold reactive autoantibody binding to RBCs occurs, should be kept warmer than the upper thermal range of the antibody. Splenectomy is typically ineffective in cold agglutinin disease because most RBC destruction occurs in the liver. Plasmapheresis is an option in severe cases of cold agglutinin disease (139). IgM cold agglutinins are large molecules confined more to the intravascular space than other immunoglobulins and therefore more effectively removed by plasmapheresis.

Thrombocytopenia and Thrombocytopathy

Immune Thrombocytopenic Purpura

Patients with ITP experience autoantibody-mediated destruction of their own platelets (142). ITP is categorized as either acute ITP, occurring predominantly in children with an acute onset often after a viral infection, and as chronic ITP, seen more commonly in adults with a more insidious onset and often with a chronic, relapsing course. Transfused platelets in ITP are destroyed rapidly just like the patient's own platelets and are of limited benefit. Attempts to maintain platelet counts above a specific value, such as 20,000/μL, with platelet transfusion can be futile. Fortunately, the bleeding time and the tendency for hemorrhage are often not increased in proportion to the thrombocytopenia (143) because the platelets in circulation are younger due to increased turnover.

Platelet transfusions are generally reserved as a treatment of last resort only for those patients with significant, life-threatening hemorrhage. Because some patients with ITP can have transient platelet count increments immediately after platelet transfusion (144), a clinical trial of platelet transfusion may be warranted for possible temporary benefit for significant bleeding. Otherwise, the treatment of ITP begins with high-dose corticosteroid and/or immunoglobulin therapy. If there is little or no response, splenectomy, chemotherapy, plasmapheresis, and staphylococcal protein A column absorption are additional treatment options (145). Routine prophylactic platelet transfusions are unwarranted, however, even during splenectomy because there is usually little operative bleeding. Moreover, the platelet count often rises after the splenic pedicle is clamped. If the patient bleeds despite clamping of the spleen, platelet transfusion may be of temporary value.

IVIG is effective in rapidly (1 to 5 days) raising platelet counts in many patients with ITP, although the effect lasts only 1 to 2 weeks. The administration of IVIG just before platelet transfusion may increase the magnitude and duration of the platelet count increment in some patients to a substantially greater degree than platelet transfusion alone (146). This combined approach has been employed for the acute treatment of hemorrhage or in preparation for surgery. The intravenous administration of RhIG to Rh-positive patients with ITP also may be of benefit in transiently raising circulating platelet counts (147). The mechanism of action of RhIG is postulated to be a competitive inhibition of uptake of IgG-coated platelets in the reticuloendothelial system by anti-D (IgG)–coated RBCs. RhIG of course is ineffective in Rh(d)-negative patients and splenectomized patients.

Thrombotic Thrombocytopenic Purpura

TTP is a microvascular occlusive syndrome characterized by thrombocytopenia and microangiopathic hemolytic anemia owing

to the presence of platelet microthrombi in small vessels. The microthrombi are composed in large part of platelets and vWF. The thrombocytopenia is apparently due to the consumption of platelets from abnormal platelet aggregation and adhesion to blood vessel walls. End-organ damage can occur because of microvascular occlusion and typically manifests itself as renal failure and neurologic deficits. Myocardial infarction and cerebrovascular accidents are possible. The platelet microthrombi have been shown to result from the platelet-aggregating activity of abnormally large vWF multimers that arise from a deficiency of vWF-cleaving protease activity (148,149,150 and 151). Patients with chronic, relapsing TTP have a chronic deficiency of this protease, whereas patients with a single acute episode have been found in many cases to have transient IgG autoantibodies to the vWF-cleaving protease that disappear after recovery. Mortality in untreated cases approaches 100%. Platelet transfusions are contraindicated in TTP and have been reported to worsen symptoms (152,153). The transfused platelets may promote the formation of additional microthrombi, especially in the cardiac or cerebral circulations, with potential fatal consequences. Antiplatelet therapy with inhibitory agents such as aspirin and dipyridamole to minimize further platelet adhesion and/or aggregation has been frequently administered but is of theoretical but undocumented benefit. Corticosteroids have been administered frequently as well, although their efficacy is also unproven.

Patients should be treated with daily plasma exchange of at least one plasma volume using FFP, SD-plasma, or cryoprecipitate-reduced plasma as replacement fluid during plasmapheresis (154). Plasma exchange may be therapeutic by providing the rapid infusion of large volumes of plasma containing exogenous vWF-cleaving protease. Plasma exchange also may serve to remove abnormally large vWF multimers and/or autoantibodies to vWF-cleaving protease. Whether the plasma removal actually provides additional therapeutic benefit over plasma infusion is yet unproven, however. Plasma exchange should continue until the clinical status of the patient improves and the platelet counts increase to normal. In addition to following platelet counts, disease status may also be assessed by measuring lactate dehydrogenase and indirect bilirubin to determine ongoing hemolysis and creatinine to assess renal status.

The therapeutic dose of plasma has not been firmly established. Therapeutic responses have been observed with doses ranging from as little as a couple of units of FFP per day to once or twice daily plasma exchange of one to two plasma volumes. Although plasma transfusions alone have been shown to be therapeutic in TTP for many patients, larger volume plasma exchange gives overall improved outcomes and lower mortality (155,156). It has been speculated that larger plasma doses may be required in patients with single acute episodes of TTP when high titers of autoantibodies are present. Conversely, smaller doses of plasma may be needed in chronic releasing patients who have a deficiency of vWF-cleaving protease without autoantibodies.

Plasma transfusions, as opposed to plasma exchange, may be administered initially if there will be a significant delay in the initiation of plasmapheresis and plasma exchange. Simple plasma transfusions have also been used to successfully keep chronic, relapsing cases of TTP in remission. Cryoprecipitate-reduced plasma, i.e., the plasma-derived product that remains after cryoprecipitate has been removed from FFP, has been used successfully in treating TTP. It has been used effectively in some patients who were unresponsive to initial treatment with FFP as well as in the initial treatment of TTP (157,158). It has the possible advantage of providing lower amounts of vWF multimers than FFP (157). SD-plasma, which also has been reported to be relatively deficient in high molecular weight multimers of vWF (157) and a reduced risk of transmission of hepatitis and acquired immunodeficiency syndrome (AIDS), also is efficacious in treating TTP (159,160). In patients with TTP who may receive as many as hundreds of units of plasma, SD-plasma's reduced risk of transmitting HIV and hepatitis may be an important consideration. For these reasons, some transfusion specialists advocate SD-plasma as the preferred plasma product for use in TTP. However, despite the reduced level of large molecular weight vWF multimers in cryoprecipitate-reduced plasma and in SD-plasma, actual improved therapeutic efficacy over FFP has not been demonstrated to date. Some patients with autoantibodies to the vWF-cleaving protease activity are relatively refractory to plasma therapy. Additional treatment modalities, such as splenectomy, vincristine, azathioprine, and protein A immunoabsorption have been employed for these refractory patients with variable results.

Hemolytic Uremic Syndrome

Hemolytic uremic syndrome (HUS) has similarities to TTP and has frequently been classified as a variant of TTP. However, unlike in TTP, renal failure predominates over neurologic signs and symptoms in HUS. Also, thrombocytopenia is typically less severe in HUS than in TTP. Unfortunately, HUS is usually less responsive to plasma exchange or plasma transfusion than TTP. Patients with HUS, unlike those with TTP, appear to have normal levels of the vWF-cleaving protease (161). Thus, the etiology and pathophysiology of HUS may prove to be different from those of TTP. More study is necessary to clarify the relationship of these two illnesses. Plasma therapy is typically not employed in childhood HUS because it is of unclear clinical benefit and because childhood HUS is generally self-limited (162,163). Plasma exchange using FFP as the plasma replacement is often reserved for the most severe cases of adult HUS. The efficacy of plasma exchange for adult HUS is not as clearly established as it is in TTP. In fact, the potential for confusion in diagnosis between TTP and HUS makes it difficult to distinguish the relative efficacy of plasma therapy in these two disorders based on published studies. Treatment of HUS often consists of dialysis as needed until recovery. Unfortunately, although most patients with TTP will have a full recovery (although some will relapse), the recovery rate is lower in HUS. Many adult patients with HUS experience chronic and irreversible renal failure, necessitating chronic dialysis or renal transplantation.

Hemolysis, Elevated Liver Enzymes, and Low Platelet Count Syndrome

Hemolysis, elevated liver enzymes, and low platelet count (HELLP) syndrome develops in approximately 4% to 12% of

women with severe preeclampsia or eclampsia at the end of pregnancy (164,165,166,167 and 168). However, HELLP is not always preceded by these other disorders. HELLP is similar to TTP in that microangiopathic hemolysis and thrombocytopenia are prominent findings. However, unlike TTP or HUS, liver abnormality is present as reflected by elevations in liver enzyme tests. It is unclear whether HELLP and TTP are related syndromes. Platelet transfusions have been employed for treating the thrombocytopenia of HELLP, with no reports to date of adverse sequelae, as seen in TTP (164,167). In most women, recovery occurs spontaneously with childbirth or after termination of pregnancy (164,168). In most cases, recovery occurs by 3 days postpartum. The persistence of microangiopathic hemolysis and thrombocytopenia beyond 3 days postpartum may suggest the diagnosis of TTP or a TTP-like syndrome, and the initiation of plasma exchange with FFP should be considered (169). Published outcomes of plasma exchange have been favorable in this setting, except when HELLP syndrome has been complicated by multisystem organ involvement (168). Plasma exchange is not otherwise an accepted therapy for HELLP syndrome and attempts to prolong pregnancies complicated by HELLP syndrome with plasma exchange have had poor outcomes (168).

Posttransfusion Purpura

Posttransfusion purpura (PTP) is an immune thrombocytopenia that can occur approximately 5 to 10 days after transfusion in previously pregnant or transfused women or less commonly in previously transfused men (170,171 and 172). The typical patient is a multiparous woman. Most commonly, the affected individual lacks the HPA-1a (PL^{A1}) platelet antigen, which is present in 98% of individuals, and has been alloimmunized to this platelet antigen by exposure during pregnancy or prior transfusion. Less commonly, platelet antigens other than HPA-1a have been implicated in cases of PTP. After subsequent transfusion, the patient has an anamnestic response owing to reexposure to the implicated platelet antigen that results in the increased synthesis of the previously made antibody. The antibody mediates destruction of transfused antigen-positive platelets as well as, paradoxically, the patient's own HPA-1a–negative platelets. Less commonly, platelet antigens other than HPA-1a have been implicated in cases of PTP.

The mechanism of destruction of the patient's own platelets is not clearly understood. Possible mechanisms include (a) immune complexes adhering to antigen-negative platelets, resulting in their accelerated removal by complement activation or phagocytosis in the reticuloendothelial system, (b) soluble transfused platelet antigens binding to autologous platelets, thereby making them targets for the antibody, and (c) production of autoantibodies or cross-reacting alloantibodies early in the immune response that target autologous platelet determinants. The resulting thrombocytopenia is often severe (10,000 to 20,000 platelets/μL or less). Although this a self-limited condition, the thrombocytopenia may lead to significant and even life-threatening hemorrhage, and the low platelet count may persist for several weeks if untreated.

The therapy of choice is high-dose IVIG (173). Corticosteroids are a possible adjunctive therapy. Plasma exchange has been employed with varying success as well. The transfusion of platelets from HPA-1a–positive or -untested donors is largely ineffective, generally avoided, and may be contraindicated. Theoretically, the infusion of additional antigen may further stimulate the immune response and may also enhance the destruction of the patient's own HPA-1a–negative platelets.

Random platelet transfusions have been associated with severe febrile and anaphylactoid transfusion reactions. Transfused HPA-1a–negative platelets may be rapidly destroyed, as would be predicted based on the destruction of the patient's own HPA-1a–negative platelets (174). However, HPA-1a–negative platelets in conjunction with IVIG have been administered with therapeutic success in severely thrombocytopenic patients (175,176). Owing to the time required to test for HPA-1a antigens and antibodies to confirm the involvement of that antigen and the difficulty in obtaining antigen-negative blood components, the transfusion of blood components, in general, should be avoided. RBCs, if not available from a donor who is negative for the implicated platelet antigen, may be saline–washed or frozen-deglycerolized and washed in an attempt to reduce platelet membrane fragments and soluble antigens. Although frozen–thawed RBCs have been reported to induce a relapse of PTP (177), washed units, nevertheless, have been administered without adverse effect to others (178). PTP has been associated with the presence of HLA B8 and DR3 antigens in the transfusion recipient (179).

Neonatal Alloimmune Thrombocytopenia

In neonatal alloimmune thrombocytopenia (NAIT), a pregnant woman becomes alloimmunized to fetal platelet antigens (179,180 and 181). As with posttransfusion purpura, HPA-1a alloimmunization is the most common cause of NAIT, although other platelet antigens less commonly have been implicated. The mother who lacks the platelet antigen HPA-1a may make antibodies to HPA-1a antigens on the fetal platelets. The antibodies cross the placenta in NAIT and mediate platelet destruction in the neonate, as well as possible suppression of megakaryocyte precursors and reduced platelet production. Thrombocytopenia may be severe and, if untreated, can last for several weeks after delivery. The major complication in severely thrombocytopenic neonates is intracerebral hemorrhage, which occurs most commonly within 24 to 36 hours after delivery (181,182). Prompt treatment with the transfusion of maternal platelets is indicated to prevent life-threatening hemorrhage. Maternal platelets will lack the involved platelet antigen and be compatible regardless of which platelet antigen is implicated. The plasma in the platelet concentrate may be replaced by saline to avoid the infusion of additional platelet antibodies. The platelets should be gamma irradiated to reduce the risk of transfusion-associated graft-versus-host disease (TA-GVHD). Platelets also may be obtained from an HPA-1a–negative donor; however, the antibody involved is not always directed against HPA-1a, and insufficient time may be available to determine the specificity of the antibody before transfusion. Furthermore, the neonate will be placed at additional risk of a new donor exposure. HPA-1a–positive or untested platelets are not contraindicated in NAIT as they are in

PTP (183). An important difference between PTP and NAIT is that the antibody is passively obtained and of limited quantity in NAIT, rather than being actively produced in the transfusion recipient as in PTP. IVIG is another treatment that has approximately a 70% response rate, although its effect is usually delayed 1 to 2 days (184). If maternal or antigen-negative platelets are not available in a timely fashion, a trial of IVIG combined with the transfusion of random platelets may be indicated.

Platelet Refractoriness

Platelets carry HLA class I antigens. Approximately 20% to 50% of patients who receive multiple transfusions become alloimmunized to HLA class I antigens (185). Approximately one third of HLA alloimmunized patients become refractory to platelet transfusion. Refractoriness is reflected by poor posttransfusion increments in platelet count measured within 10 minutes to 4 hours after transfusion, defined variably as a CCI less than 5,000 or 7,500 (Fig. 72.1) (23,24,186,187). Platelet alloimmunization apparently requires the infusion of donor antigen-presenting cells (APCs) such as monocytes, B cells, and dendritic cells, all of which express HLA class II antigens. Alloimmunization appears to occur when donor HLA class I antigens are presented to recipient T cells by donor APCs in the presence of a foreign class II antigen. When donor APCs in platelet concentrates are decreased in number by the use of leukocyte reduction filters (LRFs) or if they are inhibited by treatment with ultraviolet B radiation, the incidence of recipient HLA alloimmunization is reduced by approximately 70% (23,185). Not all patients who are refractory to platelet transfusions are HLA alloimmunized (25,26). In some cases, alloantibodies directed against platelet-specific antigens may be the cause of refractoriness. Patients with autoantibodies to platelet antigens, as in ITP, also will be refractory to platelet transfusion. Nonimmune conditions may be associated with platelet refractoriness, such as sepsis, DIC, splenomegaly, hemorrhage, fever, and the administration of amphotericin B (25,26). Suboptimal posttransfusion platelet increments also are seen with major ABO-mismatched platelets or with platelets near their storage outdate (188).

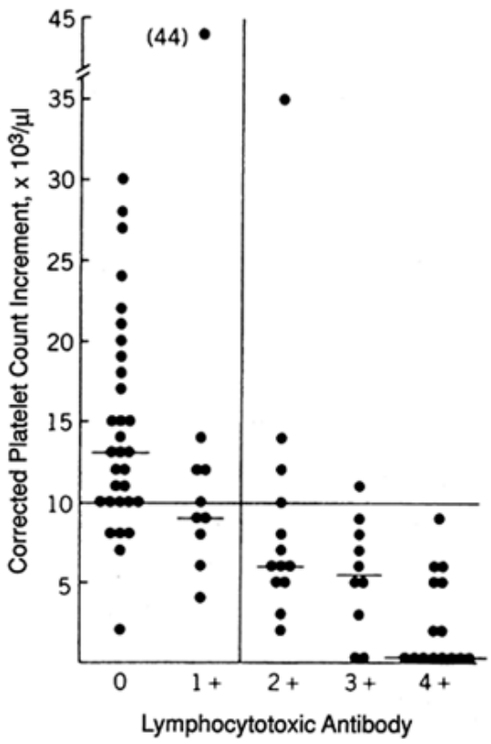


FIGURE 72.1. One-hour corrected count increments (CCI) (see text), after the transfusion of pooled, random platelet concentrates in HLA-alloimmunized thrombocytopenic recipients. The CCI is plotted as a function of the lymphocytotoxic antibody level. Reactions 2+ to 4+ are considered positive for the presence of lymphocytotoxic antibody. Median values are indicated by horizontal bars. (From Daly PA, Schiffer CA, Aisner J, et al. Platelet transfusion therapy; one-hour post-transfusion increments are valuable in predicting the need for HLA-matched preparations. *JAMA* 1980;243:435–438, with permission.)

Several algorithms for addressing platelet refractoriness have been used depending on available resources and locally accepted practices (189). The comparative cost-effectiveness of the various alternatives has not been rigorously studied. Thus, no clear consensus regarding the superiority of one approach compared with another has emerged. The different approaches generally consist of one or a combination of the following strategies: (a) provision of the closest HLA-matched single-donor (apheresis) platelets available (190), (b) provision of HLA-typed single-donor platelets that lack the HLA antigens to which the recipient is alloimmunized (191), and (c) provision of cross match-compatible single-donor platelets (192,193). Each approach has advantages and disadvantages. The initial step employed at many centers is to test and verify the presence of HLA or platelet-specific antibodies in the patient's serum before proceeding with any of these approaches.

To provide HLA-matched platelets, the HLA class I antigenic phenotype of the recipient lymphocytes must first be determined. This typing will delay the initial provision of HLA-matched platelets. An additional delay is incurred if it is necessary to call in donor(s) for apheresis. Table 72.5 lists the grading scale for degree of match of HLA class I antigens between the donor and the recipient. The significant antigens are encoded by two chromosomal loci, designated HLA A and B. Each individual has two alleles at each locus and can synthesize two A and two B antigens. An identical four-antigen match is called an A match. Donor platelets that are a B match have one or two antigens that are either related, i.e., cross-reactive, or unknown. A and B in this context do not refer to the HLA A and B loci. Detection of only one antigen at a locus indicates either that the alleles are homozygous or that the other antigen was untypable. B1U (one antigen unknown) and B2U (two antigens unknown) are potentially, but not necessarily, identical to A matches because homozygosity could explain the unknown antigen. A C match means that one of the four antigens is nonidentical and noncross-reactive with any of the recipient's antigens. In common usage, the term HLA-matched usually implies that the donor and the recipient are a C match or better or that the donor lacks the HLA antigen(s) to which the recipient is sensitized. Rarely are identical A matches available, and most HLA-matched platelets are C matches with the existing donor bases of most collecting facilities. Unfortunately, if the one mismatched HLA antigen in the C match corresponds to the antigenic specificity of the patient's HLA alloantibody, then the C match will not give improved platelet increments over randomly chosen

platelets. Clearly, this approach works best if A or B matches are available.

TABLE 72.5. HLA MATCHING NOMENCLATURE[a]

Match Grade[b,c]	Description
A	All four donor antigens are identical to those of recipient.
B1U	Three donor antigens are identical to recipient antigens; fourth donor antigen is unknown.
B1X	Three donor antigens are identical to recipient antigens; fourth donor and recipient antigens are cross-reactive.
B2U	Two donor antigens are identical to recipient antigens; third and fourth donor antigens are unknown.
B2UX	Two donor antigens are identical to recipient antigens; third donor antigen is cross-reactive with a recipient antigen; fourth donor antigen is unknown.
B2X	Two donor antigens are identical to recipient antigens; third and fourth donor antigens are both cross-reactive with recipient antigens
C	Three donor antigens are identical to those of recipient; fourth donor and recipient antigens are nonidentical and noncross-reactive
D	Two donor and recipient antigens are identical; third and fourth donor antigens are nonidentical and noncross-reactive with recipient antigens.

[a]HLA matching refers to comparisons of the four HLA class I antigens coded by the HLA A and B loci of the transfusion donor and recipient.

[b]Match grade nomenclature (i.e., A, B, C, D) is unrelated to the similarly named HLA chromosomal loci.

[c]An unknown antigen (U) indicates either homozygosity at a locus or the presence of an antigen that is nonreactive with the typing antisera. An X denotes immunologic cross-reactivity of donor and recipient antigens. A number appearing in the match grade indicates how many donor antigens are unknown or cross-reactive.

Adapted from Duquesnoy RJ, Filip DJ, Rodey GE, et al. Successful transfusion of platelets “mismatched” for HLA antigens to alloimmunized thrombocytopenic patients. *Am J Hematol* 1977;2:219–226.

To overcome the problem of incompatible C matches when using HLA-matched platelets, the antigenic specificity or specificities of the HLA alloantibodies can be determined. This can be done in a lymphocytotoxicity assay against a panel of lymphocytes representing different HLA class I phenotypes. When the specificity of the HLA antibody is determined, HLA-typed single donor platelets can be chosen that lack the antigen(s) to which the patient is alloimmunized. This approach can be combined with choosing donor platelets that are also the best HLA match available in an attempt to reduce the stimulus for further alloimmunization.

In an alternative approach, the patient's serum can be tested for platelet-reactive antibodies using, for example, a solid-phase RBC agglutination assay (193). Such an assay has the theoretical advantage of detecting both HLA- and platelet-specific antibodies, regardless of whether they are complement fixing. Platelet-specific alloantibodies mediate platelet refractoriness in a small number of chronically transfused patients and are missed by assays specific for HLA antibodies. If a panel of platelets from apheresis donors already is used in the initial platelet antibody-screening assay, there is a chance that compatible platelet donors may be identified in the process of testing for the presence of platelet alloantibody. If not, additional donors may be screened using the same assay, in what is referred to as a platelet cross match. Platelet cross matching also has the advantage of not necessarily requiring HLA typing. However, if the majority of donors are incompatible by platelet cross matching, then a combination of HLA-matching and platelet cross matching may be attempted. Platelets from the closest HLA-matched donors may be preselected for use in platelet cross matching. A possible algorithm for the use of platelet cross matching in refractory patients is shown in Fig. 72.2.

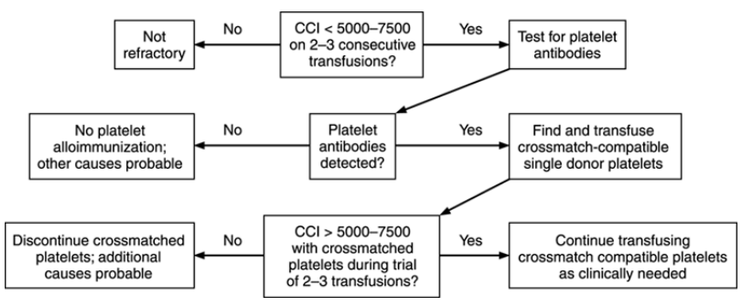


FIGURE 72.2. Algorithm for platelet cross-matching in platelet refractoriness.

Regardless of the approach taken to select compatible platelets for transfusion, the CCI of each transfusion should be measured and compared with earlier transfusions of randomly chosen platelets. Significant improvements in the CCI justify continued use of HLA-selected platelets or cross match–compatible platelets. If no improvement in platelet count increment is seen with selected platelets after a trial of two to three transfusions, this may indicate that the refractoriness is being caused by clinical factors other than HLA or platelet alloimmunization. The continued use of HLA-selected or cross match–compatible platelets might not be justified. If, however, HLA or platelet alloimmunization is likely or has been confirmed by testing, a better HLA match or platelet cross matching, if not available locally, may be sought by your blood supplier from another blood donor center.

Uremia

Patients with severe renal failure have a hemostatic defect owing in large part to platelet dysfunction, but whose etiology is likely multifactorial (194,195,196 and 197). Uremic patients have prolonged bleeding times and a variety of measured defects in platelet function, platelet–vessel wall interactions, and in vascular endothelial and smooth muscle cell metabolism. In addition, the decrease in circulating RBCs resulting from the anemia of renal failure apparently causes rheologic changes that decrease the concentration of platelets closest to the blood vessel wall where they are needed for primary hemostasis. RBCs tend to flow in the midstream, displacing platelets to the periphery of the bloodstream near the blood vessel wall. One possible explanation for platelet dysfunction is the accumulation of metabolites toxic to platelets in the blood of patients with renal failure (197).

Platelet transfusions are not considered an effective therapy in uremia because transfused platelets are expected to become dysfunctional when exposed to the same inhibitory milieu as the patient's own platelets. However, several other treatment options are available (Table 72.6). The nontransfusion treatment of choice is hemo- or peritoneal dialysis (198). Either procedure can temporarily correct the hemostatic defect, perhaps by transiently removing the toxic metabolites that cause the uremic thrombocytopathy. Bleeding times can also be shortened in uremic patients by maintaining their hematocrits 28% to 30% or more either through the use of recombinant human erythropoietin or RBC transfusions (199,200,201 and 202). Desmopressin, a synthetic analog of antidiuretic hormone that stimulates increases in endogenous vWF and factor VIII levels, also shortens the bleeding time in many uremic patients (203,204). Cryoprecipitate transfusions, at a dose of one bag per 10 kg of patient body weight, shorten bleeding times and control bleeding in some patients with uremic bleeding (205,206). The element of cryoprecipitate responsible for this effect is unknown, but vWF and factor VIII are possible candidates. The intravenous infusion of conjugated estrogens has also been shown to shorten bleeding times prolonged by uremia (207,208). The effect of estrogen is delayed compared with that of desmopressin and cryoprecipitate, but its duration of effect is as long as approximately 2 weeks, longer than that of desmopressin or cryoprecipitate. Oral and transdermal administrations of estrogen also shorten bleeding times in uremia, although the onset of action is delayed further (209,210). All these treatments are transient and not curative and also may not be effective for all patients. In some eligible patients with uremic thrombocytopathy, only renal transplantation offers hope for a cure.

TABLE 72.6. THERAPEUTIC OPTIONS IN UREMIA

Treatment	Dose	Mechanism	Time to Peak Effect	Duration
Recombinant human erythropoietin	See product insert	Increases hematocrit to ≥30%, which may increase platelet concentration at vessel wall	Weeks	Sustained if hematocrit is maintained
Desmopressin	0.3 mg/kg once or twice daily	Stimulates release of endogenous vWF and factor VIII	1–2 hr	˜6–8 hr
Red blood cells	Transfuse to ≥30% hematocrit	May increase platelet concentration at vessel wall	Immediate	Sustained if hematocrit is maintained
Cryoprecipitate	1–2 U/10 kg	Mechanism unknown (supplies vWF and factor VIII)	1–2 hr	24–36 hr
Conjugated estrogens	0.6 mg/kg/i.v. daily × 5 days	Mechanism unknown	5–7 d	˜14 d

vWF, von Willebrand factor.

Reversal of Aspirin Effects

Aspirin inhibits platelets by irreversibly acetylating and inhibiting platelet cyclooxygenase-1, which is a key enzyme in the pathway leading to the production of thromboxane A_2, a platelet-aggregating factor (211). The result is a partial inhibition of platelet activation, a prolongation of the bleeding time, and an increased tendency to bleed (212). Platelets remain inhibited for their entire lifetime, and a prolongation of the bleeding time may persist for as long as 4 to 7 days. Aspirin has a plasma elimination half-life of approximately 15 to 20 minutes and is effectively cleared from plasma after approximately 2 hours. Deacetylated metabolites persist longer in the circulation but do not inhibit cyclooxygenase-1. When aspirin is cleared, the aspirin effect is reversible by platelet transfusion.

Of note, not all patients on aspirin who undergo an invasive surgical procedure, such as CABG, will have excessive bleeding compared with patients not on aspirin (213). Since the thromboxane A_2 pathway is only one of several that mediate platelet activation, most patients have only a modest (i.e., 1 to 3 minutes) prolongation of bleeding times (212). Only patients who experience a greatly prolonged bleeding time that may put them at greater risk of surgical bleeding and who need to urgently undergo a major surgical procedure should be considered candidates for platelet transfusion. Even in that setting, platelet transfusion in most cases should be reserved for patients who actually demonstrate excessive intra- or postoperative bleeding. Desmopressin is effective in shortening an aspirin-prolonged bleeding time (214) and is a possible pharmacologic alternative to transfusion for treating or preventing bleeding attributable to aspirin effects.

Coagulation Protein Deficiencies and Defects

Hepatic Failure

Because many coagulation factors are synthesized in the liver, patients with hepatic failure develop multiple coagulation factor deficiencies (215,216 and 217), as reflected in prolongations of the PT and aPTT. Patients with liver failure are also at increased risk of developing DIC (218) owing in part to decreased synthesis of antithrombin, α_2-plasmin inhibitor, protein C, protein S, and other regulatory factors of the coagulation cascade. Vitamin K_1 administration is indicated for patients with liver failure and prolonged coagulation tests to correct a possible vitamin K deficiency. This in itself may improve coagulation factor levels and hemostasis in some patients. FFP transfusion is indicated in a patient with liver failure with PT and/or aPTT elevated 1.5-fold or more over control values in the setting of bleeding or in preparation for surgery. However, prophylactic plasma transfusions are not recommended. They are unnecessary in milder hepatic failure and futile in severe cases, except possibly on a short-term basis if liver transplantation is planned. The short-lived factor VII, with a circulating half-life of only 2 to 5 hours, is rate limiting, necessitating frequent, large-volume transfusions or plasma exchange. Platelet, RBC, and cryoprecipitate transfusions should be administered as needed according to accepted guidelines (see above).

Desmopressin is a possible adjunct to transfusion therapy for the treatment of bleeding episodes in cirrhotic patients who have prolonged bleeding times. Desmopressin has been shown to decrease the bleeding time in such patients even though their vWF and factor Vlll levels are normal or elevated (214,219). However, because a controlled trial failed to demonstrate a benefit of using desmopressin for the treatment of acute variceal hemorrhage, the clinical effect of desmopressin appears to be insufficient for large hemorrhage (220).

Disseminated Intravascular Coagulation

DIC is a state of uncontrolled thrombin generation that can be complicated by both thrombosis and bleeding (221). The thrombotic phase of DIC results from excessive activation of thrombin and the coagulation cascade. The release of procoagulant tissue factor and thromboplastins from damaged tissues is one possible cause of DIC. The hemorrhagic complications of DIC result from the consumption and depletion of clotting factors and platelets. The two main treatment modalities of DIC are supportive blood component therapy and the interruption of the coagulation cascade with antithrombin and/or heparin. The demonstration of improved outcomes from treatment have been problematic because outcomes are generally poor and highly dependent on the underlying, predisposing factors for DIC. Nevertheless, it is common clinical practice to transfuse blood components for treatment or prophylaxis of bleeding. The administration of heparin to treat thrombotic complications is more controversial. Antithrombin concentrate, also administered to decrease thrombosis, has given some promising results.

Some experts argue that providing plasma and coagulation factors only serves to “add fuel to the fire” of DIC, but there exists no clear evidence that this happens. Therefore, if the patient is bleeding seriously and the PT and/or PTT are prolonged by 1.5 to 1.8 compared with normal controls, it is common practice to replace coagulation factors with plasma transfusion. Transfused plasma also contains natural anticoagulants, such as antithrombin, protein C, and protein S, which theoretically may be of benefit in controlling thrombosis. In DIC, there is usually also some degree of thrombocytopenia, with reported average counts of 52,000 to 60,000 platelets/μL in some studies (222,223). In addition, platelet function may be significantly inhibited by circulating fibrin degradation products (224). If a patient is actively bleeding or about to undergo an invasive procedure, platelet transfusion is indicated for the restoration of the patient's platelet count to more than 50,000 /μl (222,225,226 and 227). However, because transfused platelets may undergo rapid consumption (228) and platelet count alone may be a misleading indicator of functional platelet activity owing to the concurrent thrombocytopathy, the end point of platelet transfusion in many situations may need to be more pragmatic, i.e., the control or cessation of bleeding.

Hypofibrinogenemia is common in DIC. Although some experts oppose transfusion for fibrinogen replacement on the basis of “fueling the fire” (224), others contend that restoration of the fibrinogen level to more than 100 mg/dL is desirable to prevent or treat serious bleeding (228,229). Fibrinogen repletion by cryoprecipitate transfusion may be used as a supplement to plasma. However, cryoprecipitate should not be administered alone in the treatment of DIC because it lacks significant concentrations of most of the coagulation factors, with the exception of factors Vlll, Xlll and vWF. The dose of cryoprecipitate can be calculated as described earlier in this chapter. However, because fibrinogen consumption can be greatly increased in DIC, the expected increments may not be achieved.

Interruption of the coagulation cascade can be attempted using heparin and/or antithrombin, a necessary cofactor for heparin and a powerful inhibitor of thrombin. antithrombin levels are typically low in DIC, and replacement may be indicated to help control excessive coagulation without the hemorrhagic risk of heparin. In addition, antithrombin may be administered to facilitate the action of exogenously administered heparin (230). antithrombin is found in plasma and may be supplied in part by plasma transfusion. However, adequate doses of plasma for antithrombin repletion could lead to volume overload in many cases (231). For that reason, a commercial, virus-attenuated antithrombin concentrate is the preferred treatment in DIC if antithrombin repletion is attempted (232).

Reversal of Warfarin Effects

Patients treated with coumarin anticoagulants, such as warfarin, may need to have the anticoagulant effect reversed urgently for surgery or serious bleeding or because of an abnormally high international normalized ratio (INR) that puts the patient at increased risk of intracranial hemorrhage (Table 72.7). When time permits, as with elective surgery, the anticoagulant effect of warfarin can be reversed simply by stopping the drug (233). Vitamin K_1, administered subcutaneously, parenterally, or orally, in combination with holding further warfarin doses, may also be used to reverse the warfarin effect more rapidly when greater urgency exists (233,234). More rapid (within 6 to 8 hours) effects of vitamin K_1 are seen with intravenous administration, as opposed

to the other routes of administration (234). Rare anaphylactic reactions have been reported with intravenous infusion; therefore, slow administration is recommended (1 mg per hour) (234). When 1 or more days are available for correction of the elevated INR, low dose (1.0 to 2.5 mg) oral vitamin K_1 is effective in decreasing the INR, while avoiding relative warfarin resistance that higher doses induce (235). Low-dose vitamin K_1 permits faster reestablishment of the warfarin effect, which is clinically important for many patients.

TABLE 72.7. MANAGING PATIENTS WITH HIGH INR VALUES

Clinical Setting	Recommendations
INR higher than therapeutic range but <5.0; no clinically significant bleeding; no urgent surgical intervention	Lower the dose or omit the next dose; resume warfarin at a lower dose when the INR approaches desired range If the INR is only slightly higher than therapeutic range, dose reduction may not be necessary
INR >5.0 but <9.0; no clinically significant bleeding	Patients with no additional risk factors for bleeding: omit the next dose or two of warfarin, monitor INR more frequently, and resume warfarin at a lower dose when the INR is in therapeutic range Patients at increased risk of bleeding: omit the next dose of warfarin, and give vitamin K_1 (1.0–2.5 mg orally) Patients requiring more rapid reversal before urgent surgery or dental extraction: vitamin K_1 (2–4 mg orally); if the INR remains high at 24 hr: an additional dose of 1–2 mg
INR >9.0; no clinically significant bleeding	Hold warfarin; give Vitamin K_1 (3–5 mg orally); if the INR remains high at 24 hr, vitamin K_1 dose can be repeated; resume warfarin when INR is therapeutic
Serious bleeding or major warfarin overdose (e.g., INR >20.0) requiring very rapid reversal of anticoagulant effect	Hold warfarin; give Vitamin K_1 (10 mg by slow i.v. infusion), with FFP transfusion or prothrombin complex concentrate, depending on urgency and availability; vitamin K_1 injections may be repeated q12h
Life-threatening bleeding	Prothrombin complex concentrate (if available) or FFP, with vitamin K_1 (10 mg by slow i.v. infusion); repeat if necessary, depending on the INR
Resumption of warfarin therapy after high dose vitamin K_1	Heparin, until the effects of vitamin K_1 have been reversed, and patient is responsive to warfarin

INR, international normalized ratio; FFP, fresh frozen plasma
Adapted from Ansell J, Hirsh J, Dalen J, et al. Managing oral anticoagulant therapy. *Chest* 2001;119:225–238. (from the sixth American College of Chest Physicians Consensus Conference on Antithrombatic Therapy, February 2000.)

When even the delay for vitamin K_1 to take effect puts the patient at risk of significant hemorrhage, plasma or prothrombin complex concentrate administration may be indicated (236,237). A dose of 6 to 7 mL/kg of FFP may be adequate for some patients for reversing the warfarin effect. The effect of the FFP should be verified with a posttransfusion PT and INR measurement. Additional plasma may be infused as needed to shorten the PT to approximately 1.5 times the PT control or less or to decrease the INR to the desired range. Larger doses of FFP, i.e., 12 to 15 mL/kg, may be necessary, particularly when the need arises for an immediate and more complete warfarin reversal. Prothrombin complex concentrates, which contain the vitamin K–dependent coagulation factors, can be used to reverse the warfarin effects in patients who need more urgent correction or who are at risk of hypervolemic reactions from plasma infusion (236,237). Prothrombin complex concentrates, presumably because of the higher dose possible of vitamin K–dependent coagulation factors, have been shown to shorten the PT more rapidly than plasma and may be used for the treatment of life-threatening bleeding, such as an intracerebral hemorrhage (236,237 and 238). However, the use of prothrombin complex concentrates has been associated with an increased risk of thromboembolic events and their availability may be limited. Consensus guidelines that incorporate the options for managing patients with high INR values have been published by the American College of Chest Physicians (Table 72.7) (238).

Heparins/Heparinoid/Ancrod

Plasma should not be transfused to reverse the anticoagulant effect of heparin. Heparin acts by binding to and activating antithrombin, which in turn inhibits thrombin as well as factors Xa and XIIa (239). Plasma transfusion is a source of supplemental antithrombin and has been used to treat heparin resistance owing to antithrombin deficiency, when commercial antithrombin concentrate is not available (240). Therefore, if anything, plasma transfusion potentiates the action of heparin rather than reverses it. Heparin, which has a relatively short circulating half-life (approximately 90 minutes), can be reversed by discontinuing its administration (239). Heparin can also be reversed by administering protamine, a polycationic protein that binds to and inactivates polyanionic heparin molecules (241). Protamine is commonly used in neutralizing heparin's effect at the end of open-heart surgery (127). Other methods are currently being investigated and include the administration of the enzyme heparinase, which breaks down heparin, and recombinant platelet factor 4, a cytokine that inactivates heparin (127,242). Just as with heparin, the activities of low molecular weight heparins and heparinoid are not reversed by plasma transfusion. They are, however, not as susceptible to neutralization by protamine as unfractionated heparin and have a longer half-life (243). Desmopressin is reported to shorten the aPTT prolonged by heparin and may be of use in mitigating the anticoagulant effects of these agents (244).

The defibrinogenating agent ancrod derived from the Malayan pit viper, has been used as a heparin substitute, particularly in patients with heparin-induced thrombocytopenia (245,246). Ancrod cleaves fibrinogen to water-soluble fibrin. Ancrod is infused with the goal of achieving anticoagulation by decreasing fibrinogen levels. The reversal of ancrod's effects and the treatment of ancrod overdoses is achieved with cryoprecipitate

transfusion with the goal of increasing fibrinogen to a hemostatic level (100 mg/dL or more) (245,246).

Isolated Factor Deficiencies

Plasma transfusions are indicated for the replacement of documented isolated deficiencies of factors for which no concentrates are available for clinical use, e.g., factor II, V, VII, X, or XI. Specific commercial concentrates, when they are available, are the treatment of choice for factors such as VIII and IX and antithrombin. The clotting factor deficiency should be documented by an abnormal result on the appropriate specific factor assay.

Prothrombin complex concentrates, when available, also may be used for deficiencies of factors II, VII, and X. Caution must be exercised, however, owing to reported thrombotic side effects of prothrombin complex concentrates. The need for replacement therapy varies with the particular factor and the extent of the deficiency. In some cases, the clinical manifestations of a factor deficiency vary widely between patients. For example, in factor XI deficiency, the apparent extent of the deficiency does not necessarily correlate with bleeding (247). In some factor deficiencies, the risk of serious, spontaneous hemorrhage is high, necessitating ongoing prophylactic therapy. In others, increased bleeding may occur only after trauma or surgical procedures, in which case, therapy is reserved only for those settings.

Individuals with a congenital deficiency of factor XIII are at high risk of spontaneous intracranial hemorrhage, and ongoing prophylactic therapy is therefore advisable (248). Factor XIII is a transketolase; it cross-links fibrin monomers and stabilizes fibrin clot formation (249). Factor XIII may be provided with plasma or cryoprecipitate transfusions. The hemostatic levels of factor XIII for bleeding patients are low (approximately 2 to 3 U/dL or approximately 2% to 3% of normal), and the half-life of factor XIII is long (approximately 7 to 12 days). A therapeutic dose for an average adult is four to six units of cryoprecipitate every 3 weeks (250,251). Higher levels are recommended for major surgery or trauma. Factor XIII is concentrated 1.5- to fourfold in cryoprecipitate compared with plasma. For this reason, cryoprecipitate may be preferred over plasma owing to the smaller volume needed. Therapeutic recommendations for other specific factor deficiencies are reviewed elsewhere (248).

von Willebrand's Disease

vWD is the most common inherited bleeding disorder. It is characterized by quantitative and/or qualitative defects in vWF (252,253,254 and 255). vWF is synthesized by endothelial cells, where it is stored in Weibel-Palade bodies. Megakaryocytes also produce vWF, which can be found in platelet α granules. vWF is secreted and circulates in the blood as multimers; the large multimer forms of vWF are required for full activity. vWF mediates adhesion between platelets and damaged blood vessel walls by binding to platelet GPIb and collagen of the exposed vascular subendothelium. It also binds to, stabilizes, and serves as a carrier for factor VIII.

Several treatment options exist for increasing the levels of vWF in patients with vWD who are bleeding or about to undergo an invasive procedure (252,254,256) (Table 72.8). The pharmacologic agent desmopressin is the preferred treatment of mild-to-moderate forms of vWD because it poses no infectious risk (204). Desmopressin stimulates the release of vWF from the patient's own stores of vWF in endothelial cells. Desmopressin can be administered parentally, subcutaneously, or intranasally. Approximately 20% to 25% of patients with vWD do not respond to desmopressin (254). For this reason, responsiveness should be documented with a test dose before an anticipated invasive procedure. Responsiveness can be measured with vWF or factor VIII assays or with a bleeding time. It is useful to determine the vWD type and subtype, if possible, before considering desmopressin use. Responsiveness to desmopressin is expected in type 1, variable in type 2, and unlikely in type 3 (252,254,256). Most experts consider desmopressin to be contraindicated in subtype 2B. The desmopressin-stimulated release of the mutated vWF in subtype 2B vWD can exacerbate the abnormal platelet aggregation and thrombocytopenia characteristic of this subtype and cause thrombotic complications in some patients. Nevertheless some patients with subtype 2B have been reported to respond

TABLE 72.8. THERAPEUTIC OPTIONS IN VON WILLEBRAND'S DISEASE

Treatment	Dose	Mechanism	Infectious Risk	Use	Comment
Desmopressin	0.2–2.4 μg/kg i.v. over 15–30 min.	Stimulates endogenous vWF release	No	Type I vWD; variable efficacy in type II vWD: contraindicated in type IIb; likely ineffective in type III	Not all patients respond; tachyphylaxis reported
Factor VIII/vWF concentrate (humate P)	See product insert; depends on clinical setting	Provides large vWF multimers	Reduced	All vWD types	One FDA-licensed concentrate now available; other factor VIII concentrates may not be therapeutic
Cryoprecipitate	1–2 U/10 kg	Provides large vWF multimers	Yes	All vWD types	Also provides fibrinogen, factor VIII, factor XIII, fibronectin
Platelet transfusion	1 U/10 kg	Provides vWF multimers from platelet α granules	Yes	Type III vWD	Possible adjunct if inadequate response to factor VIII/vWF concentrate

vWF, von Willebrand factor; FDA, U.S. Food and Drug Administration.

therapeutically to desmopressin without adverse effect (257).

Tachyphylaxis has been reported with repeated doses of desmopressin given at intervals of 12 to 24 hours or less. This effect varies from patient to patient and can be lessened with once-a-day dosing (204,254). The most common side effects of desmopressin are facial flushing, headache, mild tachycardia, and a drop in systolic blood pressure (252,256). These are usually mild and transient. The antidiuretic action of desmopressin may result in hyponatremia and water intoxication in some patients, especially infants, small children, and patients receiving repeated doses. For these patients, it is important to monitor serum sodium levels and to restrict their water intake. Desmopressin probably should be used with caution in elderly patients and those with atherosclerotic disease because of a few anecdotal reports of myocardial infarction and stroke associated with desmopressin administration (252,256). However, it should be noted that an attempt at a larger statistical analysis failed to show an increased risk of arterial thrombotic complications with desmopressin use (258).

In clinical settings in which desmopressin is ineffective or contraindicated, patients need to be given exogenous replacement of vWF. The product of choice for vWF replacement is a commercial, heat-treated vWF/factor VIII concentrate that has been FDA licensed for use in vWD (Humate P, Centeon Pharma GmbH) (259). This vWF concentrate has a sufficiently high amount of large molecular weight vWF multimers to be therapeutic in vWD. Because this concentrate is heat treated, it has a reduced risk of viral transmission. This concentrate may be used for all forms of vWD.

Before the availability of the vWF/factor Vlll concentrate, cryoprecipitate in a dose of one unit per 10 kg of the patient's body mass every 12 to 24 hours was the standard vWF replacement therapy (260). It remains an option should the concentrate not be available. One unit of cryoprecipitate contains in a volume of 10 to 15 mL approximately 40% to 70% of the vWF present in the original unit of whole blood from which it was prepared. Thus, vWF in cryoprecipitate is concentrated approximately 20-fold compared with plasma. The disadvantage of cryoprecipitate as a source of vWF is its potential infectious risk. For this reason, the commercial vWF/factor Vlll concentrate is preferred over cryoprecipitate. Because of the presence of vWF in platelet α granules, platelet transfusion has been reported to shorten the bleeding times of patients with type 3 vWD (261,262). vWF from platelets is presumably released locally at sites of vascular injury when platelets undergo the release reaction. Because of the infectious risk of platelet transfusion and the unclear magnitude of its effect, platelets should not be considered a front-line treatment of vWD.

Acquired vWD is an immune disorder caused by autoantibodies to vWF (263). Because of the anti-vWF antibodies, the administration of vWF from exogenous sources or the stimulation of endogenous release with desmopressin may be ineffective or of transient benefit. Some therapeutic success in acquired vWD has been reported with high-dose intravenous immune globulin (264).

Hematopoietic Progenitor Cell Transplantation

Hematopoietic progenitor cell (HPC) transplantation, which involves ablation and reconstitution of the bone marrow with either autologous or allogeneic HPCs, is a transfusion-dependent procedure. The most commonly used sources of HPCs are bone marrow and peripheral blood. Umbilical cord blood is also being investigated as a potential source. Special transfusion needs exist both pre- and posttransplantation regardless of the HPC source. The three major considerations in HPC transplantation are (a) the avoidance of transfusion-transmitted CMV, particularly in CMV-seronegative recipients, (b) the prevention of HLA alloimmunization to avoid platelet refractoriness, and (c) the prevention of TA-GVHD while the patient is immunosuppressed.

Pretransplantation

The avoidance of CMV infection pretransplantation is necessary because reactivation of latent CMV infections posttransplantation is a cause of significant morbidity in HPC transplantation. Common manifestations of CMV infections after HPC transplantation include pneumonia, enteritis, hepatitis, and retinitis (265). Most serious posttransplantation infections occur as a result of reactivation of latent CMV in previously infected patients. If the patient is CMV seronegative pretransplantation, definite precautions should be taken to avoid transfusion-transmitted CMV regardless of the HPC donor's CMV serostatus (266). This is accomplished by transfusion of only CMV-reduced-risk blood components. CMV-reduced risk can be achieved by transfusing RBCs and platelets that are leukoreduced by filtration or that are obtained from CMV-seronegative donors (267,268 and 269). CMV is a leukocyte-associated virus that is transmitted via the passenger leukocytes in cellular blood components. Therefore, the use of leukoreduced blood components, defined as those containing less than 5×10^6 leukocytes per unit (40), significantly reduces the risk of transfusion-transmitted CMV (267). A large, randomized, prospective trial in the setting of allogeneic bone marrow transplantation comparing leukoreduced to CMV-seronegative blood components showed that the two approaches are similarly effective in significantly reducing, although not eliminating, CMV disease (267). FFP has not been reported to transmit CMV, perhaps because its leukocyte content is so low and possibly also because of the freeze-thaw cycle to which it is subjected. For this reason, plasma is not screened for CMV antibodies or further leukoreduced. For transplant candidates who already have a latent CMV infection, there may be a theoretical advantage of reducing the risk of infection with a second strain of CMV. Increasingly, the trend is to transfuse CMV-reduced-risk cellular blood components achieved through filter leukoreduction to all transplant candidates, regardless of CMV serostatus.

The pretransplant patient also should receive blood components that reduce the risk of HLA alloimmunization. Virtually all transplant regimens result in severe thrombocytopenia, requiring platelet transfusion support. If the patient is HLA alloimmunized, this may cause refractoriness to platelet transfusion and increase the risk of serious bleeding (185). In addition, HLA and leukocyte alloimmunization play a role in the etiology of febrile, non-hemolytic transfusion reactions. Avoiding unnecessary transfusion reactions spares the patient considerable discomfort and avoids the time and expense required in their investigation. HLA alloimmunization can be reduced, but not eliminated, by

transfusing RBCs and platelets that have been leukoreduced by filtration to less than 5×10^6 leukocytes per unit (23,185,187). Using leukoreduced, single-donor apheresis platelets appears to provide no additional benefit in preventing HLA alloimmunization when compared with leukoreduced, pooled-platelet concentrates (23). Ultraviolet B irradiation of platelet preparations, like leukoreduction, also appears to be effective in reducing HLA alloimmunization (23).
The routine gamma irradiation of all blood components to avoid TA-GVHD is not necessary before initiation of the myeloablation in allogeneic transplants if the patient is otherwise immunocompetent. However, in peripheral blood stem cell transplantation in which autologous HPCs are collected from the peripheral blood, there is a remote possibility that passenger lymphocytes in cellular blood components transfused precollection could be subsequently collected and cryopreserved during harvesting of the peripheral blood stem cells. These lymphocytes could theoretically mediate graft-versus-host disease on reinfusion with the stem cell product when the patient is severely immunosuppressed. Therefore, it is prudent to irradiate all cellular blood components pretransplantation before autologous peripheral blood stem cell transplantation (270). When the conditioning regimen begins, at which time the transplant recipient's immune system is suppressed, it is important to transfuse gamma-irradiated cellular blood components for all HPC transplants (271,272). Cellular blood components obtained from HLA-matched or genetically-related donors should be irradiated before transfusion regardless of the stage of the transplant regimen or immunocompetency of the recipient. To streamline administrative tracking of patients' transfusion needs and thereby reduce the risk of error, some transfusion services will irradiate all cellular components for pretransplantation transfusions even before myeloablation.

Because the availability of ABO compatible donors may be limited and HLA matching is of paramount importance in HPC transplantation, ABO-incompatible HPC infusion is sometimes unavoidable (273). The harvested HPCs, whatever their source, are accompanied by large numbers of RBCs. Therefore, the infusion of HPCs across ABO types requires special attention. The most important concern is massive hemolysis of the ABO-incompatible RBCs in the infused HPC product by the recipient's naturally-occurring ABO antibodies (274). Techniques that remove incompatible RBCs from the HPC source include simple buffy coat preparation, automated cell separating devices, and various forms of gradient centrifugation in sugar, starch, or gel media. Alternatively, the incompatible ABO antibody in the recipient's plasma can be reduced by serial plasmapheresis, followed by an intentional transfusion of ABO-incompatible RBCs to adsorb any remaining anti-A or anti-B. Severe hemolytic transfusion reactions may occur during this latter process. The severity of the reactions may be reduced by premedicating the patient and transfusing the incompatible RBCs slowly in small aliquots in a carefully controlled setting. This technique has largely been replaced by RBC removal methods.

Posttransplantation

After myeloablation and before engraftment, the patient is severely immunosuppressed. CMV-reduced-risk blood components must be provided, especially to CMV-negative recipients because of the severe adverse consequences of CMV infections (see pretransplantation discussion). Despite the immunosuppression, patients receiving multiple doses of platelets are capable of becoming refractory to platelet transfusion because of HLA alloimmunization; therefore, they should continue to receive leukocyte-reduced blood products. All blood components posttransplantation must be irradiated to avoid TA-GVHD (271,272). Because engraftment does not necessarily correspond with reconstitution of normal immunity and because it is unclear when the patient is no longer at risk for TA-GVHD, it is both prudent and practical to continue irradiating cellular blood components for the rest of these patients' lives. Such a policy should not be burdensome to the blood bank because a successful transplant should eliminate the need for future transfusions.

Before successful engraftment, the patient becomes profoundly thrombocytopenic, granulocytopenic, and, to a lesser degree, anemic. The times to successful engraftment for granulocytes, defined as the time to greater than 500 cells/μL for 2 consecutive days, vary from a low of approximately 8 to 10 days in autologous peripheral blood stem cell transplants to approximately 21 to 28 days in HLA-matched, unrelated transplants. Platelet production takes slightly longer to recover; therefore, platelets are the most frequently transfused product in the HPC transplant patient (275). Granulocyte transfusions are rarely necessary with currently available antibiotics to cover infections and granulocyte colony-stimulating factors to promote more rapid engraftment. The need for RBC transfusion compared with platelets is less, in part because RBCs turn over more slowly and also because RBC precursors are the first to return in the regenerating marrow. Nevertheless, because these patients may have some increased blood loss due to their severe thrombocytopenia and are frequently anemic before the transplant owing to various myelosuppressive chemotherapy regimens, some RBC transfusion can be anticipated.

The choice of ABO type of RBCs, platelets, and plasma for transfusion in the setting of ABO nonidentical transplants can be problematic (270,273). The challenge is to transfuse components simultaneously compatible with both the recipient's original and newly engrafted ABO types and to know when to convert to the ABO blood type of the transplant donor. The simplest approach is to transfuse type O RBCs, washed of their residual plasma, to all recipients. Similarly, it is simplest to transfuse AB plasma to avoid any minor incompatibility with either the patient's own original ABO-type RBCs or with the newly produced RBCs from the engrafted marrow. However, this approach is not always practical because sufficient quantities of these components may not always be available, and this may not be the best way to manage the blood component inventory. Table 72.9 presents a possible set of guidelines for the choice of ABO blood component types to transfuse during the engraftment phase of the transplant for different HPC donor and recipient ABO types.

TABLE 72.9. BLOOD TYPE SELECTION FOR ABO-INCOMPATIBLE HEMATOPOIETIC PROGENITOR CELL TRANSPLANTS

	Hematopoietic Progenitor Cell Donor ABO Type											
	O			A			B			AB		
Recipient ABO Type	RBC Type	FFP/PLT Type	2nd Choice PLT	RBC Type	FFP/PLT Type	2nd Choice PLT	RBC Type	FFP/PLT Type	2nd Choice PLT	RBC Type	FFP/PLT Type	2nd Choice PLT
O	O	O	A or B	O	A	O	O	B	O	O	AB	A
A	O	A	O	A	A	O	O	AB	A	A	AB	A
B	O	B	O	O	AB	A	B	B	O	B	AB	B
AB	O	AB	A	A	AB	A	B	AB	B	AB	AB	A or B

FFP, fresh frozen plasma; PLT, platelet.

Priority should be given to avoiding major incompatibility of transfused RBCs while significant recipient antibodies are still present in the patient's blood capable of mediating hemolysis. Components containing significant amounts of plasma, such as platelets or FFP, ideally should be compatible with both the HPC donor and the recipient's original ABO type in the immediate

posttransplant phase. However, clinical experience with “out of group” platelet transfusions indicates that transfusing a limited volume of products with minor incompatibility (i.e., passive administration of anti-A or -B) generally does not cause harm to the patient (276,277). Despite efforts to minimize the risks of ABO-incompatible HPC transplants, some hazards still exist. For example, recipients of ABO minor incompatible transplants (e.g., O HPC donor and A recipient) may hemolyze their native erythrocytes after rapid proliferation of donor lymphocytes capable of producing the hemolytic anti-A or -B isohemagglutinins (278,279). In addition, in ABO major incompatible transplants (e.g., A HPC donor and O recipient), the patient's original ABO isohemagglutinins may cause a shortened lifespan of newly produced donor-type RBCs (278,279). When the anti-donor isohemagglutinins are no longer detectable, it is safe to administer the donor-type blood components. Rh-positive recipients with Rh-negative donors should receive Rh-negative products. Rh-negative recipients with Rh-positive donors should receive Rh-positive products, unless they have a preexisting anti-D. They should receive Rh-negative RBCs until anti-D becomes undetectable and Rh-positive RBCs are compatible by crossmatch.

ADMINISTRATION OF BLOOD COMPONENTS

Part of "72 - Transfusion Therapy"

Informed Consent

Patients should be given the opportunity to accept or reject transfusion therapy and need to be given the information necessary to make an informed choice (280,281). The Joint Commission on Accreditation of Healthcare Organizations has required informed consent for blood transfusion since 1996. The purpose of the informed consent is to initiate a discussion of the need for transfusion between the health care provider and the patient and to provide the patient with information about the risks, benefits, and alternatives to transfusion. It is important for the patient's physician to document that informed consent has been given or denied. Consent for transfusion may be included in a general consent for all invasive procedures or as a dedicated consent form just for transfusion. Documentation also may be in the form of a note in the patient's medical record. The person gaining consent must have a thorough knowledge of the risks and benefits of transfusion (282).

Recent reviews are generally available that summarize the incidences of some common risks of transfusion (283) (Table 72.10).

TABLE 72.10. RISKS OF BLOOD TRANSFUSION

Risk Factor	Estimated Frequency		No. of Deaths Per Million Units
	Per Million Units	Per Actual Unit	
Infection			
Viral			
Hepatitis A	1	1/1,000,000	0
Hepatitis B	7–32	1/30,000–1/250,000	0–0.14
Hepatitis C	4–36	1/30,000–1/150,000	0.5–17
HIV	0.4–5	1/200,000–1/2,000,000	0.5–5
HTLV types I and II	0.5–4	1/250,000–1/2,000,000	0
Parvovirus B19	100	1/10,000	0
Bacterial contamination			
Red cells	2	1/500,000	0.1–0.25
Platelets	83	1/12,000	21
Acute hemolytic reactions	1–4	1/250,000–1/1,000,000	0.67
Delayed hemolytic reactions	1,000	1/1,000	0.4
Transfusion-related acute lung injury	200	1/5,000	0.2

HIV, human immunodeficiency virus; HTLV, human T-cell lymphotropic virus.
From Goodnough LT, Brecher ME, Kanter MH, et al. Transfusion medicine (first of two parts), Blood transfusion. *N Engl J Med* 1999;340:438–447, with permission.

Because of the publicity surrounding AIDS transmission by transfusion at the start of the AIDS epidemic, a major concern of the general public is the possibility of viral transmission. However, the incidence of some other adverse effects, such as septic, allergic and febrile reactions, are higher. The benefits of RBC transfusion include the reversal of the signs, symptoms, and sequelae of anemia (inadequate oxygen-carrying capacity) and the prevention of organ ischemia. The benefit of platelet transfusion is the treatment or prevention of hemorrhage owing to thrombocytopenia or thrombocytopathy. The use of plasma, cryoprecipitate, and plasma derivatives benefits patients with coagulopathy. The use of these products should be reserved for instances in which there is no drug or coagulation factor concentrate available as an alternative. Some coagulation factors are now available that have been produced by recombinant methods or have been treated to reduce the risk of disease transmission, e.g., heat-treated or SD-treated. Health care providers should not hesitate to contact the medical director of the blood bank or the pharmacy staff for information if they have questions about the availability of virus-attenuated or pharmacologic products.

Physician's Order for Transfusion

The initiation of a blood transfusion must begin with a physician's order. Transfusion orders must be made by prescription. The ordering physician is responsible for indicating the quantity and any requested special modifications of the product (e.g., leukocyte–reduced, CMV-seronegative, irradiated, volume–reduced). In addition, the ordering physician should indicate on the chart the rate of transfusion and any premedication that should be given to the patient before transfusion. Although the ultimate responsibility for the appropriateness of a transfusion order rests with the physician who is caring for the patient, there are at least two other checks on the suitability of the transfusion in the hospital setting. First, the transfusion service and blood usage committee should monitor transfusion requests for conformance to acceptable standards. Second, the nursing staff who administer most transfusions are responsible for notifying the ordering physician of any problems that they identify regarding the transfusion.

The order for blood must contain unambiguous patient identification, e.g., patient name and a unique numeric or alphanumeric identifier. The request must be accompanied by a current sample of blood from the intended recipient. The sample must be obtained within 3 days of the intended transfusion if the patient has been transfused or pregnant within the past 3 months or if the transfusion history is unknown or unreliable. Otherwise, the sample may be obtained more than 3 days in advance. If the patient has an alloantibody to RBCs, an extension beyond 3 days should not apply.

The accurate identification and labeling of the patient sample is of critical importance in ensuring a compatible transfusion. A recent report indicated that 14% of blood transfusions administered to an unintended recipient result from errors in phlebotomy or sample identification (94). For this reason, the sample must be labeled or the label must be verified at the patient's bedside. The phlebotomist must be identified, and for that reason, the signature of the phlebotomist is often required on the specimen label. Mislabeled blood samples are 40 times more likely to have a blood group discrepancy than correctly labeled blood (284). Transfusion services must have a stringent policy for rejecting mislabeled specimens, in order to protect against such errors.

Identifying the Intended Recipient

Release of a blood component from the transfusion service should require written or electronic documentation of the request for a particular blood component for a specific patient. The American Association of Blood Banks standards require that all blood components for transfusion have a tag or label attached that has the patient's two identifiers, the donor unit number, and the compatibility test results (40). In addition, the blood component must be inspected by transfusion service personnel for any irregularities, such as the presence of clots or discoloration, before it is issued. When blood is issued, care must be taken to ensure that the blood is being transfused to the correct recipient. Transfusion to the wrong recipient is the single most common cause of mortality resulting from incompatible blood transfusions (94). It is good practice for the blood bank to limit the number of blood products released at a time to a given patient, if possible, to reduce the risk and minimize the impact of misidentification errors.

The American Association of Blood Banks standards dictate that that all information “identifying the container” be matched with the patient information in the presence of the patient (40). Most institutions require that identifying information be independently verified by two persons, often referred to as readers or verifiers. One reads the patient wristband identifying the patient and the other reads the information on the blood component container. They must ensure that the intended recipient corresponds to the patient's identification. Patient identification should never rely solely on a patient's name. Moreover, although it is appropriate to ask the patient to identify him/herself, this is not sufficient identification in itself because the patient or transfusionist may misunderstand or make a misstatement. Other methods of ensuring that only the intended recipient receives the transfusion include commercial mechanical barriers (285,286). The use of mechanical barriers requires that a code be entered into a blood administration device that connects to the blood product tubing. Because mechanical barriers can be relatively expensive, easily subverted, or cumbersome, their acceptance and use remain limited at this time. Another possible system for matching the component with the intended recipient involves bar-code reading and matching by the central hospital computer of a barcoded component tag and a bar-coded patient identification bracelet (285,287). Unfortunately, no system is foolproof. All systems require careful attention and rigorous adherence to established procedures.

Intravenous Access and Compatible Intravenous Solutions

Blood and blood products are given intravenously with rare exceptions. The exceptions are intraosseous blood transfusion for emergent transfusions in pediatric patients with no other available access (288) and intraperitoneal transfusion of fetuses *in*

utero. The choice of intravenous site is usually one of convenience and has little bearing on the efficacy of transfusion if good flow rates can be maintained. Large bore intravenous access devices, e.g., 19-gauge or larger catheters or needles, are recommended to allow rapid infusion, if necessary. Smaller intravenous access devices can be used without danger of hemolysis, but flow rates will be substantially lower. Improved flow rates can be achieved if RBCs are diluted with normal saline (289,290). However, this is usually not necessary with additive-solution RBCs because the additive solution itself serves as a diluent.

Generally, only normal (0.9%) saline should be added to blood components during infusion (40). Glucose solutions should not be used because the sugar passes through the cell membrane, making the inside of the cell hyperosmotic. When these hyperosmotic units of blood are transfused, water enters the cell to restore osmotic balance. The RBCs swell in response to the influx of water and may hemolyze. Gross clumping of RBCs also may occur in the administration set when RBCs are mixed with glucose solutions (291). Calcium-containing electrolyte solutions such as lactated Ringers' solution, which contains 3 mEq/L of calcium should not be used to dilute units of RBCs. The calcium may exceed the chelating capacity of the citrate anticoagulant in the blood component, thereby promoting clot formation. Glucose- and calcium-containing solutions also should not be coadministered with blood components for the same reasons. Medications should not be added to a unit of blood component for several reasons. If the patient should experience an adverse reaction, it may be difficult to ascertain what caused the reaction, the blood or the medication. In addition, if the transfusion needs to be halted for any reason, such as for a transfusion reaction, the patient will not receive the full dose of the medication. Finally, information regarding the compatibility of most medications with stored blood components is not readily available.

Blood Filters

Blood filters are a required part of all blood administration sets. There are three general categories of blood filters: (a) standard blood filter, (b) microaggregate filter, and (c) LRFs (47,292). The standard blood filter is a 170- to 260-μm screen that is used to remove blood clots and large aggregates of cellular debris that form in blood components during storage. A standard blood filter is required for all transfusions unless another specialty filter is used instead. Microaggregate filters, conversely, have a smaller effective pore size of 20 to 40 μm. Filters with this sieving capacity can remove microaggregates, i.e., small clumps of degenerating leukocytes, platelets, and fibrin strands that form in cellular blood components after several days of storage. Microaggregates pass through standard blood filters. Microaggregate filters were popularized during the 1960s and 1970s for use in the settings of open-heart surgery and massive transfusion (293). Microaggregate debris from large-volume blood transfusions was believed to occlude end-organ capillaries and contribute to cerebral and renal dysfunction after open-heart surgery and respiratory distress after massive transfusion (293,294,295 and 296). However, later studies showed that microaggregate debris from transfusion was unlikely to be a significant contributor to those capillary occlusion syndromes (297). Because microaggregate filters can remove 80% to 95% of leukocytes from a unit of RBCs, they have also been used for the purpose of decreasing FNHTRs (298,299 and 300). However, the increasing use of more efficient LRFs (see below) has largely eliminated this application of microaggregate filters. Microaggregate filters are still used to filter salvaged autologous blood before reinfusion and for in-line filtration of extracorporal blood during CPB.

The leukocytes in units of RBC and platelet concentrates play key roles in mediating febrile transfusion reactions, HLA alloimmunization, and CMV transmission (301). Leukocyte reduction accordingly reduces the risks of these adverse effects of transfusion. LRFs were developed to selectively remove leukocytes from units of blood components (302). They work on a principle of selective adsorption of leukocytes to the filter media. They also function as clot screens owing to the mesh created by the fibers. These filters can remove 3 logs, i.e., 99.9%, or more of leukocytes from units of cellular blood components. Blood components can be leukoreduced by filtration either prestorage during component production or at the bedside during infusion. Prestorage leukocyte-reduced blood components must still be administered with the standard 170- to 260-μm filter because clots and other aggregates might still form after filtration during subsequent storage. It is critically important that filtration be done properly because filters may clog, leak, or otherwise fail to perform. Prestorage leukoreduction in the donor center or transfusion service is advantageous because more rigorous quality control can be exercised there than at the bedside. Also, platelet concentrates that have been leukocyte-reduced before storage are expected to stimulate fewer febrile transfusion reactions. This is a result of the reduction in release of leukocyte-derived cytokines during blood component storage owing to the early removal of the leukocyte source (303,304).

For a blood product to be labeled as leukoreduced, it must contain no more than 5×10^6 residual leukocytes (40). The effectiveness of this level of leukoreduction in preventing HLA alloimmunization to platelet transfusion has been convincingly demonstrated in a recent large trial, as well as several earlier studies (23,185). In addition, a large, randomized, prospective trial of 502 allogeneic bone marrow transplant patients showed that leukoreduction filtration is an effective alternative to providing CMV-seronegative blood components for the prevention of CMV disease (267). However, neither approach is fail-safe because both leukocyte-reduction by filtration and selecting CMV-seronegative donors still permits as high as a 5% to 7% CMV infection rate from transfusion. Poststorage removal of leukocytes from units of RBCs and platelet concentrates with bedside leukoreduction filtration has not decreased FNHTRs to the extent expected by many transfusion specialists based on the level of leukocyte removal now possible. Some studies documented a decrease in FNHTRs (305,306 and 307), whereas others have not (23,308). The failures appear to be due at least in part to the inability of bedside leukoreduction to remove or prevent the accumulation of pyrogenic cytokines that are released from passenger leukocytes during storage of platelet preparations. Prestorage leukoreduction of platelets is expected to be more successful (303,304). It is important to note that LRFs cannot be used

interchangeably for filtering both RBCs and platelets. LRFs are designed specifically for RBC or platelet leukoreduction; some RBC leukoreduction filters may remove large numbers of platelets.

Another possible benefit of leukoreduction is the reduction or prevention of transmission of variant Creutzfeldt-Jakob disease (vCJD) (309). Although there is no evidence that vCJD has actually been transmitted in humans by blood transfusion, experimental studies in laboratory rodents have shown that transfusion-associated transmission is possible, particularly with buffy-coat, i.e., leukocyte-enriched, fractions (310,311). Because leukocytes, in particular B lymphocytes, are suspected potential carriers of vCJD, leukocyte-reduction may decrease the risk, if any, of transfusion-transmitted CJD.

LRFs should not be used for the prevention of TA-GVHD. TA-GVHD is mediated by passenger leukocytes, specifically T lymphocytes. However, it is unclear how many T lymphocytes are necessary to mediate TA-GVHD or whether LRFs are effective in reducing T-lymphocyte levels sufficiently to prevent TA-GVHD. Prophylaxis of TA-GVHD should instead be achieved through gamma irradiation of cellular blood components, which eliminates the proliferative potential of transfused lymphocytes (271,272). The dose of radiation should be at least 24 Gy delivered to the midplane of the canister or radiation field where the components are placed. The minimal dose at any part of the blood component should be 15 Gy (40).

Rate of Transfusion and Blood Pumps

The rate at which blood components should be administered is dictated by several factors, including the urgency of the clinical setting, the patient's tolerance of added intravascular volume, and regulatory time limits after some manipulations of blood components (e.g., thawing, pooling, removal from refrigeration). Although rapid administration of blood components may be lifesaving in emergent situations, it also leads to several unfavorable outcomes (e.g., increases in allergic transfusion reactions, febrile transfusion reactions, volume overload, hypothermia). In massive transfusion, cardiac arrhythmias have been reported to occur from transfusion of refrigerated blood at rates exceeding 100 mL/min (74,75,76 and 77). A typical and usually well-tolerated rate of transfusion in patients who are not experiencing congestive heart failure is one unit of RBC in 1 to 2 hours. The infusion should start slowly at a rate of approximately 25 to 50 mL over the first 15 minutes while the patient is closely observed for possible transfusion reactions. In patients with cardiac compromise, the infusion rate must be slowed to avoid volume overload, and the patient may need to be treated with diuretics. In addition, units of RBC can be divided into split units, i.e., half units, that can each be transfused slowly over as long as 4 hours each. The other half of a split unit may be transfused as long as 24 hours later as long as it is stored at 1° to 6°C in a blood bank–monitored refridgerator.

Blood components released from the transfusion service should be administered as soon as possible and within 4 hours of release from the transfusion service owing to the possibility of significant bacterial growth at room temperature. RBCs must not be allowed to warm up to more than 10°C if they are not transfused. When a unit of blood has been warmed to above 10°C, it must be transfused within 4 hours or else discarded because of the increased risk of bacterial growth in the unit (40). Temperature sensors are available to attach directly to RBC units to verify that 10°C has not been exceeded. Alternatively, restrictions may be placed on how long a unit may be kept unrefrigerated after release from the transfusion service. In general, the units should be refrigerated or returned to the transfusion service within 30 minutes, if not transfused (312). Platelet concentrates, whose sterility is broken when they are pooled, must be transfused within 4 hours of pooling because of concerns over bacterial growth (40).

Most transfusions flow at an acceptable rate by the force of gravity alone. In some situations, it may be desirable to control more precisely the rate of infusion with an infusion pump (313). For example, maintaining a precise, slow rate may be important for small-volume pediatric transfusions or for adults at high risk of adverse effects of transfusion, such as hypervolemia. When using electromechanical pumps for RBC transfusion, hemolysis is a consideration. It is essential to verify with the manufacturer that the pump has been tested and found safe for infusing RBCs with minimal hemolysis. Similar verification is prudent for use of pumps for other blood components as well (314,315).

Monitoring the Transfusion

Baseline vital signs should be obtained before transfusion. Appropriate documentation, including the type of component, the volume desired, the duration of transfusion, and the infusion rate, should be included in a progress or procedure note in the patient's chart. Use of any blood warmers or filters should also be appropriately documented in the chart. The patient should be observed closely throughout the transfusion. Symptoms of a severe hemolytic transfusion reaction often manifest themselves within the first 15 minutes of a transfusion. Thus, it is advisable to remain with the patient at the beginning of a transfusion. The patient's vital signs should be monitored at various times during the transfusion. In a typical hospital protocol, vital signs are taken pretransfusion, then at 15 minutes into the transfusion, and hourly thereafter. After the transfusion is completed, the patient should be tested to ensure that the transfusion outcome was appropriate and as expected. For example, a platelet count should be measured 10 minutes to 4 hours after a platelet transfusion to ensure that the patient responded to a platelet transfusion, a hematocrit should be measured after a RBC transfusion, and appropriate coagulation testing should be done after a plasma transfusion. Vital signs should be monitored for several hours after transfusion to rule out a late-developing reaction.

Blood Warming

Blood warming is not necessary for most nonemergent transfusions because the rapid transfusion of cold RBCs or plasma is unnecessary for most patients. Conversely, the urgent and massive transfusion of cold (1° to 10°C) blood at a rate of one unit every 5 to 10 minutes has been associated with the development of hypothermia, decreased sinoatrial node temperature, and ventricular fibrillation (74,75,76 and 77). Warming of cold units of RBCs can avoid these cardiovascular complications. The indications for

warming RBC units include rapid transfusion, which can be defined as more than 50 mL/kg/h in adults and more than 15 mL/kg/h in children, particularly if through a central venous catheter, and exchange transfusions in infants (292). Blood warming may also be indicated for the transfusion of RBCs in patients with pathologic cold agglutinins. Warming should only be attempted with devices that are FDA approved for transfusion applications. Blood components can be warmed in several ways (78,292,313). The simplest method is to immerse the entire bag in warm water in a thermostatically controlled water bath. It is recommended that the temperature of warmed blood components not be allowed to exceed 42°C with this or any other warming technique to avoid the possibility of hemolysis (40). Microwave warmers are available for heating plasma. Because they cause focal overheating, which can induce hemolysis in units of RBCs, they are not used for warming RBCs (316,317). Microwave devices rotate the plasma to prevent excessive heating in any one part of the bag. An additional technique for warming RBCs is rapid admixture with heated saline (78,79). This has the advantage of being one of the fastest methods and also simultaneously reduces the viscosity of RBCs, allowing rapid infusion. Admixture with heated saline has been used in the trauma setting where rapid infusion rates are necessary to treat massive hemorrhage. Other warming devices include dry heating blocks with electric warming plates and countercurrent heat exchangers, in which the heated water is rapidly circulated through water jackets surrounding the flowing RBCs (318). All these devices need to be properly quality controlled at regular intervals and monitored to ensure maintenance of the proper temperature. They should have visible thermometers and audible alarms to warn of excessive temperature (greater than 42°C). Despite the availability of warming methods, the routine use of blood warming generally should be avoided because (a) there is a risk of hemolysis when RBC units are warmed improperly, (b) warming adds expense and often time to the transfusion process, (c) warming adds a step that requires quality control monitoring, and (d) warming provides no clinical benefit for most patients.

Out-of-Hospital Transfusion

The delivery of medical care has increasingly shifted from the hospital inpatient setting to an outpatient or ambulatory setting for reasons related to cost and customer service. As part of the ambulatory care movement, transfusions are increasingly being administered in nonhospital settings, such as patients' homes, physicians' offices, convalescent homes, and freestanding outpatient surgical or hemodialysis centers (319,320 and 321). Of the various out-of-hospital settings, the patient's home offers the greatest challenges and risks because of the absence of emergency medical care facilities and staff that may be needed for the treatment of acute transfusion reactions. Nevertheless, many patients have benefited from home transfusion. Elderly patients who lack transportation and are in stable condition are potential candidates, as are some children who find the hospital setting to be stressful and intimidating. Chronically anemic patients, thrombocytopenic oncology patients, or pediatric oncology patients are among those most commonly receiving home transfusion. However, home transfusion should not be performed just for convenience. There should be a sound medical or logistical reason for performing a transfusion in the home, and that reason should outweigh the risk of a serious transfusion reaction in a setting where necessary treatment may not be immediately available. Most patients who are fully ambulatory and can travel to the hospital for transfusion should do so.

The local state laws regarding the infusion of medications by nonphysicians must be checked before initiating a home transfusion program to assure that the planned program is in compliance with local laws. Transfusions in a nontraditional setting should follow the same basic principles and practices as transfusions in the hospital. Standard operating procedures outlining all procedures of the home transfusion must be written and followed. The training requirements of the transfusionist must be established and met (322). Not only should the transfusionists be trained and experienced in administering blood components, but they also should be trained in emergency procedures, including CPR. The blood samples for typing, screening, and cross matching must be appropriately labeled, and a procedure must be in place and followed to appropriately identify the patient and blood component before transfusion. The procedure for transportation of the blood components to the transfusion site must assure that the component bags are protected and maintained within an acceptable temperature range (1° to 6°C for RBCs and 20° to 24°C for platelets) in appropriately insulated containers. A mechanism for verifying that the temperature of the components remains within acceptable limits during shipment to the transfusion site either with a thermometer or other temperature sensor should be in place. A limit to the number of units of blood component administered during home transfusion should be established, such as no more than two units of RBCs or one pool of platelets per day per patient. This will help limit the risk of adverse reactions. In addition, the disposal of waste material, such as sharps and blood-soaked gauze and needles, needs to be planned to ensure compliance with all appropriate safety codes. Quality assurance monitoring should be established for home transfusion to assure that appropriate procedures are followed during all steps of the process, the patient's needs are being met, and patient safety is being preserved.

The home must have suitable space for the patient to receive the transfusion, along with space for the transfusionist to set up a working area (321). The home must have a working telephone for possible use in obtaining emergency assistance. Physician backup, either in an office or emergency department, must be immediately available by phone. The provider should never be alone in the house with the patient; at least one other responsible and physically and mentally competent adult should be present in the home who can speak English. This allows the provider to take care of the patient while the other individual is available to call for help, if necessary. It is important that the blood product(s) be infused at an appropriate rate, i.e., within regulatory requirements (see above), but not too rapidly, to decrease the risks of transfusion reactions. Serious consideration should be given to using only prestorage leukocyte-reduced blood components to decrease the risk of febrile reactions. An adult should stay with the patient, ideally for at least several hours, after the transfusion to obtain help in case the patient suffers a late-developing transfusion reaction. Blood specimens for

relevant posttransfusion laboratory tests should be drawn at an appropriate interval after transfusion to ensure that the patient has responded as expected to the transfusion. Failure of the hematocrit to increase as expected could indicate the possibility of an overlooked RBC antibody mediating a hemolytic reaction or of some other cause of a worsening anemia that needs further treatment or additional transfusion. Failure to obtain an expected platelet count increment could indicate platelet refractoriness or ongoing consumption that needs further workup. With appropriate patient monitoring and when meticulous attention is paid to establishing and following appropriate standard operating procedures, home blood transfusion has proven to be a generally safe medical practice.

REFERENCES

1. NIH Consensus Conference. Perioperative red blood cell transfusion. *JAMA* 1988;260:2700–2703.

2. American Society of Anesthesiologists Task Force on Blood Component Therapy. Practice guidelines for blood component therapy. *Anesthesiology* 1996;84:732–747.

3. Simon TL, Alverson DC, AuBuchon J, et al. Practice parameter for the use of red blood cell transfusions. *Arch Pathol Lab Med* 1998;122:130–138.

4. Weiskopf R, Viele M, Feiner J, et al. Human cardiovascular and metabolic response to acute, severe isovolemic anemia. *JAMA* 1998;279:217–221.

5. Hébert PC, Wells G, Blachman MA, et al. A multicenter, randomized controlled clinical trial of transfusion requirements in critical care. *N Engl J Med* 1999;340:409–417.

6. Valeri CR, Crowley JP, Loscalzo J. The red cell transfusion trigger: has a sin of commission now become a sin of omission? *Transfusion* 1998;38:602–610.

7. Weiskopf R. Do we know when to transfuse red cells to treat acute anemia? *Transfusion* 1998;38:517–521.

8. Gaydos LA, Freireich EJ, Mantel N. The quantitative relation between platelet count and hemorrhage in patients with acute leukemia. *N Engl J Med* 1962;266:905–909.

9. NIH Consensus Conference. Platelet transfusion therapy. *JAMA* 1987;257:1777–1780.

10. Task Force of the College of American Pathologists. Practice parameters for the use of fresh-frozen plasma, cryoprecipitate and platelets. *JAMA* 1994;253:277–281.

11. Norfolk DR, Ancliffe PJ, Contreras M, et al. Consensus conference on platelet transfusion, Royal College of Physicians of Edinburgh, 27–28 November 1997. *Br J Haematol* 1998;101:609–617.

12. Gmur J, Burger J, Schanz V, et al. Safety of stringent prophylactic platelet transfusion policy for patients with acute leukaemia. *Lancet* 1991;338:1223–1226.

13. Heckman KD, Weiner GJ, Davis CS, et al. Randomized study of prophylactic platelet transfusion threshold during induction therapy for adult acute leukemia: 10,000/μL versus 20,000/μL. *J Clin Oncol* 1997;15:1143–1149.

14. Wandt H, Frank M, Ehninger G, et al. Safety and cost effectiveness of a 10×10^9/L trigger for prophylactic platelet transfusions compared with the traditional 20×10^9/L trigger: a prospective comparative trial in 105 patients with acute myeloid leukemia. *Blood* 1998;91:3601–3606.

15. Rebulla P, Finazzi G, Marangoni F, et al. The threshold for prophylactic platelet transfusions in adults with acute myeloid leukemia. *N Engl J Med* 1997;337:1870–1875.

16. Gil-Fernandez JJ, Alegre A, Fernández-Villalta MJ, et al. Clinical results of a stringent policy on prophylactic platelet transfusion: non-randomized comparative analysis in 190 bone marrow transplant patients from a single institution. *Bone Marrow Transplant* 1996;18:931–935.

17. Gratwohl A, Tichelli A. Prophylactic platelet transfusions in acute myeloid leukemia [Letter]. *N Engl J Med* 1998;337:1468.

18. Menitove JE, Snyder EL. Platelet transfusion practice: time for renewed consensus. *Transfusion* 1998;38:707–709.

19. Strauss RG. Clinical perspectives of platelet transfusion defining optimal dose. *J Clin Apheresis* 1995;10:124–127.

20. Slichter SJ. Optimizing platelet transfusions in chronically thrombocytopenic patients. *Semin Hematol* 1998;35:269–278.

21. Harker LA, Roskos L, Cheung E. Effective and efficient platelet transfusion strategies that maintain hemostatic protection. *Transfusion* 1998;38:619–621.

22. Hersh JK, Hom EG, Brecher ME. Mathematical modeling of platelet survival with implications for optimal transfusion practice in the chronically platelet transfusion-dependent patient. *Transfusion* 1998;38:637–644.

23. The Trial to Reduce Alloimmunization to Platelets Study Group. Leukocyte reduction and ultraviolet B irradiation of platelets to prevent alloimmunization and refractoriness to platelet transfusions. *N Engl J Med* 1997;337:1861–1869.

24. Daly PA, Schiffer CA, Aisner J, et al. Platelet transfusion therapy; one-hour posttransfusion increments are valuable in predicting the need for HLA-matched preparations. *JAMA* 1980;243:435–438.

25. Bishop JF, McGrath K, Wolf MM, et al. Clinical factors influencing the efficacy of pooled platelet transfusions. *Blood* 1988;71:383-7.

26. Alcorta I, Pereira A, Ordinas A. Clinical and laboratory factors associated with platelet transfusion refractoriness: a case-control study. *Br J Haematol* 1996;93:220–224.

27. NIH Consensus Conference. Fresh frozen plasma: indications and risks. *JAMA* 1985;253:551–553.

28. Klein HG, Dodd RY, Dzik WH, et al. Current status of solvent/detergent-treated frozen plasma. *Transfusion* 1998;38:102–107.

29. Bianco C. Choice of plasma preparations for transfusion. *Transfus Med Rev* 1999;13:84–88.

30. Byrnes JJ, Moake JL, Klug P, et al. Effectiveness of the cryosupernatant fraction of plasma in the treatment of refractory thrombotic thrombocytopenia purpura. *Am J Hematol* 1990;34:169–174.

31. Rock G, Shumak KH, Sutton DM, et al. Cryosupernatant as replacement fluid for plasma exchange in thrombotic thrombocytopenic purpura. *Br J Haematol* 1996;94:383-6.

32. Menitove JE. Preparation and clinical use of plasma and plasma fractions. In: Williams WJ, Beutler E, Erslev AJ, et al., eds. *Hematology*, 4th ed. New York: McGraw-Hill 1990:1669–1673.

33. Aggeler PM. Physiological basis for transfusion therapy in hemorrhagic disorders: a critical review. *Transfusion* 1961;1:71–86.

34. Ciavarella D, Reed RL, Counts RB, et al. Clotting factor levels and the risk of diffuse microvascular bleeding in the massively transfused patient. *Br J Haematol* 1987;67:365–368.

35. Gibble JW, Ness PM. Fibrin glue: the perfect operative sealant? *Transfusion* 1990;30:741–747.

36. Spotnitz WD, Welker RL. Clinical uses of fibrin sealant. In: Mintz PD, ed. *Transfusion therapy: clinical principles and practice*. Bethesda, MD: American Association of Blood Banks, 1999:199–222.

37. Radosevich M, Goubran HA, Burnouf T. Fibrin sealant: scientific rationale, production methods, properties, and current clinical use. *Vox Sang* 1997;72:133–143.

38. Strauss RG. Granulocyte transfusion therapy In: Mintz PD, ed. *Transfusion therapy: clinical principles and practice*. Bethesda, MD: AABB Press, 1999:81–96.

39. Strauss RG. Granulocyte transfusions. In: Rossi EC, Simon TL, Moss GS, Gould SA, eds. *Principles of transfusion medicine*, 2nd ed. Baltimore: Williams & Wilkins, 1996:321–328.

40. Menitove JE, ed. *Standards for blood banks and transfusion services*, 19th ed. Bethesda, MD: American Association of Blood Banks, 1999.

41. Dale DC, Liles WC, Llewellyn C, et al. Neutrophil transfusions: kinetics and functions of neutrophils mobilized with granulocyte-colony-stimulating factor and dexamethasone. *Transfusion* 1998;38:713–721.

42. Strauss RG. Neutrophil (granulocyte) transfusions in the new millennium. *Transfusion* 1998;38:710–702.

43. Nilsson L, Hedner U, Nilsson IM, et al. Shelf-life of banked blood and stored plasma with special reference to coagulation factors. *Transfusion* 1983;23:377–381.

44. American College of Surgeons Committee on Trauma. *Advanced trauma life support manual*. Chicago: American College of Surgeons, 1993.

45. Moore FD. The effects of hemorrhage on body composition. *N Engl J Med* 1965;273:567–577.

46. Ebert RV, Stead EA Jr, Gibson JG. Response of normal subjects to acute blood loss. *Arch Intern Med* 1941;68:578–590.

47. Vengelen-Tyler V, ed. *Technical manual*, 13th ed. Bethesda, MD: American Association of Blood Banks, 1999.

48. Schonewille H, Haak HL, van Zijl AM. Alloimmunization after blood transfusion in patients with hematologic and oncologic diseases. *Transfusion* 1999;39:763–771.

49. Vichinsky EP, Earles A, Johnson RA, et al. Alloimmunization in sickle cell anemia and transfusion of racially unmatched blood. *N Engl J Med* 1990;332:1617–1621.

50. Rosse WF, Gallagher D, Kinney TR, et al. Transfusion and alloimmunization in sickle cell disease. The Cooperative Study of Sickle Cell Disease. *Blood* 1990;76:1431–1437.

51. Blumberg N, Bove JR. Un-crossmatched blood for emergency transfusion: one year's experience in a civilian setting. *JAMA* 1978;240:2057–2059.

52. Lefebre J, McLellan B, Coovadia AS. Seven years experience with group O unmatched packed red blood cells in a regional trauma unit. *Ann Emerg Med* 1987;16:1344–1359.

53. Schwab CW, Shayne JP, Turner J. Immediate trauma resuscitation with type O uncrossmatched blood: a two-year prospective experience. *J Trauma* 1986;26:897–902.

54. Mollison PL, Engelfriet CP, Contreras M. The Rh blood group system (and LW). In: *Blood transfusion in clinical medicine*, 10th ed. Oxford: Blackwell Scientific, 1997:151–185.

55. Collins JA. Problems associated with the massive transfusion of stored blood. *Surgery* 1974;75:274–295.

56. Counts RB, Haisch C, Simon TL, et al. Hemostasis in massively transfused trauma patients. *Ann Surg* 1979;190:91–99.

57. Hewson JR, Neame PB, Kumar N, et al. Coagulopathy related to dilution and hypotension during massive transfusion. *Crit Care Med* 1985;13:387–391.

58. Murray DJ, Pennell BJ, Weinstein S, et al. Packed red cells in acute blood loss: dilutional coagulopathy as a cause of surgical bleeding. *Anesth Analg* 1995;80:336–342.

59. Hiippala S. Replacement of massive blood loss. *Vox Sang* 1998;74(Suppl 2):399–407.

60. Slichter SP. Identification and management of defects in platelet hemostasis in massively transfused patients. In: Collins JA, Murawski K, Shafer AW, eds. *Massive transfusion in surgery and trauma*. New York: Alan R. Liss, 1982:225–258.

61. Mannucci PM, Federici AB, Sirchia G. Hemostasis testing during massive blood replacement: a study of 172 cases. *Vox Sang* 1982;42:113–123.

62. Martin DJ, Lucas CE, Ledgerwood AM, et al. Fresh frozen plasma supplement to massive red blood cell transfusion. *Ann Surg* 1985;202:505–511.

63. Reed RL 2nd, Ciavarella D, Heimbach DM, et al. Prophylactic platelet administration during massive transfusion. *Ann Surg* 1986;203:40–48.

64. Dzik WH, Kirkley SA. Citrate toxicity during massive blood transfusion. *Transfus Med Rev* 1988;2:76–94.

65. Marquez J, Martin D, Virji MA, et al. Cardiovascular depression secondary to ionic hypocalcemia during hepatic transplantation in humans. *Anesthesiology* 1986;65:457–461.

66. Denlinger JK, Nahrwold ML, Gibbs PS. Hypocalcemia during rapid blood transfusion in an anesthetized man. *Br J Anaesth* 1976;48:995–1000.

67. Hester JP, McCullough J, Mishler JM, et al. Dosage regimens for citrate anticoagulants. *J Clin Apheresis* 1983;1:149–152.

68. McLellan BA, Reid SR, Land PL. Massive blood transfusion causing hypomagnesemia. *Crit Care Med* 1984;12:146–147.

69. Wolf PL, McCarthy LJ, Hafleigh B. Extreme hypercalcemia following blood transfusion combined with intravenous calcium. *Vox Sang* 1970;19:544–545.

70. Latham JT Jr, Bove JR, Weirich FL. Chemical and hematologic changes in stored CPDA-1 blood. *Transfusion* 1982;22:158–159.

71. Driscoll DF, Bistrian BR, Jenkins RL, et al. Development of metabolic alkalosis after massive transfusion during orthopedic liver transplantation. *Crit Care Med* 1987;15:905–908.

72. Blanchette VS, Gray E, Hardie MJ, et al. Hyperkalemia after neonatal exchange transfusion: risk eliminated by washing red cell concentrates. *J Pediatr* 1984;105:321–324.

73. Carmichael D, Hosty T, Kastl D, et al. Hypokalemia and massive transfusion. *South Med J* 1984;77:315–317.

74. Boyan CP. Cold or warmed blood for massive transfusions. *Ann Surg* 1964;160:282–286.

75. Collins JA. Problems associated with massive transfusion of stored blood. *Surgery* 1974;75:274–295.

76. Boyan CP, Howland WS. Cardiac arrest and temperature of bank blood. *JAMA* 1963;183:58–60.

77. Boyan CP, Howland WS. Blood temperature: a critical factor in massive transfusion. *Anesthesiology* 1961;22:559–563.

78. Iserson KV, Huestis DW. Blood warming: current applications and techniques. *Transfusion* 1991;31:558–571.

79. Cohn SM, Stack G. *In vitro* comparison of heated saline-blood admixture to a heat exchanger for rapid warming of red blood cells. *J Trauma* 1993;35:688–691.

80. Wayne AS, Kevy SV, Nathan DG. Transfusion management of sickle cell disease. *Blood* 1993;81:1109–1123.

81. Sharon BI. Transfusion therapy in congenital hemolytic anemias. *Hematol Oncol Clin North Am* 1994;8:1053–1086.

82. Rosse WF, Telen M, Ware RE. *Transfusion support for patients with sickle cell disease*. Bethesda, MD: AABB Press, 1998.

83. Fluit CR, Kunst VA, Drenthe-Schonk AM. Incidence of red cell antibodies after multiple blood transfusion. *Transfusion* 1990;30:532–535.

84. Spanos T, Karageorga M, Ladis V, et al. Red cell alloantibodies in patients with thalassemia. *Vox Sang* 1990;58:50–55.

85. Rosse WF, Gallagher D, Kinney TR, et al. Transfusion and alloimmunization in sickle cell disease. *Blood* 1990;76:1431–1437.

86. Blumberg N. Beyond ABO and D antigen matching: how far and for whom? [Editorial]. *Transfusion* 1990;30:482–484.

87. Ness PM. To match or not to match: the question for chronically transfused patients with sickle cell anemia [Editorial]. *Transfusion* 1994;34:558–560.

88. Olivieri NF, Brittenham GM. Iron-chelating therapy and the treatment of thalassemia. *Blood* 1997;89:739–761.

89. Hoffbrand AV, Wonke B. Iron chelation therapy. *J Intern Med* 1997;242(Suppl 740):37–41.

90. Goodnough LT, Brecher ME, Kanter MH, et al. Transfusion medicine. Blood conservation. *N Engl J Med* 1999;340:525–533.

91. National Heart, Lung, and Blood Institute Expert Panel on the Use of Autologous Blood. Transfusion alert: use of autologous blood. *Transfusion* 1995;35:703–711.

92. Yomtovian R. Practical aspects of preoperative autologous transfusion. *Am J Clin Pathol* 1997;107(Suppl 1);S28–S35.

93. Linden JV, Kruskall MS. Autologous blood: always safer? *Transfusion* 1997;37:455–456.

94. Linden JV. Errors in transfusion medicine. Scope of the problem. *Arch Path Lab Med* 1999;123:563–565.

95. Goodnough LT, Monk TG, Brecher ME. Autologous blood procurement in the surgical setting: lessons learned in the last 10 years. *Vox Sang* 1996;71:133–141.

96. Etchason J, Petz L, Keeler E, et al. The cost-effectiveness of preoperative autologous blood donations. *N Engl J Med* 1995;332:719–724.

97. Renner SW, Howanitz PJ, Bachner P. Preoperative autologous blood donation in 612 hospitals. *Arch Pathol Lab Med* 1992;116:613–619.

98. Kasper SM, Gerlich W, Buzello W. Preoperative red cell production in patients undergoing weekly autologous blood donation. *Transfusion* 1997;37:1058–1062.

99. Kanter MH, van Maanen D, Anders KH, et al. Preoperative autologous blood donations before elective hysterectomy. *JAMA* 1996;276:798–801.

100. Kickler TS, Spivak JL. Effect of repeated whole blood donation on serum immunoreactive erythropoietin levels in autologous donors. *JAMA* 1988;260:65–67.

101. Goodnough LT, Brittenham G. Limitations of the erythropoietic response in serial phlebotomy: implications for autologous blood donor programs. *Lab Clin Med* 1990;115:28–35.

102. Cazzola M, Mercuriali F, Brugnara C. Use of recombinant human erythropoietin outside the setting of uremia. *Blood* 1997;89:4248–4267.

103. Goldberg MA. Perioperative epoetin alfa increases red blood cell mass and reduces exposure to transfusions: results of randomized clinical trials. *Semin Hematol* 1997;34(3 Suppl 2):41–47.

104. Goodnough LT, Monk TG. Erythropoietin therapy in the perioperative setting. *Clin Orthop* 1998;357:82–88.

105. Monk TG, Goodnough LT. Acute normovolemic hemodilution. *Clin Orthop* 1998;357:74–81.

106. Goodnough LT, Monk TG, Brecher ME. Acute normovolemic hemodilution should replace the preoperative donation of autologous blood as a method of autologous-blood procurement. *Transfusion* 1998;38:473–476.

107. Rottman G, Ness PM. Acute normovolemic hemodilution is a legitimate alternative to allogeneic blood transfusion. *Transfusion* 1998;38:477–480.

108. Shah DM, Prichard MN, Newell JC, et al. Increased cardiac output and oxygen transport after intraoperative isovolemic hemodilution. *Arch Surg* 1980;115:597–600.

109. Brecher ME, Rosenfeld M. Mathematical and computer modeling of acute normovolemic hemodilution. *Transfusion* 1994;34:176–179.

110. Weiskopf RB. Mathematical analysis of isovolemic hemodilution indicates that it can decrease the need for allogeneic blood transfusion. *Transfusion* 1995;35:37–41.

111. Sowade O, WarnkeH, Scigalla P, et al. Avoidance of allogeneic blood transfusions by treatment with recombinant human erythropoietin in patients undergoing open-heart surgery. *Blood* 1997;89:411–418.

112. Kick O, Daniel E. Mathematical considerations in the practice of acute normovolemic hemodilution. *Transfusion* 1997;37:141–143.

113. AABB Autologous Transfusion Committee. *Guidelines for blood recovery and reinfusion in surgery and trauma*. Bethesda, MD: American Association of Blood Banks, 1997.

114. Stack G, Snyder EL. Alternatives to perioperative blood transfusion. *Adv Anesthesiol* 1991;8:209–240.

115. Williamson KR, Taswell HF. Intraoperative blood salvage: a review. *Transfusion* 1991;31:662–675.

116. McCarthy PM, Popovsky MA, Schaff HV, et al. Effect of blood conservation efforts in cardiac operations at the Mayo Clinic. *Mayo Clin Proc* 1988;63:225–229.

117. Page R, Russell GN, Fox MA, et al. Hard-shell cardiotomy reservoir for reinfusion of shed mediastinal blood. *Ann Surg* 1989;48:514–517.

118. Yaw PB, Sentony M, Link WJ, et al. Tumor cells carried through autotransfusion contraindications to intraoperative blood recovery? *JAMA* 1975;231:490–492.

119. Edelman MJ, Potter P, Mahaffey KA, et al. The potential for reintroduction of tumor cells during intraoperative blood salvage: reduction of risk with use of the RC-400 leukocyte depletion filter. *Urology* 1996;47:179–181.

120. Duncan SE, Edwards WH, Dale WA. Caution regarding autotransfusion. *Surgery* 1974;76:1024–1030.

121. Griffith LD, Billman GF, Daily PO, et al. Apparent coagulopathy caused by infusion of shed mediastinal blood and its prevention by washing of the infusate. *Ann Thorac Surg* 1989;47:400–406.

122. Clements DH, Sculco TP, Burke SW, et al. Salvage and reinfusion of postoperative sanguineous wound drainage. *J Bone Joint Surg Am* 1992;74A:646–651.

123. Woda R, Tetzlaff JE. Upper airway oedema following autologous blood transfusion from a wound drainage system. *Can J Anaesth* 1992;S9:290–292.

124. Wallace EL, Surgenor DM, Hao HS, et al. Collection and transfusion of blood and blood components in the United States, 1989. *Transfusion* 1993;33:139–144.

125. Stover EP, Siegel LC, Parks R, et al. Variability in transfusion practice for coronary artery bypass surgery persists despite national consensus guidelines. *Anesthesiology* 1998;88:327–333.

126. Surgenor DM, Churchill WH, Wallace EL, et al. The specific hospital significantly affects red cell and component transfusion practice in coronary artery bypass graft surgery: a study of five hospitals. *Transfusion* 1998;38:122–134.

127. Bevan DH. Cardiac bypass haemostasis: putting blood through the mill. *Br J Haematol* 1999;104:208–219.

128. Burns ER, Billet HH, Frater RW, et al. The preoperative bleeding time as a predictor of postoperative hemorrhage after cardiopulmonary bypass. *J Thorac Cardiovasc Surg* 1986;92:310–312.

129. Michaelson EL, Torosian M. Morganroth J, et al. Early recognition of surgically correctable causes of excessive mediastinal bleeding after coronary artery bypass graft surgery. *Am J Surg* 1980;139:313–317.

130. Mammen EF, Koets MH, Washington BC, et al. Hemostasis changes during cardiopulmonary bypass surgery. *Semin Thromb Hemost* 1985;11:281–292.

131. Bick RL. Hemostasis defects associated with cardiac surgery, prosthetic devices and other extracorporeal circuits. *Semin Thromb Hemost* 1985;11:249–280.

132. Harker LA, Malpass TW, Branson HE, et al. Mechanism of abnormal bleeding in patients undergoing cardiopulmonary bypass: acquired transient platelet dysfunction associated with selective alpha granule release. *Blood* 1990;56:824–834.

133. Simon TL, Akl BF, Murphy W. Controlled trial of routine administration of platelet concentrates in cardiopulmonary bypass surgery. *Ann Thorac Surg* 1984;37:359–364.

134. Gelb AB, Roth RI, Levin J, et al. Changes in coagulation during and following cardiopulmonary bypass. *Am J Clin Pathol* 1996;106:87–99.

135. Simon TL, Henderson R. Coagulation factor activity in platelet concentrates. *Transfusion* 1979;19:186–189.

136. Jefferies LC. Transfusion therapy in autoimmune hemolytic anemia. *Hematol Oncol Clin North Am* 1994;8:1087–1104.

137. Engelfriet CP, Overbeeke MA, von dem Borne AE. Autoimmune hemolytic anemia. *Semin Hematol* 1992;29:3–12.

138. Mollison PL, Engelfriet CP, Contreras M. Red cell antibodies against self antigens, bound antigens and induced antigens, In: *Blood transfusion in clinical medicine*, 10th ed. London: Blackwell Scientific Publications, 1997:213–240.

139. Schwartz RS, Berkman EM, Silberstein LE. Autoimmune hemolytic anemia, In: Hoffman R, Benz EJ Jr, Shatil SJ, et al., eds, *Hematology. Basic principles and practice*, 3rd ed. New York: Churchill-Livingstone, 2000:611–630.

140. Branch DR, Petz LD. Detecting alloantibodies in patients with autoantibodies. *Transfusion* 1999;39:6–10.

141. Leger RM, Garratty G. Evaluation of methods for detecting alloantibodies underlying warm autoantibodies. *Transfusion* 1999;39:11–16.

142. Warkentin TE, Smith JW. The alloimmune thrombocytopenic syndromes. *Transfus Med Rev* 1997;4:296–307.

143. Harker LA, Slichter SJ. The bleeding time as a screening test for evaluation of platelet function. *N Engl J Med* 1972;287:155–159.

144. Carr JM, Kruskall MS, Kaye JA, Robinson SH. Efficacy of platelet transfusions in immune thrombocytopenia. *Am J Med* 1986;80:1051–1054.

145. George SH, Woolf SH, Raskob GE, et al. Idiopathic thrombocytopenic purpura: a practice guideline developed by explicit methods for the American Society of Hematology. *Blood* 1996;98:3–40.

146. Bauman MA, Menitove JE, Aster Rh, et al. Urgent treatment of idiopathic thrombocytopenic purpura with single dose gammaglobulin infusion followed by platelet transfusion. *Ann Intern Med* 1986;104:808–809.

147. Scaradavou A, Woo B, Woloski BMR, et al. Intravenous anti-D treatment of immune thrombocytopenic purpura: experience in 272 patients. *Blood* 1997;89:2689–2700.

148. Moake JL, Chow TW. Thrombotic thrombocytopenic purpura: understanding a disease no longer rare. *Am J Med Sci* 1998;316:105–119.

149. Furlan M, Robles R, Galbusera M, et al. Von Willebrand factor-cleaving protease in thrombotic thrombocytopenic purpura and hemolytic-uremic syndrome. *N Engl J Med* 1998;339:1578–1584.

150. Tsai H-M, Lian EC-Y. Antibodies to von Willebrand factor-cleaving protease in acute thrombotic thrombocytopenic purpura. *N Engl J Med* 1998;339:1585–1594.

151. Furlan M, Robles R, Solenthaler M, et al. Acquired deficiency of von Willebrand factor-cleaving protease in a patient with thrombotic thrombocytopenic purpura. *Blood* 1998;91:2839–2846.

152. Gordon LI, Kwaan HC, Rossi EC. Deleterious effects of platelet transfusions and recovery thrombocytosis in patients with thrombotic microangiopathy. *Semin Hematol* 1987;24:194–201.

153. Harkness DR, Byrnes JJ, Lian ECY, et al. Hazard of platelet transfusion in thrombotic thrombocytopenic purpura. *JAMA* 1981;246:1931–1933.

154. Kwaan HC, Soff GA. Management of thrombotic thrombocytopenic purpura and hemolytic uremic syndrome. *Semin Hematol* 1997;34:159–166.

155. Rock GA, Shumak KH, Buskard NA, et al. Comparison of plasma exchange with plasma infusion in the treatment of thrombotic thrombocytopenic purpura. Canadian Apheresis Study Group. *N Engl J Med* 1991; 325:393–397.

156. Bell WR, Braine HG, Ness PM, et al. Improved survival in thrombotic thrombocytopenic purpura-hemolytic uremic syndrome. *N Engl J Med* 1991;325:398–403.

157. Byrnes JJ, Moake JL, Klug P, et al. Effectiveness of the cryosupernatant fraction of plasma in the treatment of refractory thrombotic thrombocytopenic purpura. *Am J Hematol* 1990; 34:169–174.

158. Rock G, Shumak KH, Sutton DM, et al. Cryosupernatant as replacement fluid for plasma exchange in thrombotic thrombocytopenic purpura. *Br J Haematol* 1996;94:383–386.

159. Harrison CN, Lawrie AS, Iqbal A, et al. Plasma exchange with solvent/detergent-treated plasma of resistant thrombotic thrombocytopenic purpura. *Br J Haematol* 1996;94:756–758.

160. Horowitz MS, Pehta JC. SD plasma in TTP and coagulation factor deficiencies for which no concentrates are available. *Vox Sang* 1998;74:231–235.

161. Sarkodee-Adoo C, Gojo I, Heyman MR. Von Willebrand factor-cleaving protease in thrombotic thrombocytopenic purpura and the hemolytic-uremic syndrome [Letter]. *N Engl J Med* 1999;340:1368.

162. Loirat C, Sonsino E, Hinglais N, et al. Treatment of the childhood haemolytic uraemic syndrome with plasma. A multicenter randomized controlled trial. *Pediatr Nephrol* 1988;2:279–285.

163. Rizzoni G, Claris-Appiani A, Edetonti A, et al. Plasma infusion for hemolytic uremia syndrome in children: result of a multicenter controlled trial. *J Pediatr* 1988;112:284–290.

164. Egerman RS, Sibai BM. HELLP syndrome. *Clin Obstet Gynecol* 1999;42:381–389.

165. Stone JH, HELLP Syndrome: Hemolysis, elevated liver enzymes, and low platelets. *JAMA* 1998;280:559–562.

166. Fish R. The HELLP syndrome: case report and review of the literature. *J Emerg Med* 1993;11:169–174.

167. Sibai BM. The HELLP syndrome (hemolysis, elevated liver enzymes, and low platelets): much ado approximately nothing? *Am J Obstet Gynecol* 1990;162:311–316.

168. Saphier CF, Repke JT. Hemolysis, elevated liver enzymes, and low platelets (HELLP) syndrome: a review of diagnosis and management. *Semin Perinatol* 1998;22:118–133.

169. Martin JN Jr, Files JC, Blake PG, et al. Plasma exchange for preeclampsia. I. Postpartum use for persistently severe preeclampsia-eclampsia with HELLP syndrome. *Am J Obstet Gynecol* 1990;162:126–137.

170. Shulman NR, Aster Rh, Leitner A, et al. Immunoreactions involving platelets: V. Post-transfusion purpura due to a complement-fixing antibody against a genetically controlled platelet antigen. A proposed mechanism for thrombocytopenia and its relevance in "autoimmunity." *J Clin Invest* 1961;40:1597–1620.

171. Mueller-Eckhardt C. Post-transfusion purpura. *Br J Haematol* 1986;64:419–424.

172. McFarland JG. Posttransfusion purpura. In: Popovsky MA, ed. *Transfusion reactions*. Bethesda, MD: AABB Press, 1996:205–229.

173. Mueller-Eckhardt C, Kietel V. High-dose IgG for post-transfusion purpura-revisited. *Blut* 1988; 57:163–167.

174. Gerstner JB, Smith MJ, Davis KD, et al. Post-transfusion purpura: therapeutic failure of PI^{A1}-negative platelet transfusion. *Am J Hematol* 1979;6:71–75.

175. Brecher ME, Moore SB, Letendre L. Posttransfusion purpura: the therapeutic value of PI^{A1}-negative platelets. *Transfusion* 1990;30:433–435.

176. Win N, Matthey F, Slater NGP. Blood components-transfusion support in post-transfusion purpura due to HPA-1a immunization. *Vox Sang* 1996;71:191–193.

177. Godeau B, Fromont P, Bettaieb A, et al. Relapse of posttransfusion purpura after transfusion with frozen-thawed red cells [Letter]. *Transfusion* 1991;31:189–190.

178. Gabriel A, Lassnigg A, Kurz M, et al. Post-transfusion purpura due to HPA-1a immunization in a male patient: response to subsequent multiple HPA-1a - incompatible red-cell transfusions. *Transfus Med* 1995;5:131–134.

179. Kunicki TJ, Beardsley DS. The alloimmune thrombocytopenias: neonatal alloimmune thrombocytopenic purpura and post-transfusion purpura. *Prog Hemost Thromb* 1989;9:203–232.

180. Goldman M, Filion M, Proulx C, et al. Neonatal alloimmune thrombocytopenia. *Transfus Med Rev* 1994;8:123–131.

181. Letsky EA, Greaves M. Guidelines on the investigation and management of thrombocytopenia in pregnancy and neonatal alloimmune thrombocytopenia. *Br J Haematol* 1996;95:21–26.

182. Reznikoff-Etievant MF. Management of alloimmune neonatal and antenatal thrombocytopenia. *Vox Sang* 1988;55:192–201.

183. Win N, Ouwehand WH, Hurd C. Provision of platelets for severe neonatal alloimmune thrombocytopenia. *Br J Haematol* 1997;97:927–940.

184. Mueller-Eckhardt C, Kiefel V, Grubet A. High-dose IgG treatment for neonatal alloimmune thrombocytopenia. *Blut* 1989;49:145–146.

185. Vamvakas EC. Meta-analysis of randomized controlled trials of the efficacy of white cell reduction in preventing HLA-alloimmunization and refractoriness to random-donor platelet transfusions. *Transfus Med Rev* 1998;12:258–270.

186. O'Connell B, Lee EJ, Schiffer CA. The value of 10-minute posttransfusion platelet counts. *Transfusion* 1988;28:66–67.

187. Sniecinski I, O'Donnell MR, Nowicki B, et al. Prevention of refractoriness and HLA-alloimmunization using filtered blood products. *Blood* 1988;71:1402–1407.

188. Lee EJ, Schiffer CA. ABO compatibility can influence the results of platelet transfusion. Results of a randomized trial. *Transfusion* 1989;29:384–389.

189. Engelfriet CP, Reesink HW. Management of alloimmunized, refractory patients in need of platelet transfusions. *Vox Sang* 1997;73:191–198.

190. Yankee RA, Grumet FC, Rogentine GN. Platelet transfusion. The selection of compatible platelet donors for refractory patients by lymphocyte HL-A typing. *N Engl J Med* 1969;281:1208–1212.

191. Petz LD, Garratty G, Clark BD, et al. The effectiveness of an antibody specificity prediction (ASP) method for selection platelets for transfusion to alloimmunized patients. *Blood* 1995;86(Suppl1):546a(abst.).

192. Friedberg RC, Donnelly SF, Mintz PD. Independent roles for platelet crossmatching and HLA in the selection of platelets for alloimmunized patients. *Transfusion* 1994;34:215–220.

193. Rachel JM, Summers TC, Sinor LT, et al. Use of solid phase red cell adherence method for pretransfusion platelet compatibility testing. *Am J Clin Pathol* 1988;90:63–68.

194. Carvalho AC. Acquired platelet dysfunction in patients with uremia. *Hematol Oncol Clin North Am* 1990;4:129–143.

195. Remuzzi G. Bleeding in renal failure. *Lancet* 1988;1:1205–1208.

196. Escolar G, Cases A, Bastida E, et al. Uremic platelets have a functional defect affecting the interaction of von Willebrand factor with glycoprotein IIb-IIIa. *Blood* 1990;76:1336–1340.

197. Weigert AL, Schafer AI. Uremic bleeding: pathogenesis and therapy. *Am J Med Sci* 1998;316:94–104.

198. DiMinno G, Martinez J, McKeon ML, et al. Platelet dysfunction in uremia: multifaceted defect partially corrected by dialysis. *Am J Med* 1985;79:552–559.

199. Livio M, Marchesi D, Remuzzi G, et al. Uremic bleeding: role of anemia and beneficial effect of red cell transfusions. *Lancet* 1982;2:1013–1015.

200. Fernandez F, Goudable C, Sie P, et al. Low hematocrit and prolonged bleeding time in uremic patients: effect of red cell transfusions. *Br J Haematol* 1985;59:139–148.

201. Moia M, Mannucci PM, Vizzott L, et al. Improvement in the haemostatic effect of uraemia after treatment with recombinant human erythropoietin. *Lancet* 1987;2:1227–1229.

202. Eschbach JW, Egrie JC, Downing MR, et al. Correction of the anemia of end-stage renal disease with recombinant human erythropoietin. *N Engl J Med* 1987;316:73–78.

203. Mannucci PM, Remuzzi G, Pusineri F, et al. Deamino-8-D-arginine vasopressin shortens the bleeding time in uremia. *N Engl J Med* 1983;308:8–12.

204. Mannucci PM. Desmopressin (DDAVP) in the treatment of bleeding disorders: the first 20 years. *Blood* 1997;90:2515–21.

205. Janson PA, Jubelirer SJ, Weinstein MJ, et al. Treatment of the bleeding tendency in uremia with cryoprecipitate. *N Engl J Med* 1980;303:1318–1322.

206. Triulzi DJ, Blumberg N. Variability in response to cryoprecipitate treatment for hemostatic effects in uremia. *Yale J Biol Med* 1990;63:1–7.

207. Liu YK, Kosfeld RE, Marcum SG. Treatment of uremic bleeding with conjugated estrogen. *Lancet* 1984;2:887–890.

208. Livio M, Mannucci PM, Vigano G, et al. Conjugated estrogens for the management of bleeding associated with renal failure. *N Engl J Med* 1986;315:731–735.

209. Shemin D, Elnour M, Amarantes B, et al. Oral estrogens decrease bleeding time and improve clinical bleeding in patients with renal failure. *Am J Med* 1990;89:436–440.

210. Dloond JA, Schiff MJ. Beneficial effect of low-dose transdermal estrogen on bleeding time and clinical bleeding in uremia. *Am J Kidney Dis* 1995;26:22–26.

211. Roth GJ, Calverley DC. Aspirin, platelets, and thrombosis: theory and practice. *Blood* 1994;83:885–898.

212. Weiss HJ, Aledort LM, Kochwa S. The effect of salicylates on the hemostatic properties of platelets in man. *J Clin Invest* 1968;47:2169–2180.

213. Despotis GJ, Filos KS, Zoys TN, et al. Factors associated with excessive post-operative blood loss and hemostatic transfusion requirements: a multivariate analysis in cardiac surgical patients. *Anesth Analg* 1996;82:13–21.

214. Mannucci PM, Vicente V, Vianello L, et al. Controlled trial of desmopressin (DDAVP) in liver cirrhosis and other conditions associated with a prolonged bleeding time. *Blood* 1986; 67:1148–1153.

215. Lechner K, Niessner H, Thaler E. Coagulation abnormalities in liver disease. *Semin Thromb Hemost* 1977;4:40–56.

216. Mammen EF. Coagulation defects in liver disease. *Med Clin North Am* 1994;48:545–554.

217. Pereira SP, Langley PG, Williams CBE. The management of abnormalities of hemostasis in acute liver failure. *Semin Liver Dis* 1996;16:403–414.

218. Straub PW. Diffuse intravascular coagulation in liver disease. *Semin Thromb Hemost* 1977;4:29–439.

219. Burroughs AK, Matthews K, Qadiri M, et al. Desmopressin and bleeding time in patients with cirrhosis. *BMJ* 1985;291:1377–1381.

220. Defranchis F, Arcidiacono PG, Carpinelli PG, et al. Randomized controlled trial of desmopressin plus terlipressin and terlipressin alone for the treatment of acute variceal hemorrhage in cirrhotic patients: a multicenter, double-blind study. *Hepatology* 1993;18:1102–1107.

221. Levi M, Cate H. Disseminated intravascular coagulation. *N Engl J Med* 1999;341:562–592.

222. Rubin RN, Colman RW. Disseminated intravenous coagulation: approach to treatment. *Drugs* 1992;44:963–971.

223. Bick RL. Disseminated intravascular coagulation: objective clinical and laboratory diagnosis, treatment, and assessment of therapeutic response. *Semin Thromb Hemost* 1996;22:69–88.

224. Bick RL. Disseminated intravascular coagulation. *Hematol Oncol Clin North Am* 1992;6:1259–1285.

225. Richey ME. Management of disseminated intravascular coagulopathy. *Clin Obstet Gynecol* 1995;38:514–520.

226. Staudinger T, Locker GJ, Frass M. Management of acquired coagulation disorders in emergency and intensive care medicine. *Semin Thromb Hemost* 1996;22:93–104.

227. Baglin T. Disseminated intravascular coagulation and treatment. *BMJ* 1996;312:683–687.

228. Colman RW, Rubin RN. Disseminated intravascular coagulation due to malignancy. *Semin Oncol* 1990;17:172–186.

229. Feinstein DI. Treatment of intravascular coagulation. *Semin Hemost Thromb* 1988;14;351–362.

230. Bucur SZ, Levy JH, Despotis GJ, et al. Uses of anti-thrombin III concentrate in congenital and acquired deficiency states. *Transfusion* 1998;38:481–498.

231. de Jonge E, Levi M, Stoutenbeck P, et al. Current drug treatment strategies for disseminated intravascular coagulation. *Drugs* 1998;55:767–777.

232. Fourrier F, Chapin C, Huart JJ, et al. Double-blind, placebo-controlled trial of anti-thrombin III concentrates in septic shock with disseminated intravascular coagulation. *Chest* 1993;104:882–888.

233. Kearon C, Hirsh J. Management of anticoagulation before and after elective surgery. *N Engl J Med* 1997;336:1506–1511.

234. Humphries JE. Treatment of acquired disorders of hemostasis. In: Mintz PD, ed. *Transfusion therapy: clinical principles and practice*. Bethesda, MD: AABB Press:1999;129–149.

235. Weibert R, Le DT, Kayser S, et al. Correction of excessive anticoagulation with low-dose vitamin K_1. *Ann Intern Med* 1997;125:959–962.

236. Fredriksson K, Norrving B, Strömblad L-G. Emergency reversal of anticoagulation after intracerebral hemorrhage. *Stroke* 1992;23: 972–977.

237. Makris M, Greaves M, Phillips WS, et al. Emergency oral anticoagulant reversal: the relative efficacy of infusions of fresh frozen plasma and clotting factor concentrate on correction of the coagulopathy. *Thromb Haemost* 1997;77:477–480.

238. Ansell J, Hirsh J, Dalen J, et al. Managing oral anticoagulant therapy. *Chest* 2001;119:225–385.

239. Hirsh J, Raschke R, Warkentin TE, et al. Heparin: mechanism of action, pharmacokinetics, dosing considerations, monitoring, efficacy, and safety. Chest 1995(Suppl 4);108:258S–275S.

240. Sabbagh AH, Chung GK, Shuttleworth P, et al. Fresh frozen plasma: a solution to heparin resistance during cardiopulmonary bypass. *Ann Thorac Surg* 1984;37:466–468.

241. Arén C. Heparin and protamine therapy. *Semin Thorac Cardiovasc Surg* 1990;2:364–372.

242. Lane D, Denton J, Flynn A, et al. Anticoagulant activities of heparin oligosaccharides and their neutralization by platelet factor 4. *Biochem J* 1984;218:725–732.

243. Massonnet-Castel S, Pelissier E, Bara L, et al. Partial reversal of low molecular weight heparin (PK 10169) anti-Xa activity by protamine sulfate: in vitro and in vivo study during cardiac surgery with extracorporeal circulation. *Haemostasis* 1986;16:139–146.

244. Schulman S, Johnsson H. Heparin, DDAVP and the bleeding time. *Thromb Haemost* 1991;65:242–244.

245. Teasdale S, Zulys V, Mycyk T, et al. Ancrod anticoagulation for cardiopulmonary bypass in heparin-induced thrombocytopenia and thrombus. *Ann Thorac Surg* 1989;48:712–713.

246. Cole C, Bormanis J, Luna G, et al. Ancrod versus heparin for anticoagulation during vascular surgical procedures. *J Vasc Surg* 1993;17:288–293.

247. Seligsohn U. Factor XI deficiency. *Thromb Haemost* 1993;70:68–71.

248. Rodgers GM, Greenberg CS. Inherited coagulation disorders. In: Lee GR, Foerster J, Lukens J, et al., eds. *Wintrobe's clinical hematology*, 10th ed. Baltimore: Williams & Wilkins, 1999:1682–1732.

249. Kitchens CS, Newcomb TF. Factor XIII. *Medicine* 1979;58:413–429.

250. Amris CJ, Hilden M. Treatment of factor XIII deficiency with cryoprecipitate. *Thromb Diath Haemorrh* 1968;20:528–533.

251. Fisher S, Rikover M, Nady S. Factor 13 deficiency with severe hemorrhagic diathesis. *Blood* 1966;28:34–39.

252. Ewenstein BM. Von Willebrand's disease. *Annu Rev Med* 1997;48:525–542.

253. Aledort LM. von Willebrand disease: from the bedside to therapy. *Thromb Haemost* 1997;78:562–565.

254. Nichols WC, Ginsburg D. Von Willebrand disease. *Medicine* 1997;76:1–20.

255. Phillips MD, Santhouse A. Von Willebrand disease: recent advances in pathophysiology and treatment. *Am J Med Sci* 1998;316:77–86.

256. Mannucci PM. Treatment of von Willebrand's disease. *J Intern Med* 1997;242(Suppl 740):129–132.

257. McKeown LP, Connaghan G, Wilson O, et al. 1-Desamino-8-arginine-vasopressin corrects the hemostatic defects in type 2B von Willebrand's disease. *Am J Hematol* 1996;51:158–163.

258. Mannucci PM, Lusher JM. Desmopressin and thrombosis. *Lancet* 1989;2:675–677.

259. Menache D, Aronson DL. New treatments of von Willebrand disease: plasma derived von Willebrand factor concentrates. *Thromb Haemost* 1997;78:566–570.

260. Perkins HA. Correction of the hemostatic defects in von Willebrand disease. *Blood* 1967;30:376–380.

261. Castillo R, Monteagudo J, Escolar G, et al. Hemostatic effect of normal platelet transfusion in severe von Willebrand disease patients. *Blood* 1991;77:1901–1905.

262. Castillo R, Escolar G, Monteagudo J, et al. Hemostasis in patients with severe von Willebrand disease improves after normal platelet transfusion and normalizes with further correction of the plasma defect. *Transfusion* 1997;37:785–790.

263. Tefferi A, Nichols WL. Acquired von Willebrand disease: concise review of occurrence, diagnosis, pathogenesis, and treatment. *Am J Med* 1997;103:536–540.

264. Federici AB, Stabile F, Castaman G, et al. Treatment of acquired von Willebrand syndrome in patients with monoclonal gammopathy of uncertain significance: comparison of three different therapeutic approaches. *Blood* 1998;92:2707–2711.

265. Gunter K, Luban N. Transfusion-transmitted cytomegalovirus and Epstein-Barr virus diseases. In: Rossi EC, Simon TL, Moss GL, et al., eds. *Principles of transfusion medicine*, 2nd ed. Baltimore: Williams & Wilkins, 1996:717–732.

266. *Leukocyte reduction for the prevention of transfusion-transmitted cytomegalovirus (TT-CMV)*. Association Bulletin 97-2. Bethesda, MD: American Association of Blood Banks Bulletin, 1997.

267. Bowden RA, Slichter SJ, Sayers M, et al. A comparison of filtered leukocyte-reduced and cytomegalovirus (CMV) seronegative blood products for the prevention of transfusion-associated CMV infection after marrow transplant. *Blood* 1995;86:3598–3603.

268. Miller WJ, McCullough J, Balfour HH, et al. Prevention of cytomegalovirus infection following bone marrow transplantation: a randomized trial of blood product screening. *Bone Marrow Transplant* 1991;7:227–234.

269. Pamphilon DH. Prevention of transfusion transmitted cytomegalovirus infection. *Transfus Med* 1999;9:115–123.

270. Friedberg RC. Transfusion therapy in hematopoietic stem cell transplantation, In: Mintz PD, ed. *Transfusion therapy: clinical principles and practice*. Bethesda, MD: AABB, 1999:233–249.

271. Przepiorka D, Leparc GF, Stovall MA, et al. Use of irradiated blood components: practice parameter. *Am J Clin Pathol* 1996;106:6–11.

272. Linden JV, Pisciotto PT. Transfusion-associated graft-versus-host disease and blood irradiation. *Transfus Med Rev* 1992;6:116–123.

273. Lasky LC, Warkentin PI, Kersey JH, et al. Hemotherapy in patients undergoing blood group incompatible bone marrow transplantation. *Transfusion* 1983;23:277–285.

274. Warkentin PI, Yomtovian R, Hurd D, et al. Severe hemolytic transfusion reaction complicating an ABO-incompatible bone marrow transplantation. *Vox Sang* 1983;45:40–47.

275. Bernestin SH, Nademanee AP, Vose JM, et al. A multicenter study of platelet recovery and utilization in patients after myeloablative therapy and hematopoietic stem cell transplantation. *Blood* 1998;91:3509–3517.

276. Shanwell A, Ringden O, Wiechel B, et al. A study of the effect of ABO incompatible plasma in platelet concentrates transfused to bone marrow transplant recipients. *Vox Sang* 1991;60:23–27.

277. Mair B, Benson K. Evaluation of changes in hemoglobin levels associated with ABO-incompatible plasma in apheresis platelets. *Transfusion* 1998;38:51–55.

278. Petz LD. Hemolysis associated with transplantation. *Transfusion* 1998;38:224–228.

279. Sniecinski IJ, Petz OL, Blume KG. Immunohematologic consequences of major ABO-mismatched bone marrow transplantation. *Transplantation* 1998;45:530–534.

280. Sazama K. Practical issues in informed consent for transfusion. *Am J Clin Pathol* 1997;107:S72–S74.

281. Holland PV. Consent for transfusion: is it informed? *Transfus Med Rev* 1997;11:274–275.

282. Eisenstaedt RS, Glanz K, Smith DG, et al. Informed consent for blood transfusion: a regional hospital survey. *Transfusion* 1993;33:558–561.

283. Goodnough LT, Brecher ME, Kanter MH, et al. Transfusion medicine (first of two parts). Blood transfusion. *N Engl J Med* 1999;340:438–447.

284. Lumadue JA, Boyd JS, Ness PM. Adherence to strict specimen-labeling policy decreases the incidence of erroneous blood grouping of blood bank specimens. *Transfusion* 1997;37:1169–1172.

285. Wenz B, Mercuriali F, AuBuchon JP. Practical methods to improve transfusion safety by using novel blood unit and patient identification systems. *Am J Clin Pathol* 1997;107(Suppl 1):S12–S16.

286. AuBuchon JP, Littenberg B. A cost-effectiveness analysis of the use of a mechanical barrier system to reduce the risk of mistransfusion. *Transfusion* 1996;36:222–226.

287. Jensen NJ, Crosson JT. An automated system for bedside verification of the match between patient identification and blood unit identification. *Transfusion* 1996;36:216–221.

288. Guy J, Haley K, Zuspan SJ, et al. Use of intraosseous infusion in the pediatric trauma patient. *J Pediatr Surg* 1993;28:158–161.

289. de la Roche MR, Gauthier L. Rapid transfusion of packed red blood cells: effects of dilution, pressure and catheter size. *Ann Emerg Med* 1993;22:1551–1555.

290. Hodge D 3rd, Fleisher G. Pediatric catheter flow rates. *Am J Emerg Med* 1985;3:403–407.

291. Ryden SE, Oberman HA. Compatibility of common intravenous solutions with CPD blood. *Transfusion* 1975;3:250–255.

292. Calhoun L. Blood product preparation and administration. In: Petz LD, Swisher SN, Kleinman S, Spence RK, et al., eds. *Clinical practice of transfusion medicine*. New York: Churchill-Livingstone, 1996:305–333.

293. Snyder EL, Bookbinder M. Role of microaggregate blood filtration in clinical medicine. *Transfusion* 1983;23:460–470.

294. Mosely RV, Doty DB. Death associated with multiple pulmonary emboli soon after battle injury. *Ann Surg* 1970;171:336–346.

295. McNamara JJ, Molot MD, Stremple JR. Screen filtration pressure in combat casualties. *Ann Surg* 1970;172:334–341.

296. Simmons RL, Heisterkamp CA 3rd, Collins JA, et al. Respiratory insufficiency in combat casualties. IV. Hypoxemia during convalescence. *Ann Surg* 1969;170:53–62.

297. Collins JA, James PM, Bredenberg CE, et al. The relationship between transfusion and hypoxemia in combat casualties. *Ann Surg* 1978;188:513–520.

298. Schned AR, Silver H. The use of microaggregate filtration in the prevention of febrile transfusion reactions. *Transfusion* 1981;21:675–681.

299. Wenz B. Microaggregate blood filtration and the febrile transfusion reaction. *Transfusion* 1983;23:95–98.

300. Parravicini A, Rebulla P, Apuzzo J, et al. The preparation of leukocyte-poor red cells for transfusion by a simple cost-effective technique. *Transfusion* 1984;24:508–509.

301. Bordin JO, Heddle NM, Blajchman MA, et al. Biological effects of leukocytes present in transfused cellular blood products. *Blood* 1994;84:1703–1721.

302. Dzik S. Leukodepletion blood filters: filter design and mechanisms of leukocyte removal. *Transfus Med Rev* 1993;7:65–77.

303. Stack G, Snyder EL. Cytokine generation in stored platelet concentrates. *Transfusion* 1994;34:20–25.

304. Heddle NM, Klama L, Singer J, et al. The role of the plasma from platelet concentrates in transfusion reactions. *N Engl J Med* 1994;331:625–628.

305. Mangano MM, Chambers LA, Kruskall MS. Limited efficiency of leuko-poor platelets for prevention of febrile transfusion reactions. *Am J Clin Pathol* 1991;95:733–738.

306. Tan KK, Lee WS, Liaw LC, Oh A. A prospective study on the use of leukocyte-filters in reducing blood transfusion reactions in multi-transfused thalassemic children. *Singapore Med J* 1993;34:109–111.

307. Reverberi R, Ferrari L, Gennari M, et al. Prevention of non-hemolytic transfusion reactions with leukocyte-poor blood: a prospective study. *Haematologica* 1989;74:283–288.

308. Goodnough LT, Riddell J, Lazarus H, et al. Prevalence of platelet transfusion reactions before and after implementation of leukocyte depleted platelet concentrates by filtration. *Vox Sang* 1993;65:103–107.

309. Murphy MR. New variant Creutzfeldt-Jakob disease (nv CJD): the risk of transmission by blood transfusion and the potential benefit of leukocyte-reduction of blood components. *Transfus Med Rev* 1999;13:75–83.

310. Brown P, Rohwer RG, Dunstan BC, et al. The distribution of infectivity in blood components and plasma derivatives in experimental models of transmissible spongiform encephalopathy. *Transfusion* 1998;38:810–16.

311. Klein MA, Frigg R, Flechsig E, et al. A crucial role for B cells in neuroinvasive scrapie. *Nature* 1997;390:687–690.

312. Hamill TR. The 30-minute rule for reissuing blood: are we needlessly discarding units? *Transfusion* 1990;30:58–62.

313. Cummings E. Platelet rotators, infusion pumps and blood warmers. *Transfus Sci* 1989;10:199–206.

314. Snyder EL, Rinder HM, Napychank P. *In vitro* and *in vivo* evaluation of platelet transfusions administered through an electromechanical infusion pump. *Am J Clin Pathol* 1990;94:77–80.

315. Snyder EL, Malech HL, Ferri PM, et al. *In vitro* function of granulocyte concentrates following passage through an electromechanical infusion pump. *Transfusion* 1986;26:141–144.

316. Staples PJ, Griner PF. Extracorporeal hemolysis of blood in a microwave blood warmer. *N Engl J Med* 1971;285:317–319.

317. Linko K, Hynynen K. Erythrocyte damage caused by the Haemotherm microwave blood warmer. *Acta Anaesth Scand* 1979;23:320–328.

318. Kruskall MS, Pacini DG, Malynn ER, et al. Evaluation of a blood warmer that utilizes a 40°C heat exchanger. *Transfusion* 1990;30:7–10.

319. Evans CS. Home and outpatient transfusion programs. In: Petz LD, Swisher SN, Kleinman S, et al., eds. *Clinical practice of transfusion medicine*, 3rd ed. New York: Churchill-Livingstone, 1996:295–304.

320. Snyder E, Menitove J, eds. *Home transfusion therapy*. Arlington, VA: American Association of Blood Banks, 1986:61–69.

321. Fridey JL. Practical aspects of out-of-hospital transfusion. *Am J Clin Pathol* 1997;107(Suppl 1):S64–S71.

322. Crocker KS, Coker MH. Initiation of home hemotherapy program using a primary nursing model. *J Intravenous Nurs* 1990;13:13–19.

73

Neonatal Transfusion

Jayashree Ramasethu

Naomi L. C. Luban

Transfusion support of the neonatal population requires an understanding of the dynamic interactions of the feto-maternal unit and the physiologic changes that accompany the transition from fetus to neonate and neonate to infant. The small blood volume, immunological immaturity, and compromised organ function resulting from either prematurity, illness, or both, provide little room for error, and unique challenges to the transfusion medicine specialist. Survivors of neonatal intensive care are more likely to suffer the long-term consequences of transfusions, unlike adults who, because of underlying disease, have a posttransfusion mortality of 24% within 1 year of transfusion and 52% within 10 years (1). Blood products utilized in neonates include packed red blood cells (RBC), platelet concentrates, granulocyte concentrates, fresh frozen plasma (FFP) and cryoprecipitate. These products may be transfused as simple small volume transfusions or as part of exchange transfusions (ETs). The indications, requirements, and transfusion techniques differ for each procedure and component (Table 73.1). Most infants who weigh less than 1,000 g at birth receive one or more RBC transfusions, especially in the first few weeks of life (2). Intensive blood bank support is necessary for neonates undergoing extracorporeal membrane oxygenation (ECMO) and cardiopulmonary bypass. Perioperative blood loss and blood product requirement, particularly for platelet concentrates and FFP, is disproportionately greater for neonates following open-heart surgery, than for any other age group (3, 4).

TABLE 73.1. FORMULA FOR TRANSFUSION OF RED CELL PRODUCTS

I. Blood Volumes
Preterm [all equal to] 100–120 ml/kg
Full Term [all equal to] 80–85 ml/kg
Adult [all equal to] 70–75 ml/kg

II. Calculation of volume of packed red blood cell (RBC) transfusion

$$\text{PRBC volume} = \frac{\text{EBV} \times (\text{Hct desired} - \text{Hct observed})}{\text{Hct of packed red-cell unit}}$$

$$\text{EBV} = \text{estimated patients blood volume}$$

III. Calculation of blood needed for exchange transfusion

Severe anemia Exchange volume of PRBC

$$= \frac{\text{EBV} \times \text{desired Hb rise}}{\text{Hb of PRBC} - \text{Hbw}}$$

$$\text{Hbw} = (\text{initial Hb} + \text{desired Hb}) / 2$$

Polycythemia Exchange volume of plasma / 5% albumin /saline

$$= \frac{\text{EBV} \times \text{desired Hct change}}{\text{starting Hct}}$$

IV. Formula for reconstitution of blood to a specific hematocrit for exchange transfusion

$$\text{Volume of packed cells} = \frac{\text{total exchange volume}}{\text{Hct of PRBC}} \times \text{desired Hct of new unit}$$

$$\text{Volume of FFP} = \text{total exchange volume} - \text{volume of packed cells}$$

SMALL VOLUME RED BLOOD CELL TRANSFUSIONS

Hemoglobin concentrations decrease in all newborn infants in the weeks after birth, reaching a nadir of 9.5 to 11 g/dL at 10 to 12 weeks of age in healthy term infants; this phenomenon is termed physiological anemia of infancy. In addition, anemia in the neonatal period may be secondary to blood loss, hemolysis, or impaired production of RBCs (Table 73.2). In preterm infants, an exaggeration of physiological anemia, called anemia of prematurity, results in an earlier and more pronounced decline in hemoglobin levels, which drop as low as 7 g/dL by 4 to 8 weeks of age (5). Anemia of prematurity occurs despite adequate numbers of erythroid progenitors, primarily because of delayed and diminished production of erythropoietin, combined with rapid postnatal expansion of plasma volume and decreased RBC survival. In sick newborn infants requiring intensive care, frequent blood sampling for laboratory tests and monitoring of blood gases overwhelms compensatory mechanisms, resulting in significant iatrogenic anemia, which increases the need for transfusions (6, 7).

In the 1980s, 80% of very low birth weight (VLBW, birth weight ≤1,500 g) infants received multiple transfusions, often from different donors (8). Over the last decade, a concerted effort has been under way to minimize the potential risks of transfusions by reducing the number of transfusions and donor exposures. A retrospective study of transfusion practices in VLBW infants admitted to one tertiary level nursery found that the mean number of transfusions per infant declined from seven in 1982, to five in 1989, to 2.3 in 1993, prior to the use of recombinant erythropoietin (2). Improvements in clinical care, laboratory micromethods, noninvasive monitoring techniques, and the adoption of conservative transfusion guidelines all have played a role in decreasing transfusions. Portable clinical analyzers that use only a fraction of the quantity of blood required for standard micromethods significantly reduce iatrogenic blood loss (9).

Sampling losses and therefore transfusion requirements may be related to the degree of prematurity and severity of illness (10). However, even when adjusted for these factors, considerable variation in transfusion practice is found in different centers, with a significant impact on the number and volume of RBC transfusions, suggesting that the utilization of small volume RBC transfusions often is discretionary (11,12 and 13).

The hemoglobin levels at which blood transfusion would be indicated and the definition of “nonphysiologic” versus “physiologic anemia” in preterm neonates is fraught with controversy. Tissue oxygen delivery is determined not only by the hemoglobin concentration, but also by the proportion of fetal to adult hemoglobin, level of RBC 2,3-diphosphoglycerate (2,3 DPG), cardiac output, and arterial oxygen tension. Fetal hemoglobin, the predominant hemoglobin at birth, has the characteristic of high oxygen affinity making it ideal for uteroplacental oxygen transport to the fetus, but its poor peripheral oxygen unloading may decrease oxygen availability to the hypoxic neonate. The physiological indicators of significant anemia include compensatory responses such as tachypnea and tachycardia, and the consequences of impaired oxygen transport manifested by poor growth, periodic breathing or apnea, decreased

activity, lethargy, poor feeding, inadequate reserve to respond to a stressor associated with increased oxygen demand, and anaerobic metabolism resulting in lactic acidosis (14). However, many of these symptoms are nonspecific and do not correlate well to hemoglobin levels or respond consistently to RBC transfusions (15).

The ability to summarize clear-cut indications for the simple transfusion of infants is hampered by the paucity of adequate randomized clinical studies and lack of acceptable laboratory parameters. Reevaluation of commonly accepted indications and the absence of clear evidence of benefit of RBC transfusions in certain situations has led to the generation of increasingly stringent, although largely empirically derived guidelines (Table 73.3) (16,17,18 and 19). The U.S. Multicenter Recombinant Human Erythropoietin trial conducted between 1991 and 1993 demonstrated that there was no adverse consequence of conservative transfusion practices in VLBW infants, including those on moderate ventilatory support or supplementary oxygen (16). However, infants with severe cardiorespiratory compromise were excluded from the trial. Definitions of significant apnea necessitating transfusion, the hematocrit levels at which transfusion is indicated for infants needing significant support for cardiorespiratory failure, and the requirement for replacing iatrogenic blood loss differ between published guidelines. Guidelines published by British authorities suggest consideration of RBC transfusions for asymptomatic newborns with hemoglobin levels less than 10.5 g/dL and support the replacement of phlebotomy losses by blood transfusion (19).

TABLE 73.2. ETIOLOGY OF ANEMIA IN THE NEONATE

Blood Loss

- Obstetric causes: abruptio placentae, placenta previa, rupture of vasa previa, incision of placenta at C section
- Feto-maternal hemorrhage
- Twin-to-twin transfusion
- Neonatal bleeding: massive subgaleal or intracranial hemorrhage, rupture of liver or spleen, gastrointestinal hemorrhage
- Phlebotomy

Impaired RBC Production

- Anemia of prematurity
- Congenital infections- Parvovirus B19, CMV, Rubella, Syphilis
- Diamond-Blackfan syndrome
- Congenital leukaemia
- Osteopetrosis

Hemolysis

- Immune hemolysis: Rh, ABO, Kell, Duffy etc., maternal lupus
- Hereditary RBC membrane defects: spherocytosis, elliptocytosis
- Hereditary RBC enzyme defects: G-6-PD, pyruvate kinase deficiency
- Hemoglobinopathies: alpha thalassemia syndromes
- Acquired hemolysis: DIC, sepsis
- Macro- or microangiopathic anemia: cavernous hemangioma, large vessel thrombi
- Vitamin E deficiency

TABLE 73.3. GUIDELINES FOR RED BLOOD CELL TRANSFUSIONS IN NEONATES

	US rHuEPO[16]	CAP[17]
Asymptomatic	Hct ≤ 20% retics <100,000/μL	Hct ≤ 20% retics <100,000/μL
Weight gain <10 g/day while receiving ≥100 kcal/kg/day	Hct ≤ 30%	Hct ≤ 25%
Significant apnea or bradycardia	Hct ≤ 30%	Hct ≤ 25%
Tachypnea or tachycardia	Hct ≤ 30%	Hct ≤ 25%
<35% hood oxygen	Hct ≤ 30%	Hct ≤ 25%
CPAP/IMV with Paw <6 cms H_2O		
>35% hood oxygen	Hct ≤ 35%	Hct ≤ 30%
CPAP/IMV with Paw ≥ 6–8 cms H_2O	Hct ≤ 35%	Hct ≤ 30%
IMV with Paw > 8 cms H_2O	not in trial	Hct ≤ 35%
Severe congenital heart disease	not in trial	Hct ≤ 35%
Surgery	Hct ≤ 30%	No recommendation
Phlebotomy loss alone	No transfusion	No transfusion

CPAP, continuous positive airway pressure; IMV, intermittent mechanical ventilation; Paw, mean airway pressure; Hct, hematocrit; Retic, absolute reticulocyte count.

Pretransfusion Testing

ABO groups and Rh types should be determined on samples obtained from both mother and baby. Cord blood may be used for initial testing. The baby's blood group is determined from the red cells alone because the corresponding isoagglutinins anti-A and anti-B in the serum usually are weak or absent (20). Maternal serum should be screened for the presence of atypical antibodies. If maternal blood is unavailable, the neonatal sample should be screened to exclude atypical antibodies. A conventional crossmatch is unnecessary if there are no atypical antibodies demonstrable. Because the formation of alloantibodies in the first 4 months of life is extremely rare, further compatibility testing for repeated small-volume transfusions can be omitted. Compatible blood may be low anti-A, B titer Group O Rh-negative blood or blood of the infant's ABO group and Rh type. If the antibody screen is positive, serological investigation to identify the antibody(ies) is necessary. Full compatibility testing is performed with blood selected to be negative for the antigens to which the antibody is directed.

Anticoagulant-Preservative Solutions

RBCs are stored in one of several anticoagulant-preservative (A-P) solutions to improve red-cell viability and extend storage for up to 42 days. The preservative solutions produce an increase in intraerythrocytic ATP, which is reflected in maintenance of red-cell membrane integrity and contributes to improved *in vivo* survival. The hematocrit of the RBCs in preservative solutions is approximately 60%, producing a less viscous product with a faster flow rate than standard packed RBCs, which usually have a hematocrit of 75 ± 5%. However, the osmotic load from the glucose, and the presence of high concentrations of mannitol and sodium in A-P solutions have raised questions about metabolic and physiologic effects in neonates (21).

Several studies have addressed the use of additive solutions in small-volume transfusions in neonates (22,23,24 and 25). In a randomized, controlled trial in 29 neonates, Strauss et al. (23) found no

significant differences in serum potassium, sodium, or osmolality in neonates transfused ADSOL (Fenwal Division, Baxter Healthcare Corp., Deerfield, IL) units stored to outdate, as compared to neonates transfused with CPDA1 units less than 7 days old. The ADSOL units were packed to remove additive with a resultant hematocrit of 85% to 87%; 47% of the 124 ADSOL transfusions were more than 14 days old. Weekly analysis of infant 2,3 DPG concentrations demonstrated similar levels in infants transfused with either CPDA1 or ADSOL. Goodstein et al. (24) studied 12 premature infants (BW <1,250 g), each of whom received 17 mL/kg transfusions of ADSOL preserved blood packed to a hematocrit of 63%. A battery of pre- and posttransfusion tests revealed no clinically relevant biochemical abnormalities, although increases in ammonia and LDH levels were noted, as well as small decreases in bicarbonate concentration, despite maintenance of pH.

For massive transfusions such as ECMO, ET, or cardiopulmonary bypass, the safety and efficacy of additive solutions have not been studied with the same scrutiny. Measurements of the various constituents of additives in packed cells stored in CPDA1, CP2D, and AS-3 over time have been published recently (26). Theoretical calculations, based on the constituents of the additives, indicate that hyperosmolality, hyperglycemia, hypernatremia, hyperkalemia, and hyperphosphatemia are likely to occur with massive transfusions in neonates (21, 26). Given the risk of toxicity, RBCs stored in additive solutions should be avoided for massive transfusions, or the additive solutions should be removed by centrifugation or by washing.

Donor Exposure

Traditional use of relatively fresh blood (<7 days old) for infants resulted in exposure to 6 to 10 different donors in small premature infants requiring multiple transfusions (27). The desire to reduce the potential risk of acquiring transfusion-transmitted diseases from multiple donor exposures has led to innovations in blood banking practices. Several techniques may be employed to prepare blood for neonates with the goal of decreasing donor exposure. Sterile connecting devices for repeatedly obtaining small aliquots of RBCs from dedicated units, blood bags with multiple satellite bags, and the use of RBCs collected in A-P solutions until the expiration date of 35 to 42 days reduce the number of donor exposures in neonates requiring multiple small transfusions, without undue wastage of blood (27,28,29 and 30). Increased K+ in older, stored blood does not appear to pose significant risks in neonates receiving small-volume transfusions of 10 to 15 mL/kg. Storage of RBCs over time causes depletion of red-cell 2,3 DPG with possible effects on oxygen dissociation (31), but the rapid regeneration of 2,3 DPG in transfused cells makes this irrelevant for small-volume transfusions.

Recombinant Erythropoietin

The benefit of exogenous erythropoietin to stimulate red-blood-cell production and overcome anemia of prematurity still is being debated, despite numerous multicenter trials (32, 33). Randomized, controlled trials in the United States, Europe, and South Africa demonstrated that treatment with recombinant human erythropoietin (rHuEpo) was associated with modest reductions in the number and volume of RBC transfusions in VLBW preterm infants, with maximum benefit being seen in larger, more stable preterm infants (16, 34, 35). However, these trials excluded infants with significant cardiorespiratory compromise, the very infants likely to experience the greatest phlebotomy losses and require the most transfusions. One trial on a small number of extremely low birth weight infants (birth weight <750 g) who were mechanically ventilated showed that early administration of erythropoietin and parenteral iron resulted in fewer transfusions in the first 3 weeks of life (36). The study again was limited by the small number of patients and the exclusion of extremely ill infants, and was too brief in duration to determine if the benefits of early erythropoietin therapy lasted through the hospitalization of these fragile infants. Although rHuEpo has not been shown to avert transfusions completely, it still shows promise as an adjuvant agent for reducing RBC transfusions in selected infants.

Whole Blood Transfusion

Whole-blood or reconstituted whole blood, prepared by adding a unit of RBCs to a compatible unit of FFP is the product of choice in the setting of massive transfusion or acute blood loss, where restoration of oxygen carrying capacity and blood volume are needed simultaneously. Often, the acute resuscitation of the patient is carried out with crystalloid or colloid solutions in conjunction with packed RBCs. Uncrossmatched O Rh-negative blood may be required for transfusion in infants with shock secondary to acute blood loss in emergencies such as splenic rupture, massive subgaleal hemorrhage, or ruptured *vasa previa*. Other blood components may need to be transfused subsequently, because the levels of coagulation factors V and VIII are low and platelet function is grossly abnormal in stored whole blood. In children under the age of 2 years, transfusion of fresh (<48 hours old) whole blood has been associated with significantly less postoperative blood loss following bypass surgery for complex congenital heart disease, than the transfusion of multiple blood components; this has been attributed to better platelet function in fresh blood (37). Other indications for transfusion of whole blood or reconstituted whole blood in the newborn include ETs, cardiopulmonary bypass, extracorporeal membrane oxygenation (ECMO), and continuous hemofiltration. The mechanics, blood product choice, and complications associated with ECMO and cardiopulmonary bypass have been discussed in detail (38).

Directed Donor Transfusions

Directed donations from first- and second-degree relatives rather than from unknown blood donors are perceived by the lay public to have a lower risk of transmitting viral infections, although there is no scientific data to support this contention. The transfusion of blood from biologic parents poses unique immunologic risks to the neonate. Transfusion-associated graft-versus-host disease (TAGVHD), a well-recognized complication of use of familial blood, is described later in this chapter. Maternal plasma may contain antibodies directed against paternal red cell, leukocyte, and platelet antigens that also may be expressed on neonatal cells (39). Transfusion of maternal blood components containing these antibodies theoretically could result in hemolysis, thrombocytopenia, or pulmonary reactions, because current pretransfusion testing may not detect these incompatibilities. Paternal red cells, WBCs, and platelets may express antigens that are not inherited by the infant, but to which the mother may have produced an IgG antibody that crosses the placenta. Transfusion of paternal cells to the infant could interact with transplacentally transmitted maternal IgG antibody in the infant's plasma. Although these dangers are theoretical, it is recommended that mothers should not provide blood components containing plasma; maternal red cells and platelets should be given as washed concentrates. Fathers and paternal blood relatives should not serve as donors for blood components containing cellular elements; a full antiglobulin crossmatch should be performed to detect red-cell incompatibilities if their use is unavoidable. In addition, all blood components obtained from biologic parents and siblings should be irradiated prior to transfusion (40).

Autologous Transfusion

The placenta contains 75 to 125 mL of blood at birth, a substantial volume of fetal blood that could be an ideal source of autologous blood for the neonate, eliminating the potential risks of transfusion-transmitted diseases and TAGVHD. The simple measure of delayed cord clamping at birth has a significant impact on red-cell mass, hematocrit, and transfusion requirement in preterm infants (41), but also may lead to circulatory overload (42). Heparinized placental blood has been used in the delivery room resuscitation of term infants with shock and profound anemia (43, 44), and placental blood stored in anticoagulant media has been used successfully for neonatal transfusion, albeit infrequently (45). Protocols for ensuring proper collection without bacterial contamination and adequate anticoagulation still are being refined (46). Although *in vitro* storage characteristics of placental blood stored for 28 days in anticoagulant media are reported to be similar to adult blood, there is no information about the posttransfusion red-cell survival and other biochemical parameters of transfused autologous placental blood (47).

ALLOIMMUNE HEMOLYTIC DISEASE OF THE NEWBORN

Part of "73 - Neonatal Transfusion"

Alloimmune hemolytic disease of the newborn (HDN) occurs as a result of sensitization of the mother's immune system to fetal red-cell antigens foreign to the mother, inherited by the fetus from the father. This sensitization results in the transplacental transfer of maternal IgG antibodies that bind to the fetal red cells causing hemolysis, and as a consequence of the hemolysis, varying degrees of anemia, extramedullary hematopoiesis, and neonatal hyperbilirubinemia. Anemia may range from mild to severe resulting in fetal or neonatal hydrops, associated with very high mortality. Neonatal jaundice secondary to hemolysis may range from minor exaggeration of physiological jaundice to extreme hyperbilirubinemia, with the risk of bilirubin encephalopathy. This disorder, also called kernicterus, is from neuronal necrosis in the basal ganglia and cerebellum, caused by the deposition of lipid-soluble free bilirubin. Kernicterus may result in death or disability characterized by choreoathetoid cerebral palsy and hearing loss.

The prevalence and severity of HDN varies widely in different races and ethnic groups, because of differences in the prevalence of RBC antigen systems. For example, 15% of Caucasians are Rh negative, compared to 7% to 8% of African-Americans, 5% of Asian Indians, and 0.3% of Chinese. Although the list of antibodies reported to cause HDN include IgG specific for virtually any known red-cell antigen, the most frequent causes of HDN are anti-D, ABO, non-D Rh antibodies (c, C, e, E, cc, and Ce) and those belonging to the Kell, Duffy, Kidd, and MNS systems (48). Transfusion of blood compatible with not only the D-antigen, but also Kell and other Rh antigens has been advocated for premenopausal women to prevent alloimmunization.

ABO HDN is limited to infants of blood type A or B born to blood group O mothers. Although ABO incompatibility exists in 15% of O-group pregnancies, ABO HDN is estimated to occur

only in 3% of births. This is attributed to the fact that most naturally occurring anti-A and anti-B antibodies are of the IgM type and do not cross the placenta; only a small proportion of group O women produce anti-A or anti-B IgG antibodies capable of transplacental transfer. Unlike Rh HDN, ABO HDN occurs with equal frequency in first or subsequent pregnancies. A higher incidence and greater severity is reported in Southeast Asians (49), Latin Americans (50), Arabs, South African, and African-American populations. The most common presentation is neonatal jaundice and anemia requiring phototherapy and small-volume RBC transfusion; ETs are rarely indicated. Severe fetal anemia and hydrops are rare (51).

Sensitization in Rh-D negative women usually is secondary to asymptomatic transplacental passage of D-positive fetal red cells and rarely from exposure to Rh-positive blood by transfusion or by the sharing of contaminated needles. Fetomaternal hemorrhage (FMH) occurs in 75% of pregnancies, increases with gestation, and is maximal at the time of delivery (52). Although the volume is less than 1 mL in 96% of women, FMH of more than 30 mL may occur in 1% of pregnancies. In the absence of Rh immunoprophylaxis, sensitization occurs in 7% to 16% of women at risk with delivery of an ABO-compatible infant, and in 2% after delivery of an ABO-incompatible infant. Fetomaternal ABO incompatibility confers a degree of protection against primary Rh sensitization because incompatible cells are destroyed rapidly by maternal anti-A and anti-B antibodies, reducing maternal exposure to Rh-D antigenic sites. Once sensitization has occurred, ABO incompatibility offers no protection against the brisk secondary immune response that occurs with further exposure to Rh-D–positive fetal cells. Obstetrical procedures such as chorionic villus sampling, amniocentesis, funipuncture, therapeutic abortion, cesarean section, manual removal of the placenta, and pathologic conditions such as abdominal trauma, spontaneous abortion, or ectopic pregnancy are associated with increased risk of FMH and thus of sensitization (53). However, large FMH may occur in women without any predisposing risk factors. Although the volume of blood required to cause sensitization often is minuscule, the diagnosis and quantification of FMH is of paramount importance to determine the appropriate dose of Rh immunoglobulin necessary to prevent sensitization to Rh antigens. Recommendations for the administration of Rh immunoglobulin have been published (54). The rosette test or an enzyme-linked antiglobulin test are recommended for screening for excessive feto-maternal hemorrhage. The Kleihauer-Betke test, based on the ability of fetal hemoglobin containing cells to resist acid elution, permits quantification of fetal red cells in a maternal blood sample. Other assays also are used with varying degrees of accuracy (Table 73.4) (55).

TABLE 73.4. TESTS TO DETECT FETOMATERNAL HEMORRHAGE

Test	Fetal Marker Detected	Advantages	Disadvantages	Sensitivity (%)
Kleihauer-Betke	HbF	easy to perform	difficult to standardize	0.1
Rosette	RhD positive cells	easy to perform	Accurate quantification problematic	0.25
Micro D^u	RhD positive cells	easy to perform	15% of tests fail to detect fetal cells	0.75
Gel agglutination	RhD positive cells	easy to perform	Accurate quantification problematic	0.2
Flow cytometry	RhD positive cells	accurate detection and quantification	dedicated equipment and technical expertise needed	0.1–0.01
FISH* and DNA amplification	cell nucleus characteristics	extreme sensitivity	currently used only as research tool	0.003

*FISH, fluorescent in situ hybridization.
Source: modified from Ref 55.

Alloimmunization in Kell-negative women is more often the result of transfusion of Kell-positive blood rather than sensitization by FMH from a Kell-positive fetus; only about 9% of individuals express the Kell antigen (56). Kell HDN is rare even in alloimmunized pregnancies because transplacentally transmitted Anti–Kell antibodies will cause anemia only in a Kell-positive fetus. However, when present, Kell HDN may be somewhat unpredictable, partly because severe fetal anemia disproportionate to the degree of hemolysis may occur, due to specific suppression of fetal erythropoiesis at the progenitor cell level by antiKell antibodies (57).

Antenatal screening programs detect clinically significant antibodies in 0.24% to 1% of pregnant women. In such women, further evaluation is necessary to determine if the fetus is affected, and if so, how severely (58). In both Rh and Kell alloimmunization, the fetus is at risk only if the father is positive for the antigen in question. The probable Rh genotype of the father may be deduced from phenotyping studies, based on gene frequencies in various populations. Fetal blood typing early in pregnancy allows early institution of monitoring and therapy for the antigen-positive fetus at risk, and the avoidance of invasive procedures for the antigen-negative fetus that will remain unaffected. Prenatal determination of the fetal blood type may be performed on samples of blood obtained by percutaneous umbilical blood sampling (PUBS), or very early in pregnancy by flow cytometry or DNA amplification of tissue obtained by chorionic villus sampling or amniocentesis (59, 60). Recently described noninvasive techniques of prenatal diagnosis using fetal cells or fetal DNA isolated from the maternal blood obviate the risk of increasing maternal sensitization or fetal loss (61, 62).

Indirect assessment of the degree of fetal hemolysis is performed by sequential maternal antibody titration, amniotic fluid spectrophotometric analysis, and ultrasonography (63). Measurement of fetal blood counts, direct antiglobulin testing, and assessment of fetal acid-base status by PUBS allows direct evaluation of disease severity. Fetal blood samples with reticulocyte counts >97.5th percentile for gestation, strongly positive direct antiglobulin test, or mild anemia (hematocrit >30%, but < 2.5th percentile for gestation) predict fetuses at high risk of having significant antenatal anemia, thus requiring frequent

monitoring to determine if intrauterine transfusion is warranted (64). Management protocols reduce the need for multiple invasive procedures, because fetal blood sampling is associated with significant risks.

Intrauterine Fetal Transfusion

Intrauterine fetal transfusion has resulted in the perinatal survival of 75% to 90% of severely affected fetuses who would have perished in the 1960s. The first umbilical blood sampling with transfusion ideally should be performed when the fetus is anemic but before hydrops has developed. Intraperitoneal fetal transfusion has been largely replaced by direct intravascular fetal transfusion (IVT) by funipuncture. The intravascular technique offers precise diagnostic evaluation of the fetal status (see PUBS above) and is effective even in hydropic fetuses by circumventing the problem of erratic and often poor absorption of red blood cells from the peritoneal cavity in such fetuses. Other techniques of fetal transfusion reported include intrahepatic venous puncture, combinations of intravascular with intraperitoneal transfusions, and even intracardiac transfusion as a last resort (65). Intraperitoneal transfusions may be necessary when intravascular access is difficult, as in early pregnancy when the umbilical vessels are narrow, or later when increased fetal size prevents access to the umbilical cord. Transfusions are performed at hematocrits of 25% to 30% or less. Generally, the hematocrit drops by 1% to 2% per day in the transfused hydropic fetus; the fall in hematocrit is rapid in fetuses with severe hemolytic disease, necessitating a second transfusion within 7 to 14 days; the interval between subsequent transfusions usually is 21 to 28 days. Very low pretransfusion fetal hematocrits, rapid large increases in posttransfusion hematocrits, and increases in umbilical venous pressure during IVT are associated with fetal-death posttransfusion. The relative merits of direct, simple intravascular transfusion versus intravascular ET have been debated but the shorter procedure time with direct simple IVT has made it the procedure of choice at most centers. Fetuses are at risk for both posttransfusion cytomegalovirus and TAGVHD, so cytomegalovirus seronegative and/or leukodepleted, irradiated blood is recommended. The blood should be as fresh as possible to maximize *in vivo* survival, antigen-negative for the offending antibody, cross-matched against the mother's blood, washed free of the anticoagulant citrate and other additives, warmed, and packed to a hematocrit of 70% to 85% in a volume calculated to increase the fetal hematocrit to between 40% and 45%. Transfusions are provided up to 33 to 34 weeks, with delivery as soon as lung maturity is achieved by antenatal steroid therapy. The decision as to when to deliver the fetus is based on gestational age, fetal weight, and lung maturity, severity of fetal disease and response to transfusions, and the ease of performing the transfusions combined with antenatal ultrasound and Doppler studies. Infants who have received multiple intrauterine transfusions are delivered closer to term, often require less phototherapy and fewer ETs in the neonatal period (66). However, some still have significant hemolytic anemia at birth requiring aggressive intervention, and many require additional simple transfusions for severe and prolonged hyporegenerative anemia secondary to suppression of fetal erythropoiesis (67).

It must be remembered that the circulating red cells in the neonate that have received multiple intrauterine transfusions are the donor cells and this may lead to misleading results on initial blood grouping and a false negative direct antiglobulin test.

Exchange Transfusion

ET for the treatment of severe neonatal anemia and hyperbilirubinemia resulting from Rh incompatibility was first employed by Diamond et al. in 1948 (68). ET removes bilirubin and free maternal antibody in the plasma, corrects anemia, and replaces the infant's antibody-coated RBCs with antigen-negative RBCs that should have normal *in vivo* survival. The indications for early ETs, performed within 9 to 12 hours of birth, although debated, have remained essentially unchanged over the last 40 years, with minor modifications. Cord hemoglobin levels ≤110 g/L, cord bilirubin levels ≥5.5 mg/dL, and rapidly rising bilirubin concentrations ≥0.5 mg/dL/hour despite phototherapy are commonly used criteria for early ETs. Late ETs are performed when serum bilirubin concentrations threaten to exceed 20 mg/L in term infants, the level at which the risk of bilirubin encephalopathy or kernicterus is approximately 10%. Prematurity, hypoxemia, acidosis, hypothermia, and sepsis predispose to kernicterus at lower levels of bilirubin.

A double volume ET calculated as (2 × 80 mL/kg in a full-term infant) replaces 87% to 90% of the infant's blood volume, but eliminates only about 50% of the intravascular bilirubin, because of a slowly equilibrating tissue-bound pool. The use of albumin prior to ET in an effort to mobilize tissue bilirubin is controversial. Equilibration of extravascular and intravascular bilirubin and continued breakdown of sensitized and newly formed red cells by persisting maternal antibodies results in a rebound of bilirubin following initial ET, often necessitating repeated ETs in severe hemolytic disease. Also, drugs are removed during an exchange; therefore, infants on cardiotonic medications or antibiotics should have replacement doses (38).

Blood chosen for the exchange should be ABO compatible, negative for the antigen responsible for the hemolytic disease, and cross-matched against the mother's blood. Irradiated citrate-phosphate-dextrose (CPD or CPDAI) blood is used as whole blood or reconstituted whole blood (red cells suspended in saline, albumin, or plasma) with a hematocrit of 40% to 50% warmed through a temperature-controlled in-line blood warmer. Hypoxemic or acidotic infants should receive blood known to lack hemoglobin S. Additive solution anticoagulants are avoided. The age of the blood may vary, depending on availability, but the blood should be as fresh as possible (<7 days) to maximize the *in vivo* survival of the transfused red cells.

ETs may be performed by the traditional push-pull method with a single vascular access, usually the umbilical vein, or by isovolumetric techniques utilizing two access sites for simultaneous removal of the infant's blood and administration of new blood (69). Aliquots of 5 to 20 mL with a maximum of 5 mL/kg is withdrawn or infused in the discontinuous method, at a rate not exceeding 5 mL/kg every 3 minutes to avoid rapid fluctuations in arterial pressure, which are accompanied by changes in intracranial pressure. When an isovolumetric exchange is being done, volumes to be removed/reinfused should not exceed 2

mL/kg/minute. The duration of the exchange usually is 1 to 2 hours.

Potential complications of ET include hypocalcemia, hyper- and hypoglycemia, hyperkalemia, thrombocytopenia, dilutional coagulopathy, neutropenia, disseminated intravascular coagulation, umbilical venous and/or arterial thrombosis, necrotizing enterocolitis, and infection. Despite advances in the management of critically ill newborn infants, morbidity and mortality associated with ETs remains high, particularly in infants who are premature or sick or both. The risk of death or permanent serious sequelae has been estimated to be as high as 12% in sick infants, compared to less than 1% in healthy infants in a recent study (70). Careful clinical judgment is required in balancing the potential risk of adverse events from ET with the risk of bilirubin encephalopathy in ill infants.

Although the vast majority of ETs are performed for the treatment of HDN, the procedure also has been advocated for neonatal sepsis and occasionally utilized for drug overdose or toxicity (71), correction of life-threatening hyperkalemia (72), and disseminated intravascular coagulation (73). The rationale for the use of whole blood ETs in neonatal sepsis includes removal of bacteria and inflammatory mediators, correction of acidosis, and improvement of tissue oxygenation by the correction of anemia. ETs increase neutrophil counts and enhance neutrophil function when fresh whole blood is used (74); enhance IgG, IgA, and IgM levels (75, 76); and clear endotoxin (77). Improved survival rates have been demonstrated in neonates with Gram-negative sepsis and neutropenia or sclerema, conditions ordinarily associated with extremely high mortality rates.

Partial Exchange Transfusion

Partial ET using packed RBCs is used to correct severe anemia in critically ill hydropic infants with severe HDN until they can be stabilized sufficiently to undergo a complete, double-volume ET.

Partial ET is used more often to decrease the hematocrit in neonates with polycythemia. A hematocrit of greater than 65% is common in infants born at high altitude, or in those with intrauterine growth retardation or maternal diabetes. Although other factors may influence the viscosity of blood, the correlation of hematocrit to viscosity is linear up to 65% and exponential thereafter (78). Because instruments to measure viscosity are not readily available in many centers, the neonatal polycythemia-hyperviscosity syndrome usually is diagnosed by using the high hematocrit as a surrogate marker for hyperviscosity and a combination of symptoms that include poor feeding, hypoglycemia, tachypnea, central nervous symptoms such as hypotonia, tremors, seizures, abnormal renal function, and necrotizing enterocolitis (79). Partial ET to reduce the hematocrit to about 55% causes rapid amelioration in the clinical manifestations of polycythemia and is associated with reversal of cerebral blood flow abnormalities in symptomatic infants (80). However, it is uncertain whether the long-term outcome, particularly in asymptomatic infants, is improved by partial ET (81, 82). The present indications for treatment include infants with hematocrit ≥70% or symptomatic infants with hematocrit between 65% and 70%, but remain controversial. 5% albumin and plasma protein solutions have been utilized as replacement solutions, but isotonic saline has been found to be as effective as colloid (83). Peripheral vessels are preferred for withdrawing and infusing replacement fluid in polycythemic infants (84), but the umbilical vein also may be used. An increased incidence of necrotizing enterocolitis has been described with the use of fresh frozen plasma administered through the umbilical vein, but it is difficult to decipher if the posttransfusion necrotizing enterocolitis was part of the syndrome or a result of the procedure or the nature of the replacement fluid (85).

GRANULOCYTE TRANSFUSION

Part of "73 - Neonatal Transfusion"

Quantitative and qualitative abnormalities in the neutrophil/phagocytic system and defects in humoral immunity make neonates, particularly those born prematurely, uniquely susceptible to developing sepsis with increased risk of mortality following infection. The neutrophil storage pool consisting of polymorphonuclear neutrophils, bands, and metamyelocytes is smaller in infants than in adults, and rapidly exhausted in the presence of bacterial infection, leading to significant neutropenia. In addition, the precursor bone marrow progenitors, colony-forming unit-granulocyte, erythroid, monocyte, megakaryocyte (CFU-GEMM), colony-forming unit-granulocyte, macrophage (CFU-GM), and the committed colony-forming unit-granulocyte (CFU-G) are decreased in number and proliferating at a maximal rate, resulting in an inability to accelerate production of PMNs in response to infection. Diminished production of colony-stimulating growth factors also may be responsible for the poor proliferative response of the bone-marrow precursors (86). Functional abnormalities of the phagocytic system in stressed neonates include decreased deformability, chemotaxis, phagocytosis, bactericidal killing, and oxidative metabolism (87). Studies of the efficacy of granulocyte transfusions for the treatment of life-threatening infections in neonatal patients are difficult to evaluate because of small numbers of patients with varying degrees of illness and supportive care regimens, differing methods of harvesting neutrophils, and different transfusion protocols. The important factors to consider in granulocyte transfusions are quantity, quality, and compatibility (88, 89). In a meta-analysis of clinical studies of the efficacy of granulocyte transfusions in the treatment of bacterial sepsis, the dose of transfused granulocytes emerged as a significant predictor of success, with transfusion of adequate doses ($>0.5 \times 10^9$/kg/dose) producing an 18-fold increase in the odds of survival in septic neutropenic neonates, as compared to controls ($p<0.05$). Transfusion of inadequate doses was not associated with any survival benefit (90). Transfusion of granulocytes collected by continuous-flow centrifugation is more effective than transfusions of buffy-coat isolated granulocytes (91,92,93 and 94). Granulocytes collected by filtration leukapheresis are functionally defective and ineffective, as are leukocytes stored for 24 hours. Donor-recipient compatibility is improved by human leukocyte (HLA) matching or by leukocyte crossmatching in addition to ensuring erythrocyte compatibility. Technical difficulties and the cost of collecting granulocytes for transfusion in a timely manner, together with concern about the potential risks of viral transmission,

TAGVHD from transfused viable lymphocytes, sensitization to cellular blood product antigens and pulmonary leukoagglutinating reactions have limited the more widespread use of white blood cell transfusions. Increased attention is being paid to the potential use of granulocyte and granulocyte-macrophage colony stimulating factors (G-CSF, GM-CSF) for the adjuvant treatment of neonatal sepsis (95).

PLATELET TRANSFUSIONS

Part of "73 - Neonatal Transfusion"

Thrombocytopenia, defined as a platelet count of <150 × 10^9/L occurs in 18% to 20% of low-birth-weight infants admitted to the neonatal intensive care unit, but is much less common in term infants, occurring in less than 1% (96, 97). Low platelet counts may be a result of impaired production or increased peripheral destruction, secondary to congenital or acquired conditions (98). A combination of mechanisms is most likely in the critically ill infant. One fifth of infants with marked thrombocytopenia have evidence of disseminated intravascular coagulation (96). Quantitative and qualitative platelet disorders may cause significant bleeding with resulting morbidity, the most serious being intracranial hemorrhage. The incidence and severity of intraventricular hemorrhage and serious neurological sequelae is greater in low-birth-weight infants with thrombocytopenia than in those with normal platelet counts (99). However, the incidence of new intracranial hemorrhages or extension of existing hemorrhages was not different between infants given prophylactic platelet transfusions to maintain platelet counts of at least 150 × 10^9/L or control infants who were not transfused until platelet counts dropped to <50 × 10^9/L or had evidence of bleeding (100). Based on this study, platelet transfusions are not recommended for nonbleeding preterm infants with platelet counts more than 60 × 10^9/L. Some of the difficulty in determining the minimum level to which platelet counts may be allowed to fall prior to transfusion is that hemostatic competence is determined not only on the platelet count, but also on platelet function, vascular integrity, and levels of coagulation factors. Sick infants may have prolonged bleeding times despite only moderately reduced platelet counts (75 to 150 × 10^9/L). Indications for platelet transfusion are therefore derived by consensus (Table 73.5) (101).

TABLE 73.5. GUIDELINES FOR PLATELET TRANSFUSION SUPPORT OF NEONATES

Prophylactic Platelet Transfusions

- Stable preterm neonates with platelet counts <30 × 10^9/L
- Stable term neonates with platelet counts <20 × 10^9/L
- Sick preterm neonates with platelet counts <50 × 10^9/L
- Sick term infants with platelet counts <30 × 10^9/L
- Preparation for an invasive procedure, e.g., lumbar puncture or minor surgery in neonates with platelet counts <50 × 10^9/L, and for major surgery in neonates with platelet counts <100 × 10^9/L

Platelet Transfusions in Neonates with Clinically Significant Bleeding

- Neonates with platelet counts <50 × 10^9/L
- Neonates with conditions that increase bleeding, e.g., DIC and platelet counts <100 × 10^9/L
- Neonates with documented significant platelet functional disorders (e.g., Glanzmann's thrombasthenia) irrespective of the circulating platelet count

"Sick" infants in this context defined as those with a history of perinatal asphyxia, extremely low birth weight (<1000 g), need for ventilatory support with an inspired oxygen concentration >40%, clinically unstable, signs of sepsis, or those who require numerous invasive interventions.
Source: ref 101.

Two neonatal conditions warrant special attention. Neonatal alloimmune thrombocytopenia (NAIT), a condition analogous to Rhesus hemolytic disease of the newborn, is the result of feto-maternal platelet incompatibility. Sensitization of the mother to paternal antigens inherited by the fetus and present on fetal platelets, result in the production of maternal antiplatelet IgG antibodies, which cross the placenta to the fetal circulation and cause fetal thrombocytopenia. Unlike Rh disease, first-born infants may be affected. The rate of recurrence of the condition among platelet-antigen–positive siblings is 100%. Antibodies to the PL^{A1} (HPA-1) antigen are the most common cause of NAIT in Caucasian populations, whereas the PEN/YUK (HPA-4) antigen is implicated in Asians. Although PL^{A1} incompatibility is present in about 1 in 42 pregnancies, NAIT occurs in only 1 in 20 to 40 incompatible pregnancies, primarily in women who are phenotypically HLA-Dw52a. Intracranial hemorrhage occurs in 10% to 20% of cases of NAIT, half of them occurring antenatally (102). A combination of noninvasive strategies including administration of IVIG and platelets to the mother with monitoring by PUBS, and intrauterine platelet transfusions coupled with cesarean section for select cases may prevent intracranial hemorrhage in the fetus at risk for NAIT. After birth, platelet counts may continue to drop in the infant with NAIT, because of continuing destruction by persisting maternal antiplatelet antibody. Random donor platelets, 98% of which are PL^{A1} (HPA-1) positive are ineffective. Transfusion of maternal platelets washed free of antibody or PL^{A1} (HPA-1) antigen-negative platelets from unrelated donors provides a sustained increase in platelet counts and is the treatment of choice. Adjuvant therapy with steroids and/or intravenous immunoglobulin may be indicated.

Neonatal autoimmune thrombocytopenia is a condition resulting from destruction of platelets by transplacentally transferred antibody, directed to platelet antigen shared by both mother and fetus, most often a result of maternal immune thrombocytopenia. Unlike NAIT, fetal hemorrhage is rare in autoimmune thrombocytopenia (103). Cesarean section is advocated if fetal scalp platelet counts are less than 50 × 10^9/L at the time of labor, but surgery is complicated by increased risk in the presence of maternal thrombocytopenia. Neonatal thrombocytopenia usually is mild to moderate. Platelet transfusions of any antigen type generally are ineffective because the antibodies are reactive against a wide range of platelet antigens.

FRESH FROZEN PLASMA AND CRYOPRECIPITATE TRANSFUSIONS

Part of "73 - Neonatal Transfusion"

The levels of vitamin-K–dependent coagulation factors II, VII, IX, and X and the contact factors XI, XII, prekallikrein and high-molecular-weight kininogen are around 50% of normal adult concentrations at birth in term infants, and are even lower in

preterm infants. Classical “hemorrhagic disease of the newborn,” from the transient deficiency of vitamin-K–dependent factors and characterized by bruising and gastrointestinal hemorrhage in neonates around 2 to 5 days of age, now has been effectively prevented by routine administration of vitamin K at birth. However, two other forms of hemorrhagic disease, secondary to vitamin-K deficiency continue to be reported. Early hemorrhagic disease, presenting within 24 hours of birth, is associated with the ingestion in pregnancy of drugs such as phenytoin, Phenobarbital, and primidone, which interfere with vitamin K metabolism. Late hemorrhagic disease appears between 2 and 8 weeks after birth and may be secondary to vitamin K malabsorption resulting from mild, undiagnosed liver disease. In both conditions, life-threatening intracranial hemorrhage may occur. Intracranial hemorrhage at birth may be the presenting feature in infants with severe forms of inherited coagulation-factor deficiencies, although most infants with inherited coagulation-factor deficiencies do not have bleeding in the perinatal period, unless they undergo an invasive procedure such as circumcision or venipuncture. The most common cause of disordered hemostasis in sick neonates is disseminated intravascular coagulation, secondary to myriad conditions, including asphyxia, sepsis, shock, respiratory-distress syndrome, or necrotizing enterocolitis. Coagulopathy also may be secondary to acute or chronic liver disease.

FFP infusions are indicated for the treatment of clinically significant bleeding as well as for correction of hemostatic defects as a result of a decrease in one or more coagulation factors, prior to invasive procedures. Rationale for the use of FFP in neonates again has been derived from evidence in adult patients (104, 105). The dose of FFP administered usually is 10 to 20 mL/kg, with further doses being determined by the clinical situation, underlying disease process, and laboratory monitoring. The amelioration of hemostatic problems by the transfusion of FFP in neonates with DIC is only temporary, and hence multiple transfusions may be required until the underlying cause resolves. Indeed, the outcome of DIC has been found to be dependent on the success of treatment of the underlying pathologic process and aggressive supportive care, and is not altered by therapy specifically directed at coagulopathy (73). The correction of coagulopathy secondary to liver failure also is temporary, in the absence of recovery of hepatic function.

FFP is not indicated for the treatment of hypovolemic shock in the absence of coagulation abnormalities, for nutritional support, treatment of immunodeficiency states, or for the treatment of isolated coagulant factor deficiencies for which virus-inactivated, plasma-derived or recombinant factor concentrates are available. Although earlier studies demonstrated that the incidence of periventricular — intraventricular hemorrhage was reduced by prophylactic FFP, a recent large, multicenter, randomized, controlled trial provided no evidence that the routine early use of FFP affected the rates of death or disability in preterm infants (106).

Cryoprecipitate is indicated for severe hypofibrinogenemia associated with DIC, and for bleeding as a result of congenital fibrinogen or factor XIII deficiency, if commercial plasma-derived, virus-inactivated concentrates are unavailable.

EXTRACORPOREAL MEMBRANE OXYGENATION

Part of "73 - Neonatal Transfusion"

Extracorporeal membrane oxygenation (ECMO) is the use of prolonged extracorporeal circulation and gas exchange through a membrane oxygenator to provide temporary life support in patients with profound cardiorespiratory failure who fail to respond to conventional therapy. The continuous contact of blood with the prosthetic surface of the ECMO circuit results in activation of the coagulation cascade. The use of heparin to prevent clot formation in the circuit and the decrease in platelet number and function from interaction with the circuit put the infant at risk for bleeding (107, 108). In addition, infants placed on ECMO are critically ill and already may have altered hemostasis secondary to sepsis, shock, or profound hypoxia, predisposing them to devastating intracranial hemorrhage (109).

Two units of group- and type-specific packed RBCs are required to prime the circuit. Subsequently, multiple small-volume transfusions of RBCs are needed to maintain the hematocrit at 40% to 45%. Platelet and FFP support is variable, with postsurgical infants, patients with sepsis, or those on prolonged bypass requiring considerably more transfusions (110, 111). The necessity for multiple blood components for neonates on ECMO often resulted in mean donor exposures as high as 22.8 ± 9.5 per ECMO run in the past (112). Changes in transfusion practices and blood-component management now have limited donor exposures to 7 to 10 in many centers (111,112 and 113). Strategies for limiting donor exposure include increasing the volume of red cells used for small-volume transfusions to 15 mL/kg, using a system of sequential aliquots from the same blood unit, using standard FFP units up to 24 hours after thawing, volume-reduced platelets, or single-donor apheresis platelet aliquots, and eliminating the empirical use of FFP and cryoprecipitate. Most institutions maintain platelet counts of at least 100×10^9/L in neonates in ECMO.

ADVERSE REACTIONS TO BLOOD TRANSFUSION

Part of "73 - Neonatal Transfusion"

The physiological immaturity of various organ systems in the newborn gives rise to significant differences in the incidence and types of adverse reactions when compared to older children or adults. The important differences will be emphasized in this section.

Hemolytic Reactions

Acute hemolytic transfusion reactions are rare in neonates, partly from the absence of naturally occurring anti-A or anti-B antibodies until the age of 4 months, and also from the practice of transfusing neonates with group O blood. Despite multiple blood transfusions and exposure to multiple blood donors, posttransfusion red-cell alloimmunization has been reported infrequently in neonates — again, a feature of immunological immaturity (114, 115).

However, severe immune-mediated hemolysis has been associated with transfusion of adult blood containing anti-T antibodies

in neonates with activation of the normally masked Thomsen-Friedenreich antigen on the surface of the neonatal red-cell membrane (116). Expression of the cryptantigen, known as T-activation, occurs by the removal of N-acetyl neuraminic acid residues from the antigen by neuraminidases produced by bacteria, particularly *Clostridia*. This phenomenon has been reported mainly in neonates with necrotizing enterocolitis, especially in those with severe disease requiring surgical intervention and with Clostridial infection. T-activation should be suspected in neonates at risk who have evidence of intravascular hemolysis with hemoglobinuria and hemoglobinemia following transfusion of blood products or unexplained failure to achieve the expected posttransfusion hemoglobin increment. Routine cross-matching techniques will not detect the polyagglutination from T-activation when monoclonal ABO antiserum is used. Minor crossmatching of neonatal T-activated red cells with donor anti-T–containing serum may show agglutination, but this is not performed routinely. Infants at risk, with discrepancies in forward and reverse blood grouping and evidence of hemolysis on smear should be suspected of T-antigen activation. The diagnosis is confirmed by specific agglutination tests using peanut lectin *Arachis hypogea* and *Glycine soja*. Further hemolysis may be prevented by using washed red cells, platelets, and low-titer anti-T plasma. Exchange transfusion with plasma-reduced components may be necessary for infants with severe ongoing hemolysis.

Numerous nonimmunological causes of hemolysis leading to morbidity and mortality have been described in neonates, including excessive infusion pressure through small-gauge needles (≥28 gauge); accidental overheating of the blood in the administration set; inappropriate storage of red-blood-cell products in freezers or unmonitored refrigerators leading to freezing and lysis; simultaneous infusion of incompatible fluids or drugs; and the transfusion of abnormal donor red blood cells (glucose-6-phosphate dehydrogenase deficiency, sickle-cell trait, or paroxysmal nocturnal hemoglobinuria) (117).

Leukocyte Alloimmunization

Transient anti-HLA antibodies have been described in multiply-transfused, preterm infants but their immunological effects either in the short or long term are not clear at present (115). The passive transfusion of blood containing leukoagglutinins directed against the recipients antigens initiates complement activation, microvascular lung injury, pulmonary edema, and severe hypoxemia, a condition termed transfusion-related acute lung injury (TRALI). This acute, life-threatening complication rarely is reported in neonates, partly because of the difficulty in distinguishing the acute respiratory distress that accompanies the syndrome from other causes of respiratory deterioration (118).

Transfusion-Associated Graft-Versus-Host Disease

TAGVHD occurs as a result of the proliferation and engraftment of viable donor lymphocytes in an immunosuppressed or immunodeficient transfusion recipient who does not recognize the cells as being foreign and is unable to reject them (119). Similarities in HLA antigens facilitate engraftment, and therefore, this phenomenon is much more likely to occur when family members serve as blood donors. TAGVHD in infants has been reported following transfusions from blood relatives, following intrauterine fetal transfusion, after postnatal transfusions in infants who received intrauterine transfusions, in neonates with primary-cell mediated immunodeficiency disease, and rarely, in infants without a definitive diagnosis of underlying primary immunodeficiency. The manifestations of TAGVHD include fever, rash, gastrointestinal symptoms, bone marrow and hepatic failure, similar to many other medical conditions in sick newborn infants, making clinical diagnosis difficult, and probably leading to considerable underestimates of the true incidence of the condition (120). The threshold number of lymphocytes necessary to produce a graft-versus-host reaction may vary depending upon the immune status of the host and the antigenic similarity of the donor-recipient pair. No leukocyte depletion filter or processing has yet been shown to uniformly decrease the lymphocyte count of a product below the critical threshold. Whole blood, red blood cells, granulocyte and platelet concentrates, and fresh plasma have been implicated in TAGVHD. Irradiation at a minimum of 2,500 cGy to the midplane of the bag currently is the only acceptable way to abrogate the disease. Current guidelines recommend the use of irradiated blood components for neonates with known or suspected congenital cell-mediated immunodeficiency states; infants with congenital leukemia, or malignancy undergoing chemotherapy; fetuses receiving intrauterine transfusions; postnatal transfusions in neonates who received intrauterine transfusions; exchange transfusions; and for recipients of cellular blood products from first- and second-degree relatives and HLA-matched donations (121). TAGVHD has not been reported in children or adults with human immunodeficiency virus (HIV) infection despite the severe immunodeficiency associated with the disease, so irradiation of blood products is not recommended for this indication at present. Although preterm infants are known to be immunoincompetent, it is unclear if all preterm infants should receive only irradiated blood products, in the absence of other factors that would predispose them to TAGVHD. Some institutions provide irradiated blood products to all neonates, to avoid TAGVHD in patients with undiagnosed immunodeficiency.

Metabolic Adverse Reactions

Inhibition of the red cell membrane bound adenosine triphosphatase pump during extended storage causes leakage of intracellular potassium, resulting in K+ levels as high as 85 mEq/L in the supernatant plasma in blood stored 42 days, particularly in blood that is irradiated on day 0, prior to storage. Small volume transfusions of red blood cells stored to their expiration date of 35 to 42 days do not result in clinically significant elevations of serum potassium, because infants receiving 15 mL/kg of packed RBCs receive only about 3 mL/kg of supernatant, containing approximately 0.15 mEq/kg of K+ (122). Life-threatening hyperkalemia has been described following rapid infusions of large volumes and exchange transfusions using stored, packed RBCs in sick term and preterm infants. Washed red cells are advisable in

infants with hyperkalemia, renal failure, or when large volumes of blood are infused rapidly.

Hypoglycemia may occur during transfusion of small volumes of CPDA-1 or AS-1 red cells, if dextrose containing intravenous infusion is interrupted. Rebound hypoglycemia is detected frequently after exchange transfusions in infants with erythroblastosis fetalis who often exhibit hyperinsulinemia. Large volume transfusions, as may occur during surgical procedures, may lead to hyperglycemia in small premature neonates with poor glucose-control mechanisms.

Hypocalcemia, secondary to the chelation of calcium by citrate in the anticoagulant, has been observed during exchange transfusions and following massive transfusions both in term and preterm infants.

The plasticizer di-(2-ethylhexyl)-phthalate (DEHP), present in blood bags, intravenous (IV) sets, and ECMO circuits made of polyvinyl chloride, is lipophilic and hence, leaches into almost all blood products. Significant exposure to plasticizer is reported in infants undergoing exchange transfusions and ECMO (123, 124). Although no short- or long-term complications have been identified in infants, studies in animals raise concerns about possible effects on hepatocellular, pulmonary, and reproductive function, and potential carcinogenicity (117).

In addition to metabolic effects, blood transfusions may cause significant alterations in temperature, blood pressure, and pulmonary compliance in fragile preterm infants.

Transfusion-Transmitted Infections

No study has ever clearly demonstrated that neonates are at any higher risk for acquiring posttransfusion infections than are older children, although premature infants appear to be more susceptible to transfusion-transmitted HIV and cytomegalovirus (CMV) infections. The practice of using one unit of blood for many infants has resulted in multiple infants being infected from a single donor. This has been reported with hepatitis A and C, malaria, Epstein Barr virus, and HIV.

Nearly all cases of HIV infection acquired from blood transfusions in the United States occurred before the implementation of routine blood donor screening practices. The median symptom-free period from birth to symptomatic infection is reported to be significantly longer in neonatal transfusion acquired infection as compared to perinatally acquired infection (125, 126). Asymptomatic HIV-1 infection has been identified as late as 5 to 9.5 years posttransfusion in studies in which cohorts of neonates transfused before routine donor screening for HIV antibody have been retrospectively tested (127). Once symptomatic, the progression of transfusion-acquired HIV disease does not appear to be different from vertically acquired infection in neonates. The potential risk of transfusion-transmitted HIV infection from a repeat blood donor with negative serologic tests in the United States is estimated to be 1 in 493,000 (128).

Improvements in donor screening have reduced the incidence of posttransfusion hepatitis, but have not completely eliminated this complication. The hepatitis viruses may be transmitted by cellular and noncellular blood products.

Posttransfusion hepatitis A is rare, and usually asymptomatic in newborn infants. However, secondary horizontal spread has led to large outbreaks in neonatal intensive care units, with symptomatic infection in nurses, physicians, parents, and relatives who have been in direct contact with infected neonates (129,130 and 131). The risk of transmitting the hepatitis B by the transfusion of blood from screened donors in the United States is estimated to be 1 in 63,000. The outcome of transfusion-acquired hepatitis B in neonates is assumed to be similar to perinatally acquired infection, with a 70% risk of chronic carrier state.

The risk of transmitting hepatitis C virus from screened blood units in the United States, using second-generation enzyme immunoassays (EIA-2) for detecting antibody to the virus, is estimated to be 1 in 103,000 (128). One cohort study of children who received transfusions of blood screened by surrogate markers of hepatitis or by EIA-1 antihepatitis C virus testing, during extracorporeal membrane oxygenation therapy in the neonatal period, found 7 of 83 children to be antihepatitis C virus seropositive, with significantly increased alanine aminotransferase levels in four of the seven (132). Another retrospective cohort study of neonates transfused with blood from donors infected with hepatitis C virus prior to the institution of routine screening indicated that infection was asymptomatic in more than 60% of children at the median age of 6.3 years. Children who were hepatitis C virus-RNA positive had mildly elevated transaminase levels; mild hepatitis was evident in those who underwent liver biopsy (133). The persistence of HCV infection in children after transfusions in the neonatal period is particularly disquieting because studies in adults reveal that 70% of hepatitis C virus-infected patients develop chronic hepatitis, 20% develop cirrhosis within two decades, and 1% to 5% may develop hepatocellular carcinoma after 20 years (134).

Persistent hepatitis G virus infection has been observed after neonatal transfusions, but the term “hepatitis” in relation to this virus may be a misnomer because no clinical or biochemical signs of liver disease have been observed on follow up (135).

CMV is widely prevalent in blood donors, and has serious consequences in low-birth-weight neonates. Viral genome has been detected in the white blood cells of CMV seropositive individuals, and even CMV seronegative donors may harbor latent CMV infection in their mononuclear cells. Risk factors for neonatal acquisition of CMV from transfusions include exposure to at least 50 mL of blood, birth weight less than 1,200 g, and maternal CMV seronegativity. Posttransfusion CMV infection in low-birth-weight infants born to seronegative mothers causes a serious clinical syndrome of fever, respiratory distress, hepatosplenomegaly, cytopenias, and may result in death (136). Disseminated CMV infection has been described in a full-term infant following ECMO (137) CMV seronegative blood or leukocyte reduced blood is recommended for transfusions in low-birth-weight infants born to seronegative mothers or those with unknown serostatus, and for intrauterine transfusions (138). Leukocyte reduction may be performed by washing, freezing followed by washing, and filtration using third-generation leukocyte depletion filters. There is a suggestion that the transfusion of CMV seronegative blood products in infants born to CMV seropositive mothers may hasten the decline in titer of protective maternal antibody, making infants susceptible to infection.

In addition to transfusion-transmitted viral infections,

syphilis, malaria, and bacterial infections from transfusions continue to be reported, albeit infrequently. Case reports of transfusion-associated parvovirus B-19 (139), babesiosis (140), Chagas disease, and the potential risk of transmission of Creutzfeldt-Jacob disease (141) emphasize the fact that the goal of zero-risk transfusion is distant.

REFERENCES

1. Vamvakas EC, Taswell HF. Long-term survival after blood transfusion. *Transfusion* 1994;34:471.

2. Widness JA, Seward VJ, Kromer IJ et al. Changing patterns of red blood cell transfusion in very low birth weight infants. *J Pediatr* 1996;129:680–687.

3. Williams GD, Bratton SL, Riley EC, et al. Association between age and blood loss in children undergoing open-heart operations. *Ann Thorac Surg* 1998;66:870.

4. Petaja J, Lundstrom U, Leijala M et al. Bleeding and use of blood products after heart operations in infants. *J Thoracic Cardiovas Surg* 1995;109:524.

5. Stockman JA. Anemia of prematurity: current concepts in the issue of when to transfuse. *Ped Clin N Am* 1986:33:111.

6. Obladen M, Sachsenweger M, Stahnke M. Blood sampling in very low birth weight infants receiving different levels of intensive care. *Eur J Pediatr* 1988;147:399.

7. Bifano EM, Curran TR. Minimizing blood donor exposure in the neonatal intensive care unit — current trends and future prospects. *Clin Perinatol* 1995;22:657.

8. Strauss RG. Transfusion therapy in neonates. *Am J Dis Child* 1991;145:904.

9. Van Newkirk LE, Bhutani VK, Husson MA, et al. Impact of reducing blood sample size on the incidence of transfusion in a neonatal ICU. *Lab Med* 1998;29:306.

10. Kling PJ, Sullivan TM, Leftwich ME, et al. Score for neonatal acute physiology and phlebotomy loss predict erythrocyte transfusions in premature infants. *Arch Pediatr Adolesc Med* 1997;151:27.

11. Ringer SA, Richardson DK, Sacher RA, et al. Variations in transfusion practice in neonatal intensive care. *Pediatrics* 1998;101:194.

12. Maier RF, Metze B, Obladen M. Low degree of regionalization and high transfusion rates in very low birth weight infants: a survey in Germany. *J Perinat Med* 1998;26:43.

13. Bednarek FJ, Weisberger S, Richardson DK, et al. Variations in blood transfusion among newborn intensive care units. *J Pediatr* 1998;133:601.

14. Alverson DC. The physiological impact of anemia in the neonate. *Clin Perinatol* 1995;22:609.

15. Ramasethu J, Luban NLC. Red cell transfusions in the newborn. *Semin Neonatol* 1999;4:5.

16. Shannon KM, Keith JF, Mentzer WC et al. Recombinant human erythropoietin stimulates erythropoiesis and reduces erythrocyte transfusions in very low birth weight preterm infants. *Pediatrics* 1995;95:1.

17. Simon TL, Alverson DC, AuBochon J, et al. Practice parameter for the use of red blood cell transfusions. Developed by the red blood cell administration practice guideline development task force of the College of American Pathologists. *Arch Pathol Lab Med* 1998;122:130.

18. Fetus and Newborn Committee, Canadian Pediatric Society. Guidelines for transfusion of erythrocytes to neonates and premature infants; position statement. *Can Med Assoc J* 1992;147:1781.

19. Voak D, Cann R, Finney D, et al. Guidelines for administration of blood products: transfusion of infants and neonates. *Transfus Med* 1994;4:63.

20. Standards Committee of the American Association of Blood Banks. Standards for blood banks and transfusion services 19th ed., 1999. AABB Bethesda, Maryland.

21. Luban NLC, Strauss RG, Hume HA. Commentary on the safety of red cells preserved in extended -storage media for neonatal transfusion. *Transfusion* 1991;31:229.

22. Lee DA, Slagle TA, Jackson TM, et al. Reducing blood donor exposures in low birth weight infants by the use of older, unwashed packed red blood cells. *J Pediatr* 1995;126:280.

23. Strauss RG, Burmeister LF, Johnson K, et al. AS-1 red blood cells for neonatal transfusions: a randomized trial assessing donor exposure and safety. *Transfusion* 1996;36:873.

24. Goodstein MH, Herman JH, Smith JF, et al. Metabolic consequences in very low birth weight infants transfused with older AS-1 preserved erythrocytes. *Ped Pathol Molec Med* 1999;18:173.

25. Goodstein MH, Locke RG, Wlodarczyk D, et al. Comparison of two preservation solutions for erythrocyte transfusions in newborn infants. *J Pediatr* 1993;133:783.

26. Rock G, Poon A, Haddad S, et al. Nutricel as an additive solution for neonatal transfusion. *Transfusion Science* 1999;20:29.

27. Strauss RG, Sacher RA, Blazina JF. Commentary on small volume red cell transfusions for neonatal patients. *Transfusion* 1990;30:565.

28. Strauss RG, Villhauer PJ, Cordle DG. A method to collect, store and issue multiple aliquots of packed red blood cells for neonatal transfusions. *Vox Sang* 1995;68:77.

29. Liu EA, Mannino FL, Lane TA. Prospective randomized trial of the safety and efficacy of a limited donor exposure transfusion program for premature neonates. *J Pediatr* 1994;125:92.

30. Wang-Rodriguez J, Mannino FL, Lane TA. A novel strategy to limit blood donor exposure and blood waste in multiply transfused premature infants. *Transfusion* 1996;36:64.

31. Delivoria-Papadopoulis M, Marrow G, Oski FA, et al. Exchange transfusion of the newborn infant with fresh and 'old' blood: the role of storage on 2, 3 -diphosphoglycerate, hemoglobin-oxygen affinity, and oxygen release. *J Pediatr* 1971;70:898.

32. Strauss RG. Recombinant erythropoietin for the anemia of prematurity: still a promise, not a panacea. *J Pediatr* 1997;131:653.

33. Asch J, Wedgewood JF. Optimizing the approach to anemia in preterm infants: is there a role for erythropoietin therapy? *J Perinatol* 1997;17:276.

34. Maier RF, Obladen M, Scigalla P, et al. The effect of epoetin beta (recombinant human erythropoietin) on the need for transfusion in very low birth weight infants. *N Engl J Med* 1994;330:1173.

35. Meyer MP, Meyer JH, Commerford A, et al. Recombinant human erythropoietin in the treatment of the anemia of prematurity: results of a double-blind placebo-controlled study. *Pediatrics* 1994;93:918.

36. Ohls RK, Harcum J, Schibler KR, et al. The effect of erythropoietin on the transfusion requirements of preterm infants weighing 750g or less: a randomized, double blind, placebo controlled study. *J Pediatr* 1997;131:661.

37. Manno CS, Hedberg KW, Kim HC. Comparison of the hemostatic effects of fresh whole blood, stored whole blood, and components after open heart surgery in children. *Blood* 1991;77:930.

38. Luban NLC. Massive transfusion in the neonate. *Transfusion Med Rev* 1995;9:200.

39. Elbert C, Strauss RG, Barrett F. Biological mothers may be dangerous blood donors for their neonates. *Acta Haematol* 1991;85:189.

40. Strauss RG, Barnes A Jr, Blanchette VS, et al. Directed and limited-exposure blood donations for infants and children. *Transfusion* 1990;30:68.

41. Kinmond S, Aitchison TC, Holland BM. Umbilical cord clamping and preterm infants: a randomised trial. *Br Med J* 1993;306:172.

42. Linderkamp OL, Nelle M, Kraus M, et al. The effects of early and late cord-clamping on blood viscosity and other hemorheological parameters in full-term neonates. *Acta Pediatr* 1992;81:745.

43. Paxson CL. Collection and use of autologous fetal blood. *Am J Obstet and Gynecol* 1979;134:708.

44. Golden SM, O'Brien WF, Metz SA. Anticoagulation of autologous cord blood for neonatal resuscitation. *Am J Obstet Gynecol* 1982;144:103.

45. Ballin A, Kenet G, Berar M, et al. Autologous umbilical cord blood transfusion. *Arch Dis Child* 1995;73:F181.

46. Anderson S, Fangman J, Wagner G, et al. Retrieval of placental blood from the umbilical vein to determine volume sterility and presence of clot formation. *Am J Dis Child* 1992;146:36.

47. Bifano EM, Dracker RA, Lorah K, et al. Collection and 28-day storage of human placental blood. *Pediatr Res* 1994;36:90.

48. Geifman-Holtzman O, Wojtowycz M, Kosmas E, et al. Female alloimmunization with antibodies known to cause hemolytic disease. *Obstet Gynecol* 1997;89:272.

49. Lin M, Broadberry RE. ABO Hemolytic disease of the newborn is more severe in Taiwan than in white populations. *Vox Sang* 1995;68:136.

50. Cariani L, Romano EL, Martinez N, et al. ABO-haemolytic disease of the newborn (ABO-HDN): factors influencing its severity and incidence in Venezuela. *J Trop Pediatr* 1995;41:14.

51. McDonnell M, Hannam S, Devane SP. Hydrops fetalis due to ABO incompatibility. *Arch Dis Child Fetal Neonatal Ed* 1998;78:F220.

52. Bowman JM, Pollack JM, Penston LE. Fetomaternal transplacental hemorrhage during pregnancy and after delivery. *Vox Sang* 1986;51:117.

53. Sebring ES, Polesky HF. Fetomaternal hemorrhage: incidence, risk factors, time of occurrence, and clinical effects. *Transfusion* 1990;30:344.

54. Hartwell EA. Use of Rh immune globulin. ASCP practice parameter. *Am J Clin Pathol* 1998;110:281.

55. Duguid JKM, Bromilow IM. Laboratory measurement of fetomaternal hemorrhage and its clinical relevance. *Trans Med Rev* 1999;13:43.

56. Bowman JM, Pollack JM, Manning FA, et al. Maternal Kell blood group alloimmunization. *Obstet Gynecol* 1992;79:239.

57. Vaughan JI, Manning M, Warwick RM, et al. Inhibition of erythroid progenitor cells by anti-Kell antibodies in fetal alloimmune anemia. *N Engl J Med* 1998;338:798.

58. ACOG Educational Bulletin. Management of isoimmunization in pregnancy. Number 227, August 1996.

59. Nelson M, Forsyth C, Popp H, et al. Rapid detection of Rh(d)- or K-positive fetal red cells in chorionic villus samples by a flow cytometric technique. *Transfusion Med* 1994;4:297.

60. Bennett PR, Le Van Kim C, Colin Y, et al. Prenatal determination of fetal RhD type by DNA amplification. *N Engl J Med* 1993;329:607.

61. Lo YMD, Bowell PJ, Selinger M, et al. Prenatal determination of fetal RhD status by analysis of peripheral blood of rhesus negative mothers. *Lancet* 1993;341:1147.

62. Lo YMD, Hjelm NM, Fidler C, et al. Prenatal diagnosis of fetal RhD status by molecular analysis of maternal plasma. *N Engl J Med* 1998;339:1734.

63. Bowman J. The management of hemolytic disease in the fetus and newborn. *Semin Perinatol* 1997;21:39.

64. Weiner CP, Williamson RA, Wenstrom RD, et al. Management of fetal hemolytic disease by cordocentesis. I. Prediction of fetal anemia. *Am J Obstet Gynecol* 1991;165:546.

65. Schumacher B, Moise Jr KJ. Fetal transfusion for red blood cell alloimmunization in pregnancy. *Obstet Gynecol* 1996;88:137.

66. Janssens HM, deHaan MJJ, van Kamp IL, et al. Outcome for children treated with fetal intravascular transfusions because of severe blood group antagonism. *J Pediatr* 1997;131:373.

67. Millard DD, Gidding SS, Socol ML, et al. Effects of intravascular intrauterine transfusion on prenatal and postnatal hemolysis and erythropoiesis in severe fetal isoimmunization. *J Pediatr* 1990;117:447.

68. Diamond LK, Allen FH Jr, Thomas Jr WO. Erythroblastosis fetalis. VII. Treatment with exchange transfusion. *N Engl J Med* 1951;244:39.

69. Edwards MC, Fletcher MA. Exchange transfusions. In: Fletcher MA, MacDonald MG, eds. *Atlas of Procedures in Neonatology*, 2nd ed. Philadelphia: JB Lippincott, 1993:363.

70. Jackson JC. Adverse events associated with exchange transfusion in healthy and healthy and ill newborns. *Pediatrics* 1997;99:e7.

71. Osborn HH, Henry G, Wax P, et al. Theophylline toxicity in a premature neonate-elimination kinetics of exchange transfusion. *J Toxicol Clin Toxicol* 1993;31:639.

72. Lorch V, Jones FS, Hoersten IR. Treatment of hyperkalemia with exchange transfusion. *Transfusion* 1985;25:390.

73. Gross SJ, Filston HC, Anderson JC. Controlled study of treatment for disseminated intravascular coagulation in the neonate. *J Pediatr* 1982;3:445.

74. Mathur NB, Subramanian BKM, Sharma VK, et al. Exchange transfusion in neutropenic septicemic neonates: effect on granulocyte functions. *Acta Paediatr* 1993;82:939.

75. Vain NE, Mazlumian JR, Swarner OW, et al. Role of exchange transfusion in the treatment of severe septicemia. *Pediatrics* 1980;66:693.

76. Sadana S, Mathur NB, Thakur A. Exchange transfusion in septic neonates with sclerema: effect on immunoglobulin and complement levels. *Ind Pediatr* 1997;34:20.

77. Togari T, Mikawa M, Iwanaga T, et al. Endotoxin clearance by exchange blood transfusion in septic shock neonates. *Acta Paediatr Scand* 1983;72:87.

78. Ramamurthy RS, Berlanga M. Postnatal alteration in hematocrit and viscosity in normal and polycythemic infants. *J Pediatr* 1987;110:929.

79. Werner EJ. Neonatal polycythemia and hyperviscosity. *Clin Perinatol* 1995;22:693.

80. Bada HS, Korones SB, Pourcyrous M, et al. Asymptomatic syndrome of polycythemic hyperviscosity: effect of partial plasma exchange transfusion. *J Pediatr* 1992;120:579–585.

81. Black VD, Lubchenco LO, Koops BL, et al. Neonatal hyperviscosity: randomized study of effect of partial plasma exchange transfusion on long term outcome. *Pediatrics* 1985;75:1048.

82. Delaney-Black V, Camp BW, Lubchenco LO, et al. Neonatal hyperviscosity association with lower achievement and IQ scores at school age. *Pediatrics* 1989;83:662.

83. Wong W, Fok TF, Lee CH, et al. Randomized controlled trial: comparison of colloid or crystalloid for partial exchange transfusion for treatment of neonatal polycythemia. *Arch Dis Child* 1997;77:F115.

84. Fok TF, So LY, Leung KW, et al. Use of peripheral vessels for exchange transfusion. *Arch Dis Child* 1990;65:676.

85. Black VD, Rumack CM, Lubchenco LO, et al. Gastrointestinal injury in polycythemic term infants. *Pediatrics* 1985;76:225.

86. Sweetman RW, Cairo MS. Blood component and immunotherapy in neonatal sepsis. *Trans Med Rev* 1995;9:251.

87. Shigeoka A, Santos J, Hill H. Functional analysis of neutrophil granulocytes from healthy, infected and stressed newborns. *J Pediatr* 1979;95:454.

88. Strauss RG. Clinical perspectives of granulocyte transfusions: efficacy to date. *J Clin Apher* 1995;10:114.

89. Price TH. Blood center perspective of granulocyte transfusions: future applications. *J Clin Apher* 1995;10:119.

90. Vamvakas EC, Pineda AA. Meta-analysis of clinical studies of the efficacy of granulocyte transfusions in the treatment of bacterial sepsis. *J Clin Apher* 1996;11:1.

91. Cairo MS, Rucker R, Bennetts GA, et al. Improved survival of newborns receiving leukocyte transfusions for sepsis. *Pediatrics* 1984;74:887.

92. Cairo MS, Worcester C, Rucker R, et al. Role of circulating complement and polymorphonuclear leukocyte transfusion in treatment and outcome in critically ill neonates with sepsis. *J Pediatr* 1987;110:935.

93. Baley JE, Stork EK, Warkentin PI, et al. Buffy coat transfusions in neutropenic neonates with presumed sepsis: a prospective randomized trial. *Pediatrics* 1987;80:712.

94. Wheeler JG, Chauvenet AR, Johnson CA, et al. Buffy coat transfusions in neonates with sepsis and neutrophil storage pool depletion. *Pediatrics* 1987;79:422.

95. Bracho F, Goldman S, Cairo MS. Potential use of granulocyte colony-stimulating factor and granulocyte — macrophage colony stimulating factor in neonates. *Curr Opin Hematol* 1998;5:215.

96. Castle V, Andrew M, Kelton J, et al. Frequency and mechanism of neonatal thrombocytopenia. *J Pediatr* 1986;108:749.

97. Oren H, Irken G, Oren B, et al. Assessment of clinical impact and predisposing factors for neonatal thrombocytopenia. *Indian J Pediatr* 1994;61:551.

98. Blanchette VS, Rand ML. Platelet disorders in newborn infants: diagnosis and management. *Semin Perinatol* 1997;21:53.

99. Andrew M, Castle V, Saigal S, et al. Clinical impact of neonatal thrombocytopenia. *J Pediatr* 1987;110:457.

100. Andrew M, Vegh P, Caco C, et al. A randomized, controlled trial of platelet transfusions in thrombocytopenic premature infants. *J Pediatr* 1993;123:285.

101. Blanchette VS, Kuhne T, Hume H, et al. Platelet transfusion therapy in newborn infants. *Trans Med Rev* 1995;9:215.

102. Johnson J-A, Ryan G, Al-Musa A, et al. Prenatal diagnosis and management of neonatal alloimmune thrombocytopenia. *Seminar Perinatol* 1997;21:45.

103. Payne SD, Resnik R, Moore TR, et al. Maternal characteristics and risk of severe neonatal thrombocytopenia and intracranial hemorrhage in pregnancies complicated by autoimmune thrombocytopenia. *Am J Obstet Gynecol* 1997;177:149.

104. Chalmers EA, Gibson BES. Clinical aspects of paediatric and perinatal transfusion: plasma products. *Vox Sang* 1994;67:54.

105. British Committee for Standards in Hematology. Guidelines for the use of fresh frozen plasma. *Transfusion Med* 1992;2:57.

106. Northern Neonatal Nursing Initiative Trial Group. Randomized trial of prophylactic early fresh-frozen plasma or gelatin or glucose in preterm babies: outcome at 2 years. *Lancet* 1996;348:229.

107. Zavadil DP, Stammers AH, Willett LD, et al. Hematological abnormalities in neonatal patients treated with extracorporeal membrane oxygenation (ECMO). *J Extra Corpor Technol* 1998;30:83.

108. Robinson T, Kickler T, Walker LK, et al. The effect of extracorporeal membrane oxygenation upon platelets in newborns. *Crit Care Med* 1993;21:1029.

109. McCune S, Short BL, Miller MK, et al. Extracorporeal membrane oxygenation therapy in neonates with septic shock. *J Pediatr Surg* 1990;25:479.

110. Meyer DM, Jessen ME. Results of extracorporeal membrane oxygenation in neonates with sepsis. The Extracorporeal Life Support Organization experience. *J Thorac Cardiovasc Surg* 1995;109:419.

111. Bjerke HS, Kelly Jr RE, Foglia RP, et al. Decreasing transfusion exposure risk during extracorporeal membrane oxygenation (ECMO). *Transfus Med* 1992;2:43.

112. Minifee PK, Daeschner 3rd CW, Griffin MP, et al. Decreasing blood donor exposure in neonates on extracorporeal membrane oxygenation. *J Ped Surg* 1990;25:38.

113. Rosenberg EM, Chambers LA, Gunter JM, et al. A program to limit donor exposures to neonates undergoing extracorporeal membrane oxygenation. *Pediatrics* 1994;94:341.

114. Ludvigsen Jr CW, Swanson JL, Thompson TR, et al. The failure of neonates to form red blood cell alloantibodies in response to multiple transfusions. *Am J Clin Pathol* 1987;87:250.

115. Strauss RG, Cordle DG, Quijana J, et al. Comparing alloimmunization in preterm infants after transfusion of fresh unmodified versus stored leukocyte-reduced red blood cells. *J Pediatr Hematol Oncol* 1999;21:224.

116. Rodwell R, Tudehope DI. Screening for cryptantigen exposure and polyagglutination in neonates with suspected necrotizing enterocolitis. *J Paediatr Child Health* 1993;29:16.

117. Holman P, Blajchman MA, Heddle N. Noninfectious adverse effects of blood transfusion in the neonate. *Trans Med Rev* 1995;9:277.

118. Wu T-J, Teng R-J, Yau K-IT. Transfusion-related acute lung injury treated with surfactant in a neonate. *Eur J Pediatr* 1996;155:589.

119. Luban NLC, DePalma L. Transfusion associated graft versus host disease in the neonate-expanding the spectrum of disease. *Transfusion* 1996;36:117.

120. Ohto H, Anderson KC. Posttransfusion graft-versus-host disease in Japanese newborns. *Transfusion* 1996;36:117.

121. BCSH Transfusion Task Force. Guidelines on gamma irradiation of blood components for the prevention of transfusion-associated graft-versus-host disease. *Transfusion Med* 1996;6:261.

122. Strauss RG. Red blood cell transfusion practices in the neonate. *Clin Pediatr* 1995;22:641.

123. Plonait SL, Nau H, Maier RF, et al. Exposure of newborn infants to di-(2-ethylhexyl)-phthalate and 2-ethylhexanoic acid following exchange transfusion with polyvinylchloride catheters. *Transfusion* 1993;33:598.

124. Karle VA, Short BL, Martin GR, et al. Extracorporeal membrane oxygenation exposes infants to the plasticizer, di (2-ethylhexyl) phthalate. *Crit Care Med* 1997;25:696.

125. Frederick T, Mascola L, Eller A, et al. Progression of human immunodeficiency virus disease among infants and children infected perinatally with human immunodeficiency virus or through neonatal blood transfusion. *Pediatr Infect Dis J* 1994;13:1091.

126. Jones DS, Byers RH, Bush TJ, et al. Epidemiology of transfusion-associated acquired immunodeficiency syndrome in children in the United States, 1981 through 1989. *Pediatrics* 1992;89:123.

127. Lieb LE, Mundy TM, Goldfinger D, et al. Unrecognized human deficiency virus type 1 infection in a cohort of transfused neonates: a retrospective investigation. *Pediatrics* 1995;95:717.

128. Schreiber GB, Busch MP, Kleinman SH, et al. The risk of transfusion transmitted viral infections. *N Engl J Med* 1996;334:1685.

129. Noble RC, Kane MA, Reeves SA, et al. Posttransfusion hepatitis A in a neonatal intensive care unit. *JAMA* 1984;252:2711.

130. Azimi PH, Roberto RR, Guralnik J, et al. Transfusion-acquired hepatitis A in a premature infant with secondary nosocomial spread in an intensive care nursery. *Am J Dis Child* 1986;140:23.

131. Rosenblum LS, Villarino ME, Nainan OV, et al. Hepatitis A outbreak in a neonatal intensive care unit: risk factors for transmission and evidence of prolonged viral excretion among preterm infants. *J Inf Dis* 1991;164:476.

132. Nelson SP, Jonas MM. Hepatitis C infection in children who received extracorporeal membrane oxygenation. *J Pediatr Surg* 1996;31:644.

133. O'Riordan JM, Conroy A, Nourse C, et al. Risk of hepatitis C infection in neonates transfused with blood from donors infected with hepatitis C. *Trans Med* 1998;8:303.

134. Tong MJ, el-Farra NS, Reikes AR, et al. Clinical outcomes after transfusion-associated hepatitis C. *N Engl J Med* 1995;332:1463.

135. Woelfle J, Berg T, Keller KM, et al. Persistent hepatitis G virus infection after neonatal transfusion. *J Pediatr Gastroenterol Nutr* 1998;26:402–407.

136. Yeager AS. Transfusion-acquired cytomegalovirus in infections in newborn infants. *J Pediatr* 1981;98:281.

137. Tiernay AJ, Higa TE, Finer NN. Disseminated cytomegalovirus infection after extracorporeal membrane oxygenation. *Ped Inf Dis J* 1992;11:241.

138. Przepiorka D, LeParc G, Werch J, et al. Prevention of transfusion transmitted cytomegalovirus infection-practice parameter. *Am J Clin Pathol* 1996;106:163–169.

139. Zanella A, Rossi F, Cesana C, et al. Transfusion-transmitted human parvovirus B 19 infection in a thalassemic patient. *Transfusion* 1995;35:769.

140. Dobroszycki J, Herwaldt BL, Boctor F, et al. A cluster of transfusion-associated babesiosis cases traced to a single asymptomatic donor. *JAMA* 1999;10:927.

141. Birchard K. Three countries to start leucocyte depletion of donated blood. *Lancet* 1998;351:1112.

74

Complications of Blood Transfusion

Jay E. Menitove

Adverse events following blood transfusion occur within a few hours, or after several days or weeks following transfusion. During the past decade, the risk of posttransfusion hepatitis and retroviral infections decreased and a better understanding of the pathophysiology of hemolytic transfusion reactions and other untoward sequelae emerged. Additional advances relate to the results of randomized, clinical trials aimed at determining whether transfusion, per se, increases the risk of postoperative bacterial infections or shortens the disease-free interval following tumor resection surgery. This chapter reviews the pathophysiology, clinical presentation, laboratory diagnosis, and treatment or prevention of the predominant complications of blood transfusion.

ACUTE TRANSFUSION REACTIONS

Part of "74 - Complications of Blood Transfusion"

Acute transfusion reactions (Table 74.1) occur minutes or hours after infusing red cell, platelet, or plasma components. These reactions involve immune, nonimmune, and infectious etiologies. The clinical presentations of these complications overlap significantly. For this reason, a laboratory evaluation is necessary for determining the specific cause and for ordering the appropriate therapeutic intervention.

TABLE 74.1. ACUTE TRANSFUSION REACTIONS

Reaction	Risk per Unit Transfused
Hemolytic transfusion reactions	~1:33,000
Fatal, hemolytic transfusion reactions	1:600,000
Febrile, nonfebrile transfusion reactions	
red cells	1:50–1:200
platelets	1:4–1:50
Allergic reactions	
red cells	1:200
platelets	1:30–1:100
Anaphylactic reactions	1:150,000
Hypervolemia	infrequent
Nonimmune hemolysis	infrequent
Air embolism	0 (per 9,000,000 observed)
Bacterial sepsis	
red cells	1:500,000
platelets	1:2,000–12,000
fatalities	1:3,000,000
Hypotensive reactions	?

Acute Hemolytic Transfusion Reaction

Hemolytic reactions caused 50% of the transfusion-related fatalities reported to the United States Food and Drug Administration (FDA) between 1990 and 1998 (1) (Table 74.2). Most of these reactions involved ABO incompatibility, in which group O recipients received blood from group A donors. Prevention of this feared complication of transfusion therapy serves as the mainstay of transfusion service policies and procedures that include proper identification of patients from whom blood samples are obtained; labeled requisition forms and tubes; compatibility testing; issuing blood components; ensuring the intended patient receives the appropriate blood component; monitoring the patient for signs and symptoms of an adverse reaction; and conducting an appropriate laboratory investigation if an adverse event occurs.

TABLE 74.2. TRANSFUSION RELATED FATALITIES: REPORTS TO FDA: 1990–1998

	Cases	%
Hemolysis	161	50
Bacterial contamination	46	14
Transfusion related acute lung injury	29	9
Nonbacterial infections	23	7
Transfusion associated graft-vs-host disease	18	6

Lee JH. FDA Workshop on Bacterial Contamination of Platelets. September 24, 1999.

In one study, based on mandated reporting of incidents, accidents, and errors to the New York State Health Department, approximately 1 per 33,000 red-cell transfusions were erroneously ABO-incompatible and an estimated 1 per 12,000 red cell units were infused into an unintended patient (2). The fatality rate was 1 per 600,000 units or 5% of incompatible transfusions. Approximately 43% of the hemolytic reactions resulted from an error involving patient identification, 11% because of a phlebotomist error, 25% as a result of laboratory error, and 15% occurred when administering or issuing the transfusion.

Pathophysiology

Transfusion of incompatible red cells triggers an antibody-antigen reaction and release of inflammatory response mediators, tumor necrosis factor-α (TNF-α), interleukins-1 and -8 (IL-1, IL-8), monocyte chemoattractant protein-1 (MCP-1), and anaphylatoxins C3a and C5a (3, 4). Hageman factor and complement may be activated. The latter is associated with intravascular hemolysis and release of red cell stroma and hemoglobin into the circulation. In the absence of complement activation, opsonized red cells are removed extravascularly within the reticuloendothelial system. In addition to anti-A and anti-B antibodies associated with intravascular hemolytic reactions include anti-K (Kell), anti-JKa (Kidd), and anti-Fya (Duffy). Those commonly associated with extravascular hemolysis include antibodies directed against Rh-antigens, anti-E, anti-C, and anti-D.

In vitro experiments indicate that cytokines play a significant role in the pathogenesis of hemolytic transfusion reactions. An IgM-mediated hemolysis model involves adding nongroup O red cells to group O whole blood, i.e., ABO-incompatibility. In an IgG-mediated hemolysis model, IgG-opsonized red cells are incubated with isolated mononuclear cells (3).

In the ABO-incompatibility experiments, TNF-α levels increased in a dose- and time-dependent pattern, with the maximal response at 2 hours. After 4 and 6 hours, IL-8 and MCP-1 levels increased and remained elevated for up to 24 hours. Complement appears necessary for cytokine production.

In the IgG-mediated red cell incompatibility experiments, IL-1, IL-6, and IL-8 levels increased within 6 hours and remained elevated for more than 24 hours. TNF levels increased transiently and did not reach levels associated with IgM-mediated hemolysis.

Clinical Presentation

Fever (elevation of at least 2°F or 1°C) or fever accompanied by chills occurs in almost all conscious or nonanesthetized patients suffering a hemolytic transfusion reaction. Nausea, vomiting, and chest pain occur less frequently. Wheezing and dyspnea, back pain, restlessness, and discomfort at the infusion site may occur. Additional clinical findings include hemoglobinuria, intravascular coagulation abnormalities, renal failure, and hypotension. In comatose or anesthetized patients, unsuspected hypotension, uncontrollable bleeding, or hemoglobinuria should alert clinicians to the possibility of a hemolytic reaction.

Disseminated Intravascular Coagulation

Two to six hours after group-A red cells are added to group-O whole blood *in vitro*, white cells generate procoagulant activity that presumably contributes to disseminated intravascular coagulation (DIC). Two mechanisms of action are postulated. Complement-binding antibodies such as anti-A or anti-B deposit C3b or iC3b on incompatible red-cell membranes that serve as ligands for white-cell CR3, that, in turn, stimulate monocytes tissue factor — generation. Alternatively or additionally, factor X is converted to factor Xa by monocytes through a mechanism requiring CD11b (a ligand for C3b and iC3b). In addition, TNF-α stimulates internalization of thrombomodulin by endothelial cells that exert a prothrombotic effect by suppressing protein-C activation (4).

Acute Renal Failure

Severe hemolytic reactions involving shock and DIC lead to renal hypoperfusion, fibrin deposition, and renal ischemia. Release of free hemoglobin and endothelial-derived vasoactive compounds has an additional role. Endothelin release and subsequent local vasoconstriction lead to parenchymal ischemia and renal failure. Likewise, circulating free hemoglobin induces vasoconstriction by binding nitric oxide with consequent renal ischemia and renal failure (4).

Vasomotor Instability/Hypotension

Hemolytic reactions activate Hageman Factor with subsequent bradykinin generation and hypotension. TNF-α release leads to capillary leak, neutrophil degranulation, and hypotension. In addition, anaphylatoxins C3a and C5a stimulate histamine release from cells that lower blood pressure.

Respiratory Failure

Although the lungs have not been regarded as a major target organ of hemolytic transfusion reactions, free hemoglobin and nitric-oxide interaction potentially leads to localized vasoconstriction with subsequent pulmonary vascular hypertension and decreased cardiac output. Neutrophil degranulation in pulmonary capillaries creates lung injury and procoagulant mediators lead to localized inflammation and thrombosis (4).

Laboratory Diagnosis

When a hemolytic transfusion reaction is suspected, the infusion must be stopped immediately and a post-transfusion blood specimen should be sent to the transfusion service laboratory (Fig. 74.1). A clerical check to confirm the patient's name, hospital number, and donor unit number should be undertaken quickly. The error may involve the cross match sample, the requisition forms, or the unit of blood. Because one or more units of blood may have been confused, mislabeled, or switched, it is important to determine immediately if another patient is at risk.

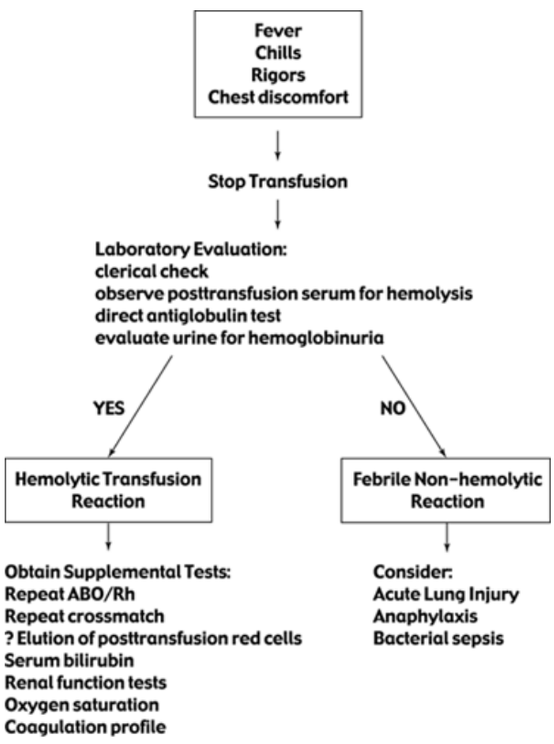


FIGURE 74.1. Algorithm for evaluating patients with fever, chills, rigors, or chest discomfort following a blood or blood-component transfusion.

The laboratory evaluation seeks evidence of hemolysis by visually observing the postreaction serum or plasma specimen for a pink or red hue. The direct antiglobin test is positive in almost all hemolytic transfusion cases. A urine specimen may be examined for the presence of free hemoglobin or, subsequently, for urinary hemosiderin.

If the clerical check, visual inspection, or direct antiglobin test support the diagnosis of a hemolytic transfusion reaction, additional testing should be conducted to determine the cause of the reaction and to provide baseline measurements for future monitoring. These tests include repeating the ABO and Rh typing on pre- and posttransfusion reaction samples and on the unit of blood. A crossmatch should be performed with prereaction and postreaction samples using red cells from the unit or from a stored, sealed segment. If the direct antiglobin test is positive, donor and patient sera should be evaluated for unexpected antibodies and consideration of eluting the antibody from the patient's cells. Plasma hemoglobin and serum haptoglobin measurements may be conducted. Posttransfusion hemoglobin and hematocrit results should be evaluated to determine if the expected

hemoglobin or hematocrit elevation occurred. The unconjugated bilirubin level should be obtained on a specimen drawn preferably 5 to 7 hours after transfusion. Renal function and oxygen saturation testing should be conducted to establish baseline information.

If hemolysis is suspected and an antibody-mediated etiology cannot be ascertained, a Gram stain of blood from the unit of blood may be evaluated, and the unit cultured.

Treatment

Treatment of hemolytic transfusion reactions includes immediate cessation of the transfusion and implementation of supportive therapy.

Monitoring vital signs, urine output, renal function, pulmonary function, coagulation status, and if necessary, cardiac filling pressures, provide data for determining needed therapeutic interventions. Hypotension should be treated with volume infusion as ascertained by invasive hemodynamic monitoring. Vasopressors such as dopamine are indicated if hypotension is not reversed by fluid infusion. Dopamine, when used in low doses, has a vasodilatory effect on renal vascular perfusion and may exert a beneficial effect on renal function. Osmotic diuretics such as mannitol and, other diuretic agents were used previously on the basis of anecdotal reports. The results of clinical findings and pulmonary function monitoring dictate the need for oxygen therapy, intubation, and ventilatory support. Treatment of DIC includes plasma, platelet, and cryoprecipitate transfusions and reversal of hypotension.

Treatment strategies under investigation involve agents aimed at inhibiting or reversing the pathophysiologic alterations (4).

Febrile Nonhemolytic Transfusion Reactions

Febrile, nonhemolytic transfusion reactions (FNHTRs) are characterized by a temperature increase of 1°C or greater during or within 1 hour after a transfusion in the absence of other medical causes of fever including a hemolytic reaction. Additional features of FNHTRs include chills, a cold sensation, rigors, or discomfort. Reports of FNHTRs vary from 0.5% to approximately 2% per red-cell unit transfused and from approximately 2% to 23% for platelet transfusions (5,6 and 7).

Pathophysiology

FNHTRs also called “chill-fever” reactions occur on the basis of at least two distinct mechanisms. Traditionally, FNHTRs result from immune recognition of infused leukocytes or platelets by recipient leukoagglutinin or HLA-antibodies. A second, more recently discovered mechanism, involves passive infusion of inflammatory cytokines generated by white blood cells contained in red cell or platelet components during storage (5,6 and 7).

FNHTRs related to antibodies in the recipient's plasma occur predominantly in patients with a history of previous transfusions or prior pregnancies. Recipient cytotoxic, agglutinating, or HLA-antibodies react against antigens on infused granulocytes, lymphocytes, or platelets. The “threshold” leukocyte count evoking “chill-fever” reactions in susceptible patients varies from 2×10^8 to 2.5×10^9 white cells per component. Most involve reactions against granulocytes. The amount, rather than the type of antibody, the number and subset of leukocyte transfused, and the transfusion rate correlate with reaction severity.

FNHTRs corresponding to passive cytokine infusion relate predominantly to platelet transfusions. Following platelet concentrate preparation in which white cells are not removed, IL-I, IL-6, IL-8, and TNF-α levels increase during storage. Passive infusion of IL-6 and other cytokines present in the plasma supernatant of the platelet concentrates presumably causes one half of the “chill-fever” reactions. Low levels of residual white cells account for the remaining reactions (6, 7). Prestorage leukocyte-reduction reduces the frequency of severe FNHTRs by 50% compared to nonleukocyte reduced platelet transfusions (7).

Clinical Presentation

Classically, FNHTRs consist of an immediate transient reaction marked by a flush, palpitation, tachycardia, cough, chest discomfort, or neutropenia. Following a latent interval of approximately 15 to 60 minutes, the diastolic blood pressure increases and headache, chillness, or rigors occur. Subsequently the temperature rises, apathy, irritability, impaired mentation, and prostration follow and may persist for several hours. A neutrophilic

leukocytosis with a shift to the left occurs 2 to 5 hours following infusion.

FNHTRs complicate and interrupt transfusion therapy, because fever may be the presenting sign of a hemolytic reaction. Onset of fever and/or chills should result in cessation of the transfusion and a subsequent laboratory evaluation to determine whether hemolysis occurred. The signs and symptoms of a FNHTR usually are self-limited. Fortunately, more severe consequences are infrequent.

Laboratory Diagnosis

Despite advances in understanding the etiology of FNHTRs, the diagnosis is one of exclusion, and is made by demonstrating the absence of hemolysis during the evaluation of a possible hemolytic transfusion reaction.

Treatment

Currently, leukocyte-reduction by filtration conducted, preferably, prior to red-cell or platelet storage serves as a preventive measure. Previously used methods of leukocyte-reduction such as buffy-coat removal and freezing and washing are significantly less effective than filtration with specially designed leukocyte-reduction filters.

Treatment of reactions involves symptomatic therapy with nonaspirin antipyretics. Severe rigors respond to meperidine.

Transfusion-Related Acute Lung Injury

Transfusion-related acute lung injury (TRALI or noncardiogenic pulmonary edema, pulmonary “hypersensitivity,” or severe allergic pulmonary edema) is a clinical constellation of an adult respiratory distress–like syndrome that occurs within hours of a blood transfusion (8, 9).

Affected patients develop fever, substernal chest pain, marked dyspnea, cyanosis, cough, blood-tinged sputum and hypoxemia, hypotension, and bilateral pulmonary edema within 4 hours (usually 1 to 2 hours but up to 6 hours) after transfusion. TRALI has been reported following infusion of whole blood, red blood cells, and fresh frozen plasma, as well as granulocytes, cryoprecipitate, platelet concentrates, and apheresis platelets. TRALI occurs at a frequency of approximately 1 per 5,000 units infused or from 0.04% to 0.16% per patient transfused. The mortality rate is approximately 5%.

Pathogenesis

Granulocyte and/or lymphocyte antibodies are present in the plasma of 50% to 90% of infused components. These antibodies include anti-NA2, anti-5b, anti-NB2, anti-NB1, and anti-HLA antibodies.

The antileukocyte antibodies, when infused into a recipient with corresponding leukocyte antigens, result in white-cell pulmonary sequestration. Complement activation occurs in TRALI animal models. Lung injury results from release of granulocyte contents, an increase in vascular permeability, and fluid and protein exudation into the alveoli.

TRALI does not universally follow infusion of antileukocyte antibodies and leukocyte antibodies are not detected in all cases of TRALI. These observations led to an alternative TRALI hypothesis that involves a “two-hit” model (10). Patients have a predisposing condition such as recent surgery, active inflammation, or infection, and then a second event, such as an infusion of biologically active lipids present in stored blood components leads to neutrophil priming and adherence to endothelial cells. However, the “two-hit” model is inconsistent with clinical observations because TRALI occurs at a relatively low frequency among patients receiving stored blood components with the implicated conditions.

Clinical Presentation

The clinical presentation resembles the adult respiratory distress syndrome and occurs shortly after blood transfusion. Less frequently, the syndrome is delayed for several days after transfusion. The severity relates primarily to the extent of hypoxemia. TRALI is a form of noncardiogenic pulmonary edema and is distinguished from other forms of pulmonary edema on the basis of normal to decreased pulmonary capillary wedge pressures; normal pulmonary artery pressure; absence of jugular venous distention, murmurs or gallops; normal cardiac silhouette on chest radiograph; absence of pulmonary vascular congestion; and no evidence of myocardial infarction.

The protein content of TRALI pulmonary edema fluid is elevated compared to cardiogenic pulmonary edema. The ratio of protein in edema fluid compared to protein in blood usually is greater than 0.7 in noncardiogenic edema and less than 0.5 in cardiogenic pulmonary edema. This relates to increased vascular permeability. Arterial oxygen saturation is decreased, acidosis may occur, and hemoconcentration has been observed.

Treatment

Intubation and mechanical ventilatory assistance is required in as many as three-fourths of patients. TRALI usually responds to supportive therapy within 24 to 48 hours. Pulmonary infiltrates resolve in 80% of patients and arterial blood-gas measurements return to baseline within 96 hours.

Diuretic therapy has been associated with further decreases in pulmonary capillary wedge pressures, cardiac output, and hypotension. Instead these parameters are more likely to improve with saline infusion.

Laboratory Finding

The diagnosis is based on clinical presentation. The laboratory investigation involves detecting antileukocyte antibodies. This usually proceeds in a step-wise fashion in which donors with predisposing conditions for antibody formation are identified and tested for anti-HLA and antigranulocyte antibodies. This includes women with prior pregnancies and donors with a history of blood transfusion. When donors with antibodies are identified, a crossmatch should be performed between donor serum and recipient white cells, if feasible. A positive crossmatch confirms the diagnosis, but a negative result does not rule out the diagnosis.

Future blood donations from implicated donors, if used, should avoid transfusion of components containing plasma. Alternatively, the donor should be deferred from future blood donation.

Allergic Reactions

Allergic reactions, manifest by urticaria and pruritus, occur at a frequency of approximately 0.5% per red-cell transfused and approximately 1% to 3% per platelet transfusion presumably reflecting the lower plasma volume in red blood cell transfusions compared to platelets (5,6 and 7). Severe anaphylactic reactions occur in IgA deficient transfusion recipients who have class-specific anti-IgA. The estimated frequency of transfusion-associated anaphylactic reactions is 1 per 150,000 units and ranges from 1 per 47,000 transfusion to 1 per 770,000 units (11).

Urticaria

Hives and pruritus, usually in the absence of fever, result presumably from a reaction between a protein in infused plasma and a corresponding IgE antibody in the recipient. Significant morbidity results if hives occur in anatomically compromised positions such as the larynx and requires prompt intervention. Otherwise, routine treatment involves stopping the transfusion, administering an antihistamine, waiting for the pruritus to subside, and in most instances continuing the transfusion. Because urticaria is not a sign of hemolysis, laboratory investigation for a hemolytic transfusion reaction is unnecessary. If hives do not respond to symptomatic therapy, recur frequently following subsequent transfusions, or cause significant clinical problems, plasma should be removed from future red cell and platelet transfusions by washing with saline solutions.

Anaphylactic Transfusion Reactions

An immediate, generalized reaction during a blood transfusion should raise suspicion of a hemolytic or an anaphylactic reaction. Anaphylactic reactions occur immediately or may be delayed for as long as 1 hour after exposure. Other causes of anaphylaxis including exposure to allergens such as drugs, foods, or latex gloves should be considered. Signs and symptoms of IgA anaphylactic reactions include skin manifestations (hives, rash, pruritus, angioedema), respiratory findings (dyspnea, stridor, wheezing), cardiovascular events (hypotension, syncope, arrhythmia, shock), and gastrointestinal complications (cramps, diarrhea, vomiting). Anaphylactic reactions have been reported following infusion of whole blood, red cells, plasma, platelets, cryoprecipitate, and IV-gamma globulin.

Anaphylactic reactions may occur in previously untransfused patients and in patients transfused previously without complications. The clinical diagnosis requires laboratory confirmation. Because most affected patients are IgA deficient (serum IgA less than 0.05 mg/dL) and have class-specific anti-IgA, the laboratory evaluation consists of testing for IgA levels and the presence of anti-IgA antibodies (11). IgA deficiency in blood donors occurs at a frequency of approximately 1 per 400. One in 1,200 blood donors are IgA-deficient and also have class-specific anti-IgA. This frequency far exceeds that of observed anaphylactic reactions suggesting that detection of anti-IgA does not predict a risk for anaphylactic transfusion reactions.

Treatment of anaphylactic reactions includes immediate cessation of the transfusion and administration of epinephrine and other supportive therapy.

Further blood transfusions should be delayed until a pretransfusion sample is tested for IgA concentration. If IgA is present, consideration of other causes should be entertained. If transfusions cannot be withheld, red cells or platelets washed with large amounts of saline may be infused. Fresh frozen plasma and cryoprecipitate should be obtained from IgA deficient donors. If IgA test results indicate detectable IgA levels, the likelihood of an anaphylactic reaction related to IgA infusion is considered low and washed components or components from IgA deficient donors generally are not recommended.

Red Eye Syndrome

A cluster of adverse ocular reactions following transfusion occurred in late 1997 through early 1998 (12). Patients experienced severe bilateral conjunctival erythema and/or hemorrhage, eye pain, headache, periorbital edema, arthralgias, nausea, dyspnea, and rash within 24 hours of transfusion. The symptom complex resolved within 5 days (range 2 to 21 days). All of the patients received blood components subjected to leukocyte-reduction filtration using filters from the same manufacturer. The filters were recalled and no subsequent reactions were reported. A toxic or allergic reaction was suspected. A decade previously, severe reactions during platelet apheresis collections were attributed to allergic reactions to the sterilization agent, ethylene oxide gas (13).

Hypervolemia

Patients with impaired myocardial or renal function are at risk of hypervolemia and heart failure. An average unit of whole blood contains 56 mEq sodium, a unit of red blood cells has 8 to 20 mEq sodium, and a unit of red cells with an additive solution has 24 to 30 mEq sodium. With the exception of acute blood loss situations, infusion rates should be 2 to 4 mL/kg/hour but reduced to 1 mL/kg/hour in patients known to be at risk for hypervolemia.

Nonimmune Hemolysis

Transfused red cells subjected to osmotic, mechanical, or thermal stress may burst and release hemoglobin into the circulation. Hypotonic solutions, such as 25% albumin diluted with sterile water, or medications added to blood may lead to osmotic rupture (14). Only normal saline or agents shown to be safe and to not adversely affect the blood component should be added.

Overheating of red cells by blood warmers may result in red cell destruction; only devises with alarms and found in compliance with the FDA should be used.

Infusion pumps should be selected that do not disrupt red cells or have tubing that interferes with red cell flow.

Of note, hemolysis has been reported during cell salvage

procedures in association with turbulence at the air/fluid interface.

A negative laboratory evaluation for immune hemolysis in conjunction with an investigation of the factors associated with hemolysis should provide a cause for the nonimmune hemolytic event and provide information for appropriate corrective action.

Air Embolism

Air embolism is considered by most as an archaic complication associated with vented glass bottles and is not considered to occur with modern plastic containers. Observations by a state health department with mandatory error, accident, and severe adverse reaction reporting support this presumption (15). No cases of air embolism were associated with 9 million routine transfusions. However, fatal air embolism occurred at a rate of 1 per 30,000 intraoperative blood recovery procedures and 1 per 38,000 postoperative blood recovery procedures between 1990 and 1995. In each of the fatal cases, blood recovery and subsequent infusion occurred with a pressure cuff or a pressure infusion devise applied to the bag. Using a model system constructed to stimulate operative blood recovery and pressurized readministration, 200 mL air were infused within 4 seconds; an amount considered fatal in humans.

Techniques for preventing air embolism involve avoidance of direct reinfusion from red cell recovery containers, limiting the reinfusion from the recovery containers, limiting the reinfusion of blood under pressure, and maintaining vigilance.

Symptoms of air embolism include chest pain, cough, and acute onset of dyspenea. The infusion should be stopped immediately by clamping the administration tubing. The patient should be placed on the left side in a head down position and appropriate surgical consultation to remove air from the right atrium and pulmonary artery should be obtained.

Complications of Massive Transfusion

Massive blood loss is defined as replacement of a patient's total blood volume, transfusion of more than 10 units of blood within 24 hours, replacement of 50% of the circulating blood volume in less than 3 hours, or transfusion of more than 150 mL per minute. Both the volume and the time period are relevant to the clinical urgency and the metabolic and coagulation consequences of massive transfusion (16).

Blood bank considerations of massive transfusion include logistic concerns about providing multiple blood components for one or more patients urgently, coagulation abnormalities arising from massive blood replacement, metabolic consequences of infusing large amounts of anticoagulated, stored blood components, and the effect of rapidly infusing large amounts of relatively cold fluids.

Logistical Issues

Urgent situations heighten the probability of misidentifying patient or patients, blood samples, and blood components. Preparing and implementing massive transfusion protocols by multidisciplinary teams should decrease mishaps involving identification errors in emergency situations. Use of group O red cells in lieu of “type-specific” blood and decisions about use of Rh-negative red cells for emergency release should be decided during the planning stages. The time needed to transfer red cells from the transfusion service to the patient, thaw frozen plasma, and prepare platelet transfusions should be part of the discussions with the multidisciplinary team. The ability to restock necessary blood components should be considered by hospital transfusion services in determining the routine level of components stored in hospital inventory.

Hemostasis

Stored red cells are devoid of functional platelets and have minimal plasma and coagulation factor content. Platelet concentrates prepared from whole blood contain an equivalent of one unit of plasma. Factor VIII levels fall to 50% of initial levels after one day of storage and further loss occurs during the 5-day storage period. Factor V levels fall to 50% of initial activity 14 days after storage.

Following massive blood loss, blood volume is expanded by infusing crystalloid solutions. Infusion of red cells lacking platelets and coagulation factors leads to dilution of these hemostatic elements. In addition, platelets, fibrinogen, and factors V and VIII are consumed during the coagulation process. Dilutional thrombocytopenia develops during massive transfusion; platelet counts should be monitored. Platelet transfusion, one unit per 12 kg body weight or one single donor (apheresis) platelet unit should be given to maintain the platelet count above 50,000/uL or 100,000 u/L in patients with central nervous systems injury. The prothrombin time and partial thromboplastin times should be monitored, and, in general, plasma, 10 to 15 mL/kg body weight, should be given when these parameters exceed 1.5 times control values. The fibrogen concentration should be maintained above 75 to 100 mg/dL by infusing one unit cryoprecipitate per 5 kg body weight (16).

Metabolic Factors

Patients suffering massive blood loss are at risk for hypoperfusion with resultant acidosis and hyperkalemia. Red blood cells are stored at acidotic pH ranges and have elevated potassium levels following storage. Red cells containing the anticoagulant/preservation solution CPDA-1 contain 33.2 g sodium citrate compared to only 5.31 g in red cells containing the additive solution AS-1 (17). Reperfusion reverses systemic acidosis, citrate is metabolized by the liver, muscle, and kidneys to bicarbonate with resultant hydrogen ion binding, and stored red cells reabsorb potassium when returned to the circulation. Citrate metabolism is impaired by liver disease and hypothermia. In a healthy 70-kg adult, infusion of two units of citrated whole blood at a rate of 2 mL/kg/minute results in citrate toxicity with circulatory depression, decreased cardiac output, reduced pulse pressure, and hypotension. Citrate also binds magnesium, which can result in hypomagnesemia. Infusion of calcium gluconate (or magnesium) in practice, may be indicated for treatment of severe reactions. Usually, the outcome in massive transfusion is alkalosis and hypokalemia. Infusion of glucose in anticoagulated red cells and low doses of mannitol contained in AS-1 solutions does not result

in hyperglycemia or clinical diuresis. Red cell 2,3 DPG levels diminish during storage but the levels are restored within 8 to 12 hours following transfusion.

Hypothermia

Infusion of cold blood and intravenous fluids lowers body temperature. Hypothermia slows citrate metabolism, potentiates the harmful myocardial depressant effects of hypothermia and hypocalcemia, and reduces oxygen release from red cells. Blood warmers are used, in massive transfusion situations, to prevent these complications.

Bacterial Sepsis

Bacterial contamination of blood and blood components causes 14% of transfusion related fatalities reported to the FDA or 5 to 10 deaths per year in the United States (1). Overall, one death occurs per 3 million units transfused. Of the 86 bacterial contamination-associated fatalities reported to the FDA between 1976 and 1998, 58 occurred among platelet transfusion recipients and 28 among red cell recipients. Clinical sepsis related to red-cell and platelet transfusion is estimated to occur 10 to 100-fold more frequently than fatal reactions.

Pathophysiology

Bacterial contamination of red cell and platelet components occurs as a result of bacterial entry through skin plugs at the time of phlebotomy, transient bacteremia during blood or platelet collection, contaminated containers, minute holes in blood containers, or nonsterile pooling or aliquoting. Bacteria may contaminate bag ports when frozen components are thawed in a water bath (18).

Microorganisms contaminating platelet concentrates include *Staphloccus epidermidis*, *Escherichia coli*, *Bacillus* species, *Salmonella* species, *Protease mirabilis*, *Klebsiella* species, *Staphloccus aureas*, *Serratia* species, *Pseudomonas* species, *Enterobacter*, and *Streptococcus*.

Microorganisms contaminating red cells or whole blood include *Yersinia enterocolitica*, *Clostridium perfringens*, *Pseudomonas acnes*, *Enterococcus*, *Enterobacter*, *Serratia*, and *Klebsiella* (18, 19, 20).

Assuming bacteria enter the containers at the time of phlebotomy, the bacteria enter log phase growth after several days in room temperature stored platelets (Fig. 74.2). *In vitro* studies indicate platelet concentrates stored for 1 to 4 days had 2.6×10^3 to 1.0×10^4 colony-forming units per milliliter compared to 1.2×10^7 to 4.4×10^{13} colony-forming units in concentrates stored for 5 days (21). The contamination rate varies between 0.04% and 0.3% per unit for platelet concentrate prepared from whole blood and from 0% to 0.49% per apheresis (single donor) platelet concentrate.

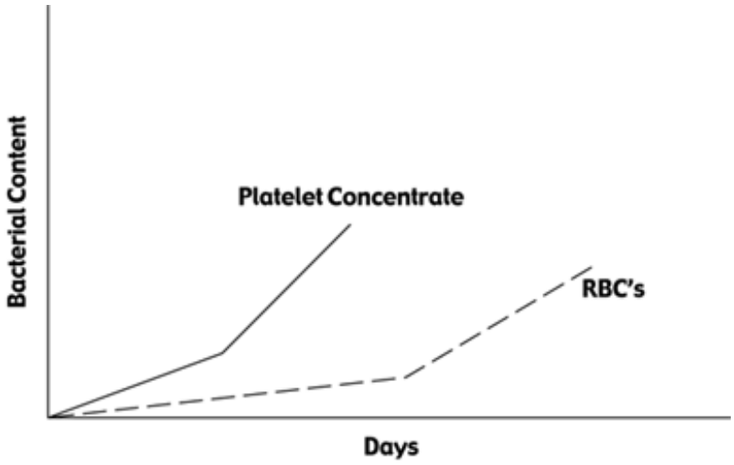


FIGURE 74.2. Bacteria proliferate more rapidly in platelet concentrates (stored at room temperatures) compared to red blood cells (stored at refrigerated temperatures).

Y. enterocolitica causes the majority of septic reactions associated with red cell transfusions. Fourteen deaths caused by *Yersinia*-contaminated red cells were reported to the FDA between 1976 and 1998. Because red cells are stored at 1 to 6°C, psychrophilic organisms such as *Yersinia* and *Pseudomonas* are the most likely pathogens. Assuming bacterial entry occurs at phlebotomy, an approximate 3-week lag period ensues prior to significant bacterial proliferation and accumulation of endotoxin sufficient to result in clinical symptoms.

Clinical Presentation

Criteria for suspecting a septic transfusion reaction include temperature elevations of 1 to 3.5°F or 1 to 2°C accompanied by rigors, tachycardia, and greater than 30 mmHg decreases or increases in blood pressure. Additional clinical signs or symptoms include tachycardia, nausea and vomiting, lumbar pain, shortness of breath, diarrhea, chest pain, and cyanosis.

Laboratory Diagnosis

The container and remaining blood should be evaluated for bacterial contamination by Gram stain and microbiological culture. A Gram stain is most likely to reveal bacteria when more than 10^6 colony forming units per milliliter are present. An acridine orange stain may detect bacteria at 10^5 colony-forming units per milliliter.

Investigation of biochemical parameters associated with bacterial growth are in progress to detect and thereby interdict of infected components. Included are reduced glucose and pH measurements, loss of platelet swirling, color changes, and evolved gases. Additional parameters under study include ribosomal RNA detection and bacterial culturing following storage but prior to release (18).

Treatment

When septic reactions are suspected, the transfusion must be discontinued immediately and should not be restarted. Supportive care and broad-spectrum antibiotic coverage should be instituted promptly. The diagnosis should be sustained despite a negative Gram stain, because this test reveals bacteria in only two thirds of symptomatic cases.

Hypotensive Reactions Following Transfusion

Severe hypotensive reactions following passage of red cells, plasma, or platelets through bedside leukocyte reduction filters have been reported in patients receiving angiotensin-converting enzyme (ACE) inhibitor medications. Severe hypotension occurs within minutes of starting the transfusion and may be accompanied by skin flushing and loss of consciousness. The incidence is unknown but considered infrequent.

The pathophysiology appears related to bradykinin is generation as plasma passes over leukocyte-reduction filters. This vasodilatory, nonapeptide has a plasma half-life of less than 30 seconds. The breakdown of its active metabolite, des-Arg^{9}-bradykinin, is significantly prolonged in the presence of ACE inhibitors (22).

Hypotensive reactions occur more frequently following platelet transfusions than red-cell transfusions, presumably because more plasma is present in platelets than red cells. The reactions have been reported following filtration with negatively charged and positively charged filters. Not all of the reported events involve patients taking ACE inhibitors. However, the hypotensive reactions occurred exclusively in patients receiving components leukocyte-reduced with bedside filters and not with prestorage leukocyte reduction or when filtration was performed in the laboratory. These findings relate, probably, to the short half-life of bradykinin and its metabolites and their degradation prior to infusion.

Leukocyte reduction performed prior to storage should obviate these reactions. Treatment may require intravenous vasopressor agents.

DELAYED TRANSFUSION REACTIONS

Part of "74 - Complications of Blood Transfusion"

In addition to immediate adverse consequences of blood transfusion, some untoward events occur days to years after transfusion (Table 74.3). Nonacute immunologic complications include delayed hemolytic reactions, bystander hemolysis in patients with sickle cell anemia, alloimmunization, graft-versus-host disease (GvHD), and immunosuppressive effects of blood transfusion. Nonimmune complications involve iron overload and transfusion-transmitted diseases.

TABLE 74.3. DELAYED TRANSFUSION REACTIONS

	Risk per Unit Transfused
Delayed serologic and hemolytic transfusion reaction	
serologic	1:1600
hemolytic	1:6700–1:9100
“Bystander hemolysis”	?
Alloimmunization	
red cells	
hospitalized patients	1%–2%
multiply transfused patients	10%
platelets*	
HLA-antibodies	
nonleukocyte-reduced	˜ 45%
leukocyte-reduced	˜ 18%
platelet specific antibodies	6%–11%
refractoriness	
nonleukocyte-reduced	7%
leukocyte-reduced	16%
Graft-vs-host disease	infrequent
Iron overload	after ˜80 units RBC's
Post transfusion purpura	rare
Immunosuppressive effects	?

*TRAP Study results[30]

Delayed Serologic and Hemolytic Transfusion Reactions

Delayed hemolytic transfusion reactions refer to destruction of transfused red cells days to weeks after transfusion by an immune response against recently infused foreign antigens. A delayed serologic transfusion reaction involves production of antibodies against recently infused foreign alloantigens and possibly autoantibodies in the absence of clinical hemolysis.

Diagnostic criteria defining delayed hemolytic/serologic transfusion reactions require that three of the following five criteria are met: (i) compatible crossmatch with pretransfusion serum; (ii) negative antibody screen using recipient's pretransfusion serum; (iii) incompatible crossmatch and a rising titer of the incriminated antibody in postreaction serum samples; (iv) history of blood transfusion, pregnancy, or both; and (v) clinical or laboratory evidence of hemolysis. The incidence of delayed serologic transfusion reactions is approximately 1 per 1,600 red-cell units transfused, and the incidence of delayed hemolytic transfusion reactions ranges from 1 per 6,700 to 1 per 9,100 units on the basis of reports released during the 1990s (23, 24).

Of note, persistence of positive direct antiglobulin tests, long after the infused cells would be expected to have been removed from the circulation, suggests autoantibodies are produced or nonspecific uptake of alloantibodies occurs in patients with delayed hemolytic or serologic reactions (23).

Clinical Presentation

The onset of anemia, fever, and jaundice approximately one week after a transfusion should raise suspicion about a delayed hemolytic transfusion reaction. Fever, at least 2°F (1°C) occurs in up to 75% of patients with delayed hemolytic reactions. Decreases in the hematocrit or hemoglobin occur in almost all of the patients, the bilirubin concentration increases in up to two thirds, and approximately one sixth of patients develop oliguria or an increase in serum creatinine levels. Red-cell destruction occurs 3 to 21 days after transfusion. For most patients, this complication is relatively benign and does not involve the severe reactions associated with an acute hemolytic reaction. However, increased levels of proinflammatory mediators have been observed in the setting of a delayed hemolytic transfusion reaction (25).

Laboratory Diagnosis

All patients with delayed hemolytic/serologic reactions have a positive direct antiglobulin test with anti-IgG reagents and many show reactivity for complement. In two-thirds to three-fourths of cases a single antibody is involved, in the others multiple antibodies have been found through elution studies. Antibodies against Rh phenotypes E, C, e, D, and Cw and antibodies against Kell(K), Kidd (Jka), and Duffy (Fya and Fyb) are involved predominantly. Previously, anti-Jka and anti-Fya antibodies were considered more likely to result in hemolytic reactions rather than isolated serologic findings. However, studies conducted in the mid-1990s show the ratio of hemolytic: serologic reactions decreased from that observed a decade previously. Explanations for this observation include a decrease in the average length of hospital stay in which Fya sensitization would not be reported and use of more sensitive serologic test reagents that detect previously undetected, nonhemolytic Jka antibodies (24).

In many patients, the direct antiglobin test remains positive for 6 months or longer posttransfusion. Eluates from the red cells demonstrate an alloantibody with specificity against transfused red cells or a panagglutinin (23).

Treatment

Treatment rarely is needed because of the relatively benign nature of the hemolytic event. Supportive care is indicated only for serious complications.

Severe Reactions Following Transfusion in Patients with Sickle Cell Disease

Alloimmunization against red cell antigens occurs in approximately 25% of patients with sickle cell disease. Antibodies form, most frequently against E, C, K, and against Fya, Jkb, S, and D antigens. In addition, delayed reactions occur as a result of antibodies against Jka, Fyb, and s antigens.

Pain crises, posttransfusion hemoglobinemia, pulmonary infiltrates, and disseminated intravascular coagulation have been reported after exchange transfusions in patients with sickle cell disease (26,27 and 28). In many cases the direct antiglobulin test is nonreactive. The etiology of these events is not understood completely and, conflicting hypotheses are under consideration. These include suppression of endogenous erythropoiesis and the subsequent fall in hemoglobin concentration, and “hyper-hemolysis” in which transfused red cells are destroyed and endogenous red cell production decreases.

Alternatively, “bystander-hemolysis” occurs in which presumably innocent, bystander cells undergo lysis as a result of a non–red-cell alloantibody-antigen reaction that causes complement activation. Antibody reactions against infused alloantigens, such as HLA and plasma protein antigens, may cause complement activation. Sickle cells have high affinity for complement activation in that the densest sickle cells have increased C5b-7 and C9 binding and subsequent increased susceptibility to C5b-9 mediated lysis. This parallels findings of complement on red cells involved in delayed hemolytic transfusion reactions presumably opsonized by alloantibodies not usually associated with complement binding (Rh system antibodies). As such, the postulated mechanism for “bystander hemolysis” may occur during any delayed hemolytic transfusion reaction.

Several of the reported patients developed profound, life-threatening anemia; transfusions did not result in sustained increases in hematocrit or hemoglobin levels. One approach to treatment involved discontinuing transfusions and prescribing a short course of corticosteroid therapy.

Alloimmunization

Red Cell Transfusions

One to 2% of hospitalized patients have clinically significant red cell alloantibodies. The frequency increases to approximately 10% among transfused patients and approximately 25% in patients with sickle cell anemia. Red cell allo antibodies develop 2 to 24 weeks post transfusion (29).

Compatibility testing prior to transfusion is designed to detect clinically significant alloantibodies. Frequently encountered alloantibodies are directed against, Rh, Kell, Duffy, and Kidd antigens.

Platelet Transfusions

The largest, randomized clinical trial investigating alloimmunization and refractoriness to platelet transfusions, the Trial to Reduce Alloimmunization to Platelets (TRAP), demonstrated that leukocyte reduction by filtration reduced alloimmunization rates in patients with acute myelocytic leukemia (30). Lymphocytotoxic HLA-antibodies were detected in 45% and 18% of patients randomized to receive non-leukocyte–reduced and leukocyte-reduced platelet concentrates. Refractoriness to platelet transfusions was reduced significantly from 16% to 7%, respectively. Antibodies against platelet-specific antigens occurred in 11% of those receiving non-leukocyte–reduced platelets compared to 6% of those receiving leukocyte-reduced platelets; a nonstatistically significant difference. These findings confirmed other reports showing that leukocyte reduction to levels below 5×10^6 white blood cells per transfusion effectively prevents alloimmunization and refractoriness to platelet transfusions.

Graft-versus-Host Disease

Graft-versus-host disease (GvHD) occurs after bone marrow transplantation, presumably by engraftment in recipients of donor-derived immunocompetent donor T lymphocytes.

In transfusion-associated GvHD, bone marrow involvement occurs in addition to skin, liver, and gastrointestinal tract dysfunction seen in marrow-related GvHD. Fever, erythematous maculopapular rashes, anorexia, vomiting, right upper-quadrant pain, cough, and diffuse diarrhea occur. Liver function enzymes, bilirubin, and alkaline phosphatase are elevated. The hepatocellular damage and elevate transaminase levels reflect liver damage in transfusion-associated GvHD in contrast to marrow transplant GvHD where biliary damage and obstruction predominate. Pancytopenia reflects marrow involvement (31).

The pathophysiology includes multidirectional interaction of CD4+ and CD8+ T cells activated by autologous and allogeneic events, cytotoxic antibodies, elaboration of cytokines IL-1, IL-2, and TNF. Host HLA antigens serve as targets of donor effective cells. Different CD4+/CD8+ ratios and Th-1 and Th-2 subpopulations occur in patients with varying underlying disorders (32).

Fever occurs 4 to 23 (median 10) days after the implicated transfusion. A rash develops 2 days later and may evolve to generalized erythroderma, bulla formation, and epidermolysis. Leukopenia occurs 11 to 31 (median 16) days after transfusion. Hepatomegaly and jaundice occur frequently, but lymphadenopathy and splenomegaly occur less often. The mortality rate is approximately 90% at a median 21 days after transfusion. Death usually results from infection or hemorrhage secondary to marrow aplasia and pancytopenia.

The diagnosis is clinically supported by finding donor cells in affected tissues, HLA DNA typing of the donors and the patient, or demonstrating Y chromosome regions in female patients.

An estimated minimal infusion of 10^4 lymphocytes per kilogram of recipient weight results in transfusion-associated GvHD. Implicated blood components include unirradiated whole blood, red blood cells, platelets, granulocytes, and fresh, nonfrozen plasma.

Transfusion-associated GvHD has been reported in immunocompetent patients including those undergoing cardiovascular surgery, cholecystectomy, tumor resection surgery, and an otherwise healthy pregnant woman with mild preeclampsia. These cases may reflect the frequency of transfusions from HLA-homozygous donors to HLA-heterozygous recipients who share a haplotype with the donor. The estimated risk of this event in United States whites is 1 in 17,700 to 1 in 39,000. If the donation involves parents and children (directed donations), the rate is increased 21-fold (33).

Most reports of transfusion-associated GvHD involve patients with impaired cellular immunity, hematologic malignancies, or immaturity (34). Generally accepted categories of patients considered at high risk of transfusion-associated GvHD include bone marrow transplant recipients, congenital cellular immunodeficiency syndrome patients, intrauterine transfusion recipients, granulocyte transfusion recipients, patients with Hodgkin's Disease or nonHodgkin's lymphoma, patients with certain solid tumors such as neuroblastoma and glioblastoma, and patients receiving HLA-matched components or components from biologic relatives. A consensus has not been reached about the degree of risk for patients with hematologic malignancies such as leukemia, term newborns on extracorporeal membrane oxygenators, organ transplant recipients, or patients receiving crossmatch compatible platelet transfusions. Most agree that there is no added risk for patients with acquired immunodeficiency syndrome (AIDS) or HIV infection, most patients receiving chemotherapy, patients with aplastic anemia not receiving immunosuppressive therapy, and full term infants without other risks.

Prevention is the best treatment modality and is accomplished by subjecting cellular blood components intended for patients at risk of transfusion-associated GvHD to 2,500 cGy gamma irradiation to the central midplane of a cannister with a minimum 1,500 cGy irradiation elsewhere. Irradiated red cells undergo an enhanced potassium efflux during storage. This may pose problems for neonates receiving transfusions. Cell washing or irradiation immediately prior to transfusion should be considered for these patients. Irradiated red cells are stored for the lesser of 28 days or the maximum storage interval (35). Use of photochemical inactivation using psoralens and UV-A irradiation to prevent transfusion-associated GvHD is under investigation (36).

Iron Overload

Manifestations of iron overload include increased skin pigmentation, hepatic disease, diabetes mellitus, cardiac dysfunction, arthropathy, gonadal insufficiency, and other endocrine disorders. In hereditary hemochromatosis, symptoms of organ damage develop when iron stores increase from the normal range of less than 1 gram iron to 15 to 20 grams or more. Further increments may be fatal, but some patients tolerate body stores of 40 to 50 grams iron. Transfused patients not receiving iron chelation therapy have cardiac and possibly liver, pancreas, and endocrine organ iron deposition when they have received 100 units of blood, 20 to 25 grams iron based on an average of 200 to 250 mg iron per unit of red cells (37).

Plasma ferritin and iron-binding capacity are used to measure body iron status. However, ineffective erythropoiesis may produce elevations of plasma ferritin unrelated to increases in body iron. The iron chelator, deferoxamine, is available for clinical use, but requires prolonged parenteral infusion and is considered inconvenient and inefficient. Judgment is required for prescribing deferoxamine therapy and should take into consideration the anticipated number of red-cell units needed and the patient's age and prognosis.

Posttransfusion Purpura

Posttransfusion purpura (PTP) is a syndrome in which profound thrombocytopenia appears 5 to 14 days after transfusion. The majority of cases occur in women with a history of prior pregnancy and, almost all patients have a platelet-specific antibody, predominantly anti-HPA-1a and less frequently anti-HPA-1b, anti-HPA-3a, anti-HPA-3b, or anti-HPA-5b (38).

The thrombocytopenia is associated with alloantibodies against platelet-specific antigens not present on the patient's platelets. Hence, it is unclear why patients with such antibodies are thrombocytopenic. Potential etiologies include coexistence of autoantibodies or cross-reacting antibodies, the presence of immune complexes, or binding of transfused soluble platelet antigens to the patient's platelets.

Testing for platelet-specific antibodies includes use of platelet immunofluorescence tests, the monoclonal antibody-specific immobilization of platelet antigens assay, or modifications of these tests. Oligonucleotide hybridization testing is used to determine platelet genotypes in thrombocytopenic patients.

Infusion of intravenous immunoglobins is the treatment of choice. Therapeutic apheresis has been used successfully.

Immunosuppresive Effects of Blood Transfusion

Potential immunomodulatory effects associated with allogeneic blood transfusion include enhanced survival of renal allografts, fewer spontaneous abortions, increased recurrence rates following tumor resection, higher postoperative bacterial infection rates, fewer recurrences of Crohn's disease, and accelerated progression of HIV infection (39). Recent reports contradict earlier studies showing a benefit of alloantigen exposure for preventing spontaneous abortions and for alloantigen exposure accelerating HIV infections (40, 41).

The specific mechanism of transfusion-related immunomodulation is unknown, but is considered to be mediated, in part, by allogeneic white cells. Studies evaluating the role of leukocyte reduction are clouded by lack of consistency in dealing with confounding variables. Meta-analyses of randomized clinical trial data addressing tumor recurrences and postoperative infections do not support an adverse effect of allogeneic transfusions. However, based on the statistical power of the meta-analyses, an adverse effect of transfusion of less than 20 to 33% would not be discernible (39, 42).

One retrospective observational study of patients undergoing coronary artery bypass grafting found the risk of pneumonia increased by 5% for each unit of nonleukoreduced allogeneic red cells and/or platelets infused (43). The risk of pneumonia increased by 1% per day in relationship to the mean storage time of infused red cells. The authors postulated the retained white cells produced bioactive substances that contributed to this effect, and that leukocyte-reduction prior to storage would have prevented it. Alternatively, the authors questioned whether red cells with increased oxygen affinity, nitric oxide depletion, and increased rigidity during storage impeded blood flow in the lungs following infusion that predisposes patients to pulmonary ischemia and pneumonia.

In another study, patients undergoing coronary artery bypass grafting who received leukocyte reduced red cells had a lower incidence of postoperative infections, regardless of whether leukocyte-reduction was performed pre- or poststorage in comparison to recipients of non-leukocyte–reduced blood (44).

The inconsistencies result, in part, because of differences in study design and potential bias attributed to confounding variables. *In vitro* data, in contrast, provide data for determining mechanisms of action. For example, transfused donors cells circulate in immunocompetent surgery patients for up to 2 weeks after blood transfusion and persist for up to 18 months in multitransfused trauma patients (45). Transfusion initiates a microchimeric state in which some transfused cells circulate in the recipient for prolonged intervals (Fig. 74.3). It has been suggested that this leads to an *in vivo* mixed lymphocyte reaction in which T-cells become activated and immunomodulatory cytokines are produced with subsequent differentiation of inflammatory Th 1 cells or helper Th 2 cells. Preliminary data suggest a shift from a Th 1:Th 2 balance towards a Th 2 response. That is, graft rejection, tumor suppression, and bacterial cytotoxicity shift toward graft, infection, and cancer tolerance (46). Conclusions based on published data require careful consideration. Implementation of universal leukocyte reduction of all cellular blood components will alter clinical practice but should not preclude further investigation of the role played by leukocytes in transfusion-related immunomodulation.

TRANSFUSION-TRANSMITTED INFECTIONS

Part of "74 - Complications of Blood Transfusion"

The risk of acquiring hepatitis or a retroviral infection through blood transfusion decreased significantly during the past decade (Table 74.4). Improved tests for hepatitis C, HIV-1 and-2, HTLV I/II and HIV p24 antigen testing were implemented. Concern about emerging infections, such as HIV group O, and the theoretical risk of acquiring new variant Creutzfeldt Jacob Disease (nvCJD) led to deferral of blood donors born in, residing or traveling to endemic areas.

TABLE 74.4. TRANSFUSION TRANSMITTED INFECTIONS

	Window Period	Risk per Unit Transfused Prior to NAT
Hepatitis		
A		Infrequent
B	56–59 days (41–53 days)	1:63,000–1:233,000
C	70–82 days (35–47 days)	˜ 1:120,000
D		Infrequent
Retroviral Infection		
HIV-1/-2		˜ 1:676,000
anti-HIV	22 days	
p24 antigen	16 days (˜11 days)	
HTLV-I/-II		˜ 1:640,000
Cytomegalovirus		Various according to patient's status
Malaria		1:4,000,000
Chagas' disease		Infrequent
Syphilis		No recent cases
Parvovirus		Infrequent
Creutzfeldt Jakob disease (CJD) and new variant CJD		No cases reported

() = estimated window period post Nucleic Acid Amplification Testing-(NAT) (47, 48, 54)

In addition, testing for viral nucleic acids by nucleic acid amplification testing (NAT) began in 1999 under an FDA-approved Investigation New Drug procedure. Using this approach, NAT was performed on mini-pools containing aliquots from 24 or more individuals at approximately 20 locations in the United States. Logistic and cost restraints required pooled samples rather than single samples. NAT is intended to reduce the residual risk of hepatitis C, HIV, and hepatitis B by detecting donations made during the preseroconversion “window period,” the interval between infectivity and serodetection. Sufficient hepatitis C (10^3 to 10^4 copies per mL) and HIV (approximately 10^4 copies per mL) genomic copies are present in mini-pools, whereas hepatitis B copies are detected best in nonpooled samples (47, 48). Initial estimates indicated the HIV “window” would be reduced by 6 days and the window for HCV by approximately 35 days. Animal model experiments suggested the HIV window might be closed sufficiently to interdict transmission (49).

Hepatitis

Hepatitis A

A small nonenveloped RNA-containing virus belonging to the picornavirus family causes hepatitis A. A brief viremic phase occurs 5 to 28 days before onset of clinical disease. Transfusion-associated hepatitis A occurs infrequently because of the short viremic period before symptoms and the lack of an asymptomatic carrier state.

Transmission by donor blood components and by pooled clotting factor concentrates has been observed. Solvent/detergent treatment alone, as a viral inactivation process, is ineffective against hepatitis A.

The case-fatality rate is relatively low. However, fulminant liver failure occurred in some patients with chronic hepatitis C infection who later became infected with hepatitis A (50).

Hepatitis B

Hepatitis B, a partially double-stranded DNA virus, is a member of the Hepadnaviridae family. The infective virion or Dane particle, has a 42-mm diameter and contains surface and core, or nucleocapsid components (51).

Approximately one half of the 140,000 to 320,000 persons with acute hepatitis B annually in the United States develop symptoms and 2% to 10% become chronic carriers. Hepatitis B is transmitted through percutaneous or mucous-membrane exposures including heterosexual and male homosexual contact and intravenous drug use.

The period between infection and clinical symptoms ranges from 45 to 180 days. Hepatitis B surface antigen (HBsAg), hepatitis B DNA, and hepatitis B e antigen are present approximately 6 weeks after infection. ALT levels increase 2 weeks later, and symptoms and jaundice follow by 2 to 4 weeks. Antibodies against hepatitis B core (anti-HBc) appear at the onset of symptoms or ALT elevation and persist indefinitely. The infection resolves within 12 to 20 weeks of infection in all but approximately 5% who develop chronic hepatitis.

Transfusion-associated hepatitis B occurs at a frequency of 1 in 63,000 or 1 in 233,000 depending on the mathematical model selected for determining the risk (52, 53). The 1 in 63,000 estimate does not take into account the impact of anti-HBc testing, which is used routinely. Anti-HBc testing closes the serological windows in the setting of acute, resolving hepatitis B infection when HBsAg levels decrease before anti-HBs appears; infections with low levels of circulating HBsAg; and some hepatitis B variant infections. NAT may reduce the window between infection and hepatitis B infection by 6 to 15 days (54).

Hepatitis C

Hepatitis C, a single-stranded RNA flavivirus, causes the majority of infections formally termed non-A, non-B hepatitis. Approximately 3.9 million American have antibodies against hepatitis C, and 74% are positive for hepatitis C RNA, indicating that 2.7 million persons in the United States are infected chronically (55). Infection with genotype 1 accounts for 74% of those infected chronically (57% with genotype 1a and 17% with genotype 1b).

Risk factors for hepatitis C infection include injection drug use, cocaine use, smoking marijuana 100 times or more, early age at first intercourse, and 50 or more lifetime sexual partners. Employment in a health-related occupation, previously considered a risk factor for hepatitis C, did not have a similar association in a recent survey (55). Persons with evidence of hepatitis B infection are six times more likely to be positive for hepatitis C infection compared to those without hepatitis B infection. Estimates about perinatal transmission are low but vary considerably. In general, the efficiency of sexual transmission of hepatitis C is considered low, approximately 5% (56).

Although the incidence of transfusion-associated hepatitis C infections decreased dramatically following introduction of the initial and improved hepatitis C antibody tests in 1990 and 1992, transfusion-transmitted hepatitis C accounts for approximately 7% of the 3.9 million infected persons.

In addition to testing, efforts to reduce transmission of hepatitis C by blood transfusion incorporate a multi-layered approach. Potential donors are deferred from donation by providing predonation educational information and asking direct questions about risk behaviors for hepatitis C including intranasal cocaine use. Hepatitis C enzyme immunoassays (EIA) available in 1990 detected antibodies directed against epitopes expressed on the recombinant polypeptide, c100, derived from the NS-4 region (Fig. 74.4). Multiantigen or second-generation tests available in 1992 include recombinant polypeptides derived from the NS3 region (c33c), and the core region (c22-3), in addition to c100, contained in a composite antigen c200 comprised of c100 and c33c epitopes. Third-generation assays, providing minimal improvement over second-generation assays, use

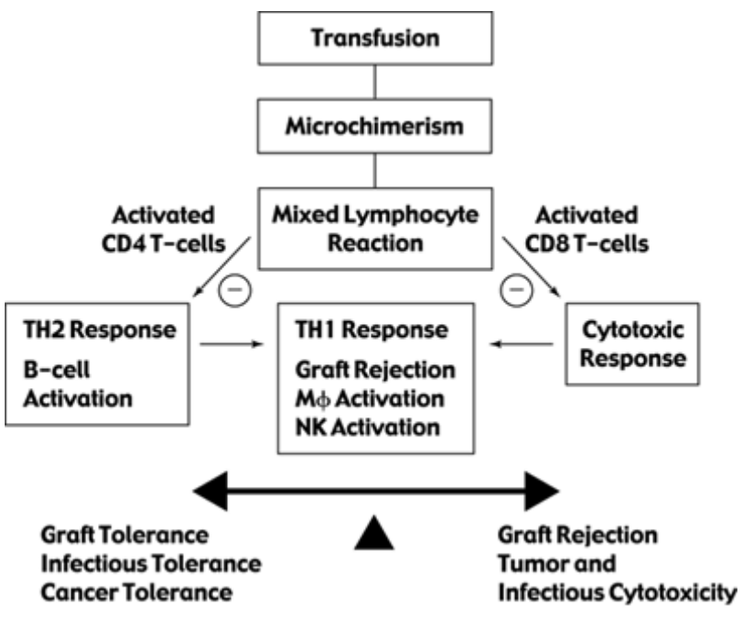


FIGURE 74.3. The prevailing pathophysiologic mechanism for transfusion related immunomodulation situates allogeneic leukocytes in a pivotal role. Transient microchimerism results in a shift towards a TH2 response. Presumably, this action is ameliorated by leukocyte-reduction (46). Mf = macrophage; NK = natural killer.

three recombinant antigens, c22-3, c200, and NS5, derived from four different regions of the HCV genome (57).

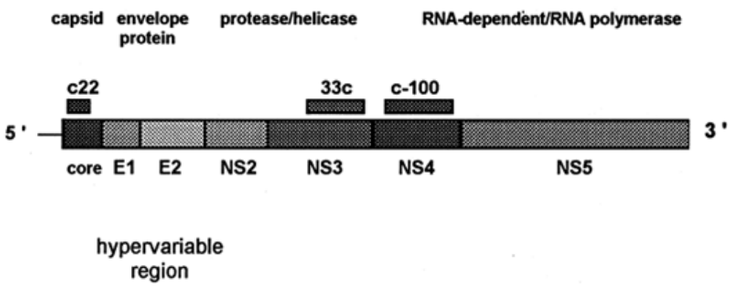


FIGURE 74.4. The hepatitis C virus genomic structure conta core, envelope (E), and various nonstructural (NS) regio c22, 33c, and c-100 refer to expressed viral epitopes. illustrated: c200 comprises c100 and c33c epitopes. (Modif from Centers for Disease Control and Prevention home pa http://www.cdc.gov/ncidod/diseases/hepatitis/hepatitis.h

Nonspecificity of EIA testing is resolved by using recombinant immunoblot assays (RIBA). Detection of hepatitis C RNA is conducted by nucleic amplification testing. Hepatitis C RNA is detectable in serum within 1 to 2 weeks of exposure and several weeks before ALT levels increase. At seroconversion, hepatitis C levels exceed 10^6 copies/mL and those infected chronically have 10^3 to 10^4 hepatitis C genome copies/mL, levels within the sensitivity of current NAT assays (47).

ALT testing, one of the two surrogate tests to prevent non-A, non-B hepatitis, is no longer mandated in the United States, although it is still performed by many blood centers. Anti-HBc has been retained, not to interdict potential donors infected with hepatitis C, but to reduce the risk of posttransfusion hepatitis B.

“Lookback” to identify recipients of blood from donors found hepatitis C antibody positive after 1990 was mandated in the United States by 1999. The “lookback” extends to recipients of blood a decade or more previously, records permitting, despite data suggesting as many as three-fourths of those patients are deceased at the current time and the conflicting data about long-term consequences of hepatitis C infection (58,59,60,61 and 62). Several studies of patients infected by blood transfusion or blood derivatives indicate severe complications such as cirrhosis and hepatocellular carcinoma do not appear for up to two decades after infection (60,61 and 62). Alcohol consumption appears to adversely affect hepatitis C infected individuals.

Hepatitis D

The delta agent, a “defective” RNA passenger virus, requires the hepatitis B virus as a “helper” for assembling envelope proteins. Delta hepatitis occurs only in patients with HBsAg. Efforts aimed at reducing the risk of posttransfusion hepatitis B infection, thereby, have a similar affect on hepatitis D transmission (63).

Hepatitis E

Hepatitis E is a single-stranded RNA calicivirus. Transmission occurs predominantly through the fecal-oral route, although viremia has been detected (64). To date, no cases of posttransfusion hepatitis E have been reported.

Hepatitis G (GBV-C)

Because the etiology of viral hepatitis remains elusive for 10% to 15% of posttransfusion hepatitis cases, the search remains open for detecting the etiologic agent(s) (65, 66). A novel, single-stranded RNA flavivirus has been designated both the GBV-C virus and the hepatitis G virus. Hepatitis G prevalence, determined by finding hepatitis G RNA in serum, approaches 1% to 2% in volunteer blood donors and is efficiently transmitted by blood transfusion. Antibodies directed against E2 serve as a marker of past infection since they occur infrequently in those RNA positive, but appear subsequently when RNA is no longer present. Viremia was found in 15% to 20% of hemophilic patients and intravenous drug users; 80% to 90% have anti-E2.

Although hepatitis G occurs relatively frequently in blood donors and is transmitted by transfusion, there does not appear to be an association with biochemical liver disease. Therefore, causality between the hepatitis G virus and clinical hepatitis has not been established and the term may be a premature misnomer (66).

TT Virus

In pursuit of agents causing “non A-E” hepatitis, a nonenveloped, single-stranded DNA virus in the Parvoviridae family, “TT,” was discovered by means of molecular testing methods (67). The name of the virus reflects the initials of the patient from whom the virus was isolated. While present in up to two thirds of multitransfused patients such as hemophiliacs and 12% of blood donors, it does not appear to be associated with elevated ALT levels in the blood donor population. Further investigation is required to establish whether the “TT” virus is related causally to hepatitis.

Retrovirus

HIV-1/-2

The initial reports of transfusion-associated AIDS appeared in 1982. Since that time, improved donor education about risk behaviors,

donor history questioning, and HIV testing have reduced the risk of transfusion transmitted HIV infections.

The “window” period between HIV infection and occurrence of HIV antibodies is approximately 22 days compared to 56 days when donor HIV testing started in 1985 (53). Testing for HIV-2 was introduced in 1992; 12 persons are known to be infected with HIV-2 in the United States (68). The introduction of HIV p24 antigen testing in 1996 narrowed the “window” to 16 days, and NAT should detect HIV within 11 days of infection (53).

Through June 30, 1999, the Centers of Disease Control and Prevention (CDC) reported 8,430 cases of AIDS in adult/adolescent blood transfusion or tissue recipients and 376 cases in children including 40 who received blood from HIV seronegative donors (68). There were an additional 5,243 hemophiliac and other patients with coagulation disorders who acquired HIV as a result of plasma-derivative therapy. Currently, coagulation factor concentrates are prepared by recombinant technology or undergo viral inactivation procedures such as heat (pasteurization), or physical-chemical disruption (solvent/detergent treatment).

HIV-1 group O viruses were isolated from central and West African patients in 1994. These HIV-1 variants share 65% to 70% homology with HIV-1 and 56% homology with HIV-2. Until test kits incorporating group O antigens are available, donors are asked if they were born in, lived in, traveled to, and received a blood transfusion, or had risk exposures in HIV-1 group O endemic African countries (69). Those providing affirmative answers are deferred from blood donation.

Human T-Lymphotropic viruses-I and -II

Human T-lymphotropic virus type I (HTLV-I) preferentially infects CD4+ lymphocytes and causes adult T-cell leukemia (ATL) or tropical spastic paraparesis/HTLV-associated myelopathy (TSP/HAM). ATL occurs in 2% to 4% of infected persons following a latent period that may extend to several decades. TSP/HAM occurs in less than 1% of HTLV-I infected persons. Those with transfusion associated TSP/HAM develop neurologic symptoms a median 3.3 years after transfusion.

HTLV-II infects CD8+ and CD4+ lymphocytes and macrophages and has been associated with TSP/HAM but not with hematologic malignancies. Persons infected with this virus appear to have an increased incidence of bronchitis, bladder and kidney infections, and oral herpes infections (70).

By 1998, donors were screened routinely with test kits designed specifically to detect both HTLV-I and -II. HTLV-I and -II are intracellular viruses, and are not transmitted by noncellular components such as plasma and cryoprecipitate.

Herpes Virus Infection

Cytomegalovirus

Cytomegalovirus (CMV) is a large, enveloped, double-stranded DNA herpes virus that resides intracellulary in leukocytes. Acute infections and latent recurrences occur (71). Whether secondary cases or reinfections occur is debated. Cellular components transmit CMV, not plasma or cryoprecipitate. CMV-seronegative persons receiving blood not screened for CMV have a 30% (range 10% to 70%) likelihood of seroconversion. CMV seronegative cellular blood components have a less than 1% to 4% risk of CMV. Cellular products subjected to leukocyte reduction by currently used leukocyte reduction filters (residual leukocyte count < 5×10^6) have an approximate 2% risk of transmitting CMV (72).

Approximately 50% of blood donors are CMV seropositive; CMV-reduced risk components (seronegative or leukocyte reduced by filtration) are reserved for selected patients (Table 74.5).

TABLE 74.5. INDICATIONS FOR CMV-RISK REDUCED COMPONENTS

Premature infants (≤1200 g) born to CMV-seronegative mothers
CMV-seronegative allogeneic and autologous progenitor cell transplant recipients
CMV-seronegative, HIV-positive patients
CMV-seronegative recipients of solid organ transplants from CMV-seronegative donors
CMV-seronegative patients who are potential progenitor cell transplant recipients
CMV-seronegative pregnant women

Epstein-Barr Virus Infection

Approximately 90% of the adult population has evidence of infection with Epstein-Barr virus (EBV) (73). This virus may establish latent infections in lymphocytes and cause proliferation of infected cells. Transfusion transmission, although reported, is unlikely because of the high incidence of prior infection in adults and deference of infection by host virus-specific cytotoxic T lymphocytes.

Parasitic Infections

Malaria

Approximately three transfusion-associated malaria cases occur per year in the United States. Malarial parasites, *Plasmodium falciparum*, *P. vivax*, *P. ovale*, and *P. malariae* maintain viability in red cells stored at 4°C, in platelet concentrates stored at room temperature, or following red cell cryopreservation and thawing. Malaria is not transmitted by noncellular products. The incubation period following transfusion ranges from 7 to 50 days (average 20 days) (74).

Babesiosis should be considered in the differential diagnosis. The diagnosis may be delayed if fever is attributed to the underlying illness or postoperative infection. Prevention of transfusion-transmitted malaria results from deferring donors emigrating or returning from malaria endemic regions.

Babesiosis

Babesiosis is caused by *Babesia microtia* or *B. equi* that is transmitted by a bite from an infected *Ixodes scapularius* or *I. pacificus* tick. At least 24 cases of transfusion-transmitted babesiosis have been reported in the United States. Persons with a history of

babesiosis are deferred from blood donation, as the infection may persist for months to years (75, 76).

Toxoplasmsis

Toxoplasmosis is caused by the intracellular protozoan parasite, *Toxoplasma gondii*, whose usual host is the domestic cat. Transfusion-associated cases have occurred in immunocompromised patients (71).

Leishmaniasis

No cases of transfusion-transmitted viscerotropic cases of leishmaniasis have been reported.

Chagas Disease

Trypanosoma cruzi, the etiologic agent of Chagas' disease, is endemic in Central and South America. *T. cruzi* is transmitted to humans primarily by a reduviid bug vector, but is transmitted also by blood transfusion and congenitally (77). Acute infections generally resolve in 1 to 2 months, but infected persons have a lifelong, asymptomatic, indeterminate phase in which low-level parasitemia persists. Approximately 10% to 30% of infected individuals develop chronic infections manifest by cardiac and central nervous system involvement.

Since 1987, at least four cases of transfusion-transmitted Chagas' disease have occurred in the United States and Canada. In three of these cases, platelet transfusions were provided by previous residents of endemic areas. This finding raised concern that other new immigrants to the United States could be asymptomatic carriers, and, could transmit the agent through transfusion. Seroprevalence studies indicate that up to 5.6% of donors at some centers may be at risk for *T. cruzi* infection.

Tests for screening blood donors for Chagas' disease are not available in the United States. Screening donors by history questioning to determine previous residence in endemic areas or living in dwellings where vector insects reside may not be effective in light of possible congenital transmission by decedents of women infected several generations previously.

Creutzfeldt-Jakob Disease (CJD) and New Variant CJD

CJD is a subtype of the spongiform encephalopathies caused, presumably, by a prion protein (PrP) or proteinaceous infections particle. The median age of onset is 61 to 67 years. The incidence of chronic CJD is approximately one case per million population, an estimate that has held constant for some time negating the likelihood of a new epidemic (78).

The disease occurs spontaneously, has a familial association, or follows an exposure. The acquired cases involve treatment with contaminated human pituitary-derived growth hormones, dura mater transplants, corneal transplant, and exposure to contaminated depth electroencephalograph electrodes. The latent period following human growth hormone administration ranges from 3 years to more than 20 years and for dura mater graft placement is 89 months (range 16 to 193 months) (79).

No cases of transfusion-transmitted CJD have been reported including recipients of blood from donors who later developed CJD. Regardless, the FDA requires deferral from blood donation by those at high risk for developing CJD including persons who received human pituitary-derived growth factors and dura-mater transplants, and persons whose families are at increased risk of CJD. If a donor develops CJD subsequent to a donation, in-date blood components are destroyed. Physicians caring for the recipients of prior donations are notified so they can decide whether to inform the patient that they received blood from someone who later developed CJD (80).

In 1996, a previously unrecognized variant of CJD was described in the United Kingdom. Patients presented predominantly with psychiatric and sensory symptoms. The median age of death was 29 years. The evidence linking bovine spongiform encephalopathy in the United Kingdom (mad cow disease) and new variant CJD (nvCJD) and experimental studies suggesting nvCJD and CJD maybe dissimilar led to concerns that nvCJD may be transmitted by blood transfusion. Approximately 47 patients died as a result of nvCJD in Europe through 1999. Although no cases of transfusion-transmitted nvCJD have been reported, the FDA issued guidance in June 1999 mandating deferral of persons who traveled or resided in the United Kingdom for 6 months or longer between 1980 and 1996. This affects an estimated 2.2% of the current U.S. blood supply or 285,000 whole-blood collections (80, 81).

Spirochete Infections

Syphilis

Serologic tests for syphilis were introduced for blood screening in the United States in 1938. Only two cases of transfusion-transmitted syphilis have been reported in the past 30 years (82).

Treponema pallidum, the spirochete causing syphilis loses viability at refrigerated temperature after seven days. This observation and the impact of concomitant antibiotic therapy in many patients receiving transfusions makes estimates about the efficiency of syphilis testing difficult to determine.

Lyme Borrelosis

Borrelia burgdorferi, the etiologic agent of Lyme disease, is transmitted to humans by *Ixodus scapularis* tick bites. Spirochetemia probably occurs postinfection, but reports of transfusion-transmitted Lyme disease have not appeared (83).

Other Tick-borne Illnesses

Rocky Mountain spotted fever and ehrlichiosis are transmitted by tick bites. Rocky Mountain spotted fever has been transmitted by transfusion. In a recent report involving blood donors with an extensive history of tick exposure, none of the recipients developed transfusion-transmitted illness (83).

Parvovirus Infection

Parvovirus B19, a nonenveloped DNA virus in the genus *Erythrovirus*, causes fifth disease and has been associated with transient

aplastic crisis, pure red cell aplasia, thrombocytopenia, pancytopenia, and arthropathy. Most infection occurs between 6 and 15 years of age. Up to 90% of adults are seropositive for parvovirus B19 antibodies. The acute infection is self-limiting. The virus is tropic for the globoside conferring the P blood group antigen. Some patients with red-cell hemolytic disorders such as sickle cell anemia and thalassemia develop acute aplastic or hypoplastic anemia following infection. Patients with HIV infection, solid organ transplants, or children with malignancies are at risk for aplasia. Fetal loss may follow acute parvovirus infection during the second trimester of pregnancy (84).

Neither solvent/detergent nor heat treatment inactivate parvovirus. In 1999, the FDA approved NAT for testing parvovirus B19 for the derivative, solvent/detergent treated plasma. Acute viremia occurs in less than 1 per 1,000 donors. Transfused patients adversely affected by parvovirus infection may respond to therapy with intravenous immunoglobulin.

CONCLUSION

Part of "74 - Complications of Blood Transfusion"

Despite the benefits of transfusion, inherent risks occur concomitantly. This chapter provides information about transfusion complications. Recognition of these events is needed to access the risk:benefit ratio associated with transfusion therapy.

REFERENCES

1. Lee JH. Centers for Biologics Evaluation and Research, FDA Workshop on Bacterial Contamination of Platelets, September 24, 1999.

2. Linden, JV. Errors in transfusion medicine. *Arch Path Lab Med* 1999;123:563–565.

3. Davenport RD, Kunkel SL. Cytokine roles in hemolytic and nonhemolytic transfusion reactions. *Trans Med Rev* 1994;8:157–168.

4. Capon SM, Goldfinger D. Acute hemolytic transfusion reaction, a paradigm of the systemic inflammatory response: new insights into pathology and treatment. *Transfusion* 1995;35:513–520.

5. Dzieczkowski JS, Barrett BB, Nester D, et al. Characterization of reactions after exclusive transfusion of white cell-reduced cellular blood components. *Transfusion* 1995;35:20–25.

6. Federowicz I, Barrett BB, Anderson JW, et al. Characterization of reactions after transfusion of cellular blood components that are white cell reduced before storage. *Transfusion* 1996;36:21–28.

7. Heddle NM, Klama L, Meyer R, et al. A randomized controlled trial comparing plasma removal with white cell reduction to prevent reactions to platelets. *Transfusion* 1999;39:231–238.

8. Kopko PM, Holland PV. Transfusion-Related Acute Lung Injury. *British J of Haem* 1999;105:322–329.

9. Dry SM, Bechard KM, Milford EL, et al. The pathology of Transfusion-Related Acute Lung Injury. *Am J Clin Pathol* 1999;112:216–221.

10. Silliman CC, Paterson AJ, Dickey WO, et al. The association of biologically active lipids with the development of transfusion-related acute lung injury: a retrospective study. *Transfusion* 1997;37:719–726.

11. Sandler SG, Mallory D, Malamut D, et al. IgA Anaphylactic Transfusion Reactions. *Trans Med Reviews* 1995;9:1–8.

12. Adverse ocular reactions following transfusions — United States, 1997–1998. *CDC MMWR* 1998;47.

13. Leitman SF, Boltansky H, Alter HJ, et al. Allergic reactions in healthy plateletpheresis donors caused by sensitization to ethylene oxide gas. *N Engl J Med* 1986;315:1192–1196.

14. Centers for Disease Control and Prevention. Hemolysis associated with 25% human albumin diluted with sterile water — United States, 1994–1998. *MMWR* 1999;48:157–59.

15. Linden JV, Kaplan HS, Murphy MT. Fatal air embolism due to perioperative blood recovery. *Anesth Analg* 1997;84:422–426.

16. Hiippala S. Replacement of massive blood loss. *Vox Sang* 1998;74:399–407.

17. Buchholz DH, Borgia JF, Ward M, et al. Comparison of Adsol and CPDA-1 blood preservatives during simulated massive resuscitation after hemorrhage in swine. *Transfusion* 1999;39:998–1004.

18. Klein HG, Dodd RY, Ness PM, et al. Current status of microbial contamination of blood components: summary of a conference. *Transfusion* 1997;37:95–101.

19. Wagner SJ, Friedman LI, Dodd RY. Transfusion-associated bacterial sepsis. *Clin Microbio Rev* 1994;7:290–302.

20. Centers for Disease Control and Prevention. Red blood cell transfusions contaminated with *Yersinia enterocolitica* — United States, 1991–1996, and initiation of a national study to detect bacteria-associated transfusion reactions. *MMWR* 1997;46:553–555.

21. Yomtovian R, Lazarus HM, Goodnough, et al. A perspective microbiologic surveillance program to detect and prevent the transfusion of bacterially contaminated platelets. *Transfusion* 1993;33:902–909.

22. Cyr M, Hume HA, Champagne M, et al. Anomaly of the des-Arg^9 — bradykinin metabolism associated with severe hypotensive reactions during blood transfusions: a preliminary study. *Transfusion* 1999;39:1084–1088.

23. Ness PM, Shirley RS, Thoman SK, et al. The differentiation of delayed serologic and delayed hemolytic transfusion reactions: incidence, long-term serologic findings, and clinical significance. *Transfusion* 1990;30:688–693.

24. Pineda AA, Vamvakas EC, Gorden LD, et al. Trends in the incidence of delayed hemolytic and delayed serologic transfusion reactions. *Transfusion* 1999;39:1097–1103.

25. von Zabern I, Ehlers M, Grunwald U, et al. Release of mediators of systemic inflammatory response syndrome in the course of a severe delayed hemolytic transfusion reaction caused by anti-D. *Transfusion* 1998;38:459–468.

26. Garratty G. Severe reactions associated with transfusion of patients with sickle cell disease. *Transfusion* 1997;37:357–361.

27. King KE, Shirey RS, Lankiewicz MW, et al. Delayed hemolytic transfusion reactions in sickle cell disease: simultaneous destruction of recipients' red cells. *Transfusion* 1997;37:376–381.

28. Petz LD, Calhoun L, Shulman IA, et al. The sickle cell hemolytic transfusion reaction syndrome. *Transfusion* 1997;37:382–392.

29. Redman M, Regan F, Contreras M. A prospective study of the incidence of red cell allo-immunisation following transfusion. *Vox Sang* 1996;71:216–220.

30. The Trial to Reduce Alloimmunization to Platelets Study Group. Leukocyte reduction and ultraviolet B irradiation of platelets to prevent alloimmunization and refractoriness to platelet transfusions. *N Engl J Med* 1997;337:1861–1869.

31. Linden J, Pisciotto PT. Transfusion-associated graft-versus-host disease and blood irradiation. *Trans Med Rev* 1992;6:116–123.

32. Vogelsang GB, Hess AD. Graft-versus-host disease: new directions for a persistent problem. *Blood* 1994;84:2061–2067.

33. Wagner FF, Flegel WA. Transfusion-associated graft-versus-host disease: risk due to homozygous HLA haplotypes. *Transfusion* 1995;35:284–291.

34. Ohto H, Anderson KC. Survey of transfusion-associated graft-versus-host disease in immunocompetent recipients. *Trans Med Rev* 1996;10:31–43.

35. Moroff G, Luban NLC. The irradiation of blood and blood components to prevent graft-versus-host disease: technical issues and guidelines. *Trans Med Rev* 1997;11:15–26.

36. Grass JA, Wafa T, Reames A, et al. Prevention of transfusion-associated Graft-versus-Host Disease by photochemical treatment. *Blood* 1999;93:3140–3147.

37. Bottomley SS. Secondary iron overload disorders. *Sem Hematol* 1998;35:77–86.

38. Taaning E, Tønnesen F. Pan-reactive platelet antibodies in post-transfusion purpura. *Vox Sang* 1999;76:120–23.

39. Vamvakas EC. Transfusion-associated cancer recurrence and postoperative infection: meta-analysis of randomized, controlled clinical trials. *Transfusion* 1996;36:175–86.

40. Ober C, Karrison T, Odem RR, et al. Mononuclear-cell immunisation in prevention of recurrent miscarriages: a randomised trial. *Lancet* 354;365–369.

41. Busch MP, Collier A, Gernsheimer T, et al. The viral activation transfusion study (VATS): rationale, objectives, and design overview. *Transfusion* 1996;36:854–859.

42. McAlister FA, Clark HD, Wells PS, et al. Perioperative allogeneic blood transfusion does not cause adverse sequelae in patients with cancer: a meta-analysis of unconfounded studies. *Brit J Surg* 1998;85:171–178.

43. Vamvakas EC, Carven JH. Transfusion and postoperative pneumonia in coronary artery bypass graft surgery: effect of the length of storage of transfused red cells. *Transfusion* 1999;39:701–710.

44. van de Watering LMG, Hermans J, Houbiers GA, et al. Beneficial effects of leukocyte depletion of transfused blood on postoperative complications in patients undergoing cardiac surgery. *Circulation* 1998;97:562–568.

45. Lee TH, Paglieroni T, Ohto H, et al. Survival of donor leukocyte subpopulations in immunocompetent transfusion recipients: frequent long-term microchimerism in severe trauma patients. *Blood* 1999;93:3127–3139.

46. Klein HG, Dzik S, Slichter SJ, et al. Leukocyte-reduced blood components: Current status. *Hematology* 1998;154–177.

47. Sun R, Schilling W, Jayakar H, et al. Simultaneous extraction of hepatitis C virus (HCV), hepatitis B virus, and HIV-1 from plasma and detection of HCV RNA by a reverse transcriptase-polymerase chain reaction assay designed for screening pooled units of donated blood. *Transfusion* 1999;39:1111–1119.

48. Stramer SL, Porter RA, Brodsky JP, et al. Replacement of HIV-1 p24 antigen screening with HIV-1 RNA nucleic acid testing (NAT) for whole blood donations. *Transfusion* 1999;39:2S.

49. Murthy KK, Henrard DR, Eichberg JW, et al. Redefining the HIV-infectious window period in the chimpanzee model: evidence to suggest that viral nucleic acid testing can prevent blood-borne transmission. *Transfusion* 1999;39:688–693.

50. Vento S, Garofano T, Renzini C, et al. Fulminant hepatitis associated with hepatitis A virus superinfection in patients with chronic hepatitis C. *N Engl J Med* 1998;338:286–290.

51. Mahoney FJ. Update on diagnosis, management, and prevention of hepatitis B virus infection. *Clin MicroBio Rev* 1999;12:351–366.

52. Schreiber GB, Busch MP, Kleinman SH, et al. The risk of transfusion-transmitted viral infections: the retrovirus epidemiology donor study. *N Engl J Med* 1996;334:1685–1690.

53. United States General Accounting Office. Report to the Ranking Minority Member, Committee on Commerce, House of Representatives. Blood supply, transfusion-associated risks. *GAO/PEMD*, 97-2, 1997.

54. Flanagan P, Barbara J. PCR testing of plasma pools: From concept to reality. *Trans Med Rev* 1999;13:164–176.

55. Alter Mj, Kruszon-Moran D, Nainan OV, et al. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N Engl J Med* 1999;341:556–562.

56. Recommendations for prevention and control of hepatitis C (HCV) infection and HCV-related chronic disease. *CDC MMWR* 1998;47:RR-19.

57. Vrielink H, Reesink HW, van den Burg PJM, et al. Performance of three generations of anti-hepatitis C virus enzyme-linked immunosorbent assays in donors and patients. *Transfusion* 1997;37:845–49.

58. Christensen PB, Groenbæk K, Krarup HB, et al. Lookback Group. Transfusion-acquired hepatitis C: the Danish lookback experience. *Transfusion* 1999;39:188–193.

59. Long A, Sprull G, Demers H, et al. Targeted hepatitis C lookback: Quebec, Canada. *Transfusion* 1999;39:194–200.

60. Seeff LB, Buskell-Bales Z, Wright EC, et al. Long-term mortality after transfusion-associated non-A, non-B hepatitis. The National Heart, Lung, and Blood Institute Study Group. *N Engl J Med* 1992;327:1906–1911.

61. Kenny-Walsh E, for the Irish Hepatology Research Group. Clinical outcomes after hepatitis C infection from contaminated anti-D immune globulin. *N Engl J Med* 1999;340:1228–1233.

62. Vogt M, Lang T, Frösner G. Prevalence and clinical outcome of hepatitis C infection in children who underwent cardiac surgery before the implementation of blood-donor screening. *N Engl J Med* 1999;341:866–870.

63. Dingle K, Bichko V, Zuccola H, et al. Initiation of hepatitis Delta virus genome replication. *J Virology* 1998;72:4783–4788.

64. Kwo PY, Schlauder GG, Carpenter HA, et al. Acute hepatitis E by a new isolate acquired in the United States. *Mayo Clin Proc* 1997;72:1133–1136.

65. Alter HJ, Gallagher M, Morris TT, et al. Acute non-A-E hepatitis in the United States and the role of hepatitis G virus infection. *N Engl J Med* 1997;336:741–746.

66. Alter HJ, Nakatsuju Y, Melpolder J, et al. The incidence of transfusion-associated hepatitis G virus infection and its relation to liver disease. *N Engl J Med* 1997;336:747–754.

67. Nishizawa T, Okomoto H, Konishi K, et al. A novel DNA viruses (TTV) associated with elevated transaminase levels in post transfusion hepatitis of unknown etiology. *Biochem Biophys Res Comm* 1997;241:92.

68. Centers for Disease Control. HIV/AIDS Surveillance Report 1999;11:1.

69. Centers for Disease Control and Prevention: Identification of HIV-1 Group O infection — Los Angeles County, California, 1996. *MMWR* 1996;45:561–565.

70. Murphy EL, Glynn SA, Friday J, et al. Increased incidence of infectious diseases during prospective follow-up of human T-Lymphotropic virus type II-and I-infected blood donors. *Arch Intern Med* 1999;159:1485–1491.

71. Hillyer CD, Lankford KV, Roback JD, et al. Transfusion of the HIV-seropositive patient: immunomodulation, viral reaction, and limiting exposure to EBV (HHV-4), CMV (HHV-5), and HHV-6, 7, and 8. *Trans Med Rev* 1999;13:1–17.

72. Bowden RA, Slichter SJ, Sayers M, et al. A comparison of filtered leukocyte-reduced and cytomegalovirus (CMV) seronegative blood products for the prevention of transfusion-associated CMV infection after marrow transplant. *Blood* 1995;86:3598–3603.

73. Alfieri C, Tanner J, Carpentier L, et al. Epstein-Barr virus transmission from a blood donor to an organ transplant recipient with recovery of the same virus strain from the recipient's blood and oropharynx. *Blood* 1996;87:812–817.

74. Centers for Disease Control and Prevention. Transfusion-transmitted malaria — Missouri and Pennsylvania, 1996–1998. *MMWR* 1999;48:253–256.

75. Dobroszycki J, Herwaldt BL, Boctor F, et al. A cluster of transfusion-associated babesiosis cases traced to a single asymptomatic donor. *JAMA* 1999;281:927–930.

76. Krause PJ, Spielman A, Telford SR. Persistent parasitemia after acute babesiosis. *N Engl J Med* 1998;339:160–165.

77. Leiby DA, Fucci MH, Stumpf RJ. Trypanosoma cruzi in a low- to moderate-risk blood donor population: seroprevalence and possible congenital transmission. *Transfusion* 1999;39:310–315.

78. Johnson RT, Gibbs CJ. Creutzfeldt-Jakob disease and related transmissible spongiform encephalopathies. *N Engl J Med* 1998;339: 1994–2004.

79. Centers for Disease Control and Prevention. Creutzfeldt-Jakob disease associated with cadaveric dura mater grafts — Japan, January 1979–May 1996. *MMWR* 1997;46:1066–1069.

80. Center for Biologics Evaluation and Research. Guidance for Industry: revised precautionary measure to reduce the possible risk of trans-mission of Creutzfeldt-Jakob disease (CJD) and new variant Creutzfeldt-Jakob disease (nvCJD) by blood and blood products. http://www.fda.gov/cber/guidelines.htm

81. Collinge J. Variant Creutzfeldt-Jacob disease. *Lancet* 1999;354: 317–323.

82. Aberle-Grasse J, Orton SL, Notari E, et al. Predictive value of past and current screening tests for syphilis in blood donors: changing from a rapid plasma reagin test to an automated specific treponemal test for screening. *Transfusion* 1999;39:206–211.

83. Arguin PM, Singleton J, Rotz RD, et al. An investigation into the possibility of transmission of tick-borne pathogens via blood transfusion. *Transfusion* 1999;39:828–833.

84. Azzi A, Morfini M, Mannucci PM. The transfusion-associated transmission of parvovirus B19. *Trans Med Rev* 1999;13:194–204.

75
Immune Hemolysis

Susan T. Johnson

Immune hemolysis is defined as increased destruction of red blood cells (RBCs) caused by antibody directed against antigens on RBCs. Antibody-mediated destruction of RBCs can be divided into four major categories: autoimmune hemolytic anemia, drug-induced immune hemolytic anemia, hemolytic transfusion reactions, and hemolytic disease of the newborn. Investigation as to the cause of immune hemolysis tells the history of transfusion medicine. Transfusion reactions and hemolytic disease of the newborn led to the discovery of most human blood groups known today. Studying autoimmune hemolytic anemias and drug-induced immune hemolytic anemia led to advances in serologic techniques.

INTRAVASCULAR AND EXTRAVASCULAR IMMUNE HEMOLYSIS

Part of "75 - Immune Hemolysis"

Antibody binding to antigen alone does not cause destruction of RBCs. This complex promotes other interactions to occur. Immune hemolysis is divided into two types depending on the mechanism by which RBCs are destroyed *in vivo*.

Intravascular Hemolysis

Intravascular hemolysis occurs after either IgG or IgM antibody binds to antigen on RBCs and the complement system is activated. Once complement is activated, destruction of red cells in the vasculature occurs and free hemoglobin and red-cell stroma are released into the plasma. This type of hemolysis may be more rapid and severe.

Extravascular Hemolysis

Most immune hemolysis occurs extravascularly. IgG antibody attaches to antigen and occasionally activates complement. Complement activation occurs to a lesser extent than in intravascular hemolysis. When activated, it generally is not complete activation, stopping with only C3b sensitizing the RBCs. IgG and C3b-coated RBCs are gradually removed from the circulation by the reticuloendothelial system. IgG antibody binds to corresponding antigen by its Fab portion, leaving the Fc portion of the antibody free. Fc receptors on macrophages in the spleen and liver capture these RBCs and phagocytize them or remove a portion of the red-cell membrane and release them back into the circulation as microspherocytes. C3b-coated cells are removed by C3b receptors on macrophages in the liver.

CLINICAL FEATURES OF IMMUNE HEMOLYSIS

Part of "75 - Immune Hemolysis"

Intravascular Hemolysis

Activation of the complement cascade resulting in C3 activation causes subsequent release of anaphylatoxins, C5a, and C3a. These proteins bind to receptors on mast cells and basophils causing them to release histamine and serotonin, vasoactive substances that can contribute to hypotension (See Fig. 75.1).

Antigen-antibody complexes also are capable of activating Factor XIIa (Hageman factor). Factor XIIa acts on the kinin system to produce bradykinin, which increases capillary permeability and dilates arterioles causing a drop in blood pressure. This drop in blood pressure stimulates the sympathetic nervous system. Catecholamines (norepinephrine and others) then are released, which causes vasoconstriction, especially in the kidneys.

If RBCs are lysed, circulating red-cell stroma, activated Hageman factor or thromboplastic substances can activate the intrinsic clotting cascade and may result in disseminated intravascular coagulopathy (DIC).

There is increasing evidence that cytokines also play a role in the pathophysiology of immune hemolysis (1). Cytokines are released from leukocytes that have been exposed to antigen-antibody complexes. They can induce fever, cause hypotension, mobilize neutrophils, activate endothelial cells to express adhesion molecules, and activate T and B lymphocytes.

Other clinical features associated with intravascular hemolysis and severe extravascular hemolysis are anemia caused by RBC destruction and subsequent hemoglobinemia. Plasma haptoglobin binds free hemoglobin, but it has a limited capacity. If free hemoglobin levels exceed this capacity, hemoglobin spills into the urine causing hemoglobinuria. Renal failure can result from acute hemolysis as a result of vasoconstriction, hypotension, and thrombi formation.

Extravascular Hemolysis

In extravascular hemolysis, the patient may exhibit symptoms related to a drop in hemoglobin, including fatigue/weakness, pallor,

palpitations, dyspnea, chest pain, and tachycardia. Symptoms related to RBC destruction may include jaundice, history of dark urine, hepatosplenomegaly, and fever.

LABORATORY FINDINGS

Part of "75 - Immune Hemolysis"

Common laboratory findings associated with immune hemolysis, whether intravascular or extravascular, are a drop in hemoglobin and hematocrit values and an increase in reticulocyte count. The peripheral blood smear contains microspherocytes, indicating ongoing red cell destruction, polychromasia, and occasionally nucleated and fragmented red cells. The bone marrow manifests erythroid hyperplasia. Chemistry findings include an increase in total bilirubin as a result of hemoglobin breakdown and a characteristic rise in unconjugated (indirect) bilirubin. Serum haptoglobin may be decreased or absent and lactose dehydrogenase (LDH) levels are increased.

The direct antiglobulin test (DAT) is key to defining immune hemolysis. This test is used to determine if IgG and/or complement is/are sensitizing the patient's red cells *in vivo*. In all but a few unusual instances discussed in this chapter, the DAT will be positive. Strength of reactivity and protein (IgG, C3, etc.) detectable on patient RBCs are unique to each type of immune hemolysis.

Another extremely useful test is the antibody detection test or antibody screen to determine if the patient has either alloantibody or autoantibody detectable in their plasma. In addition to detecting antibody, the reactivity (temperature at which it is detected) of the antibody is also important. Antibody in the plasma, a positive DAT and the patient's clinical history including age, race, sex, diagnosis, pregnancies, transfusions, and medications will present beneficial information when deciding on the type of immune hemolysis occurring in a patient. These characteristics will be highlighted in each section of this chapter.

AUTOIMMUNE HEMOLYTIC ANEMIA

Part of "75 - Immune Hemolysis"

Autoimmune hemolytic anemia (AIHA) occurs when a patient makes antibody (autoantibody) against antigen present on their own RBCs. It occurs in 1 per 40,000 to 80,000 individuals (3). AIHA is classified by the optimal reactivity of the autoantibody causing red cell destruction. Warm autoimmune hemolytic anemia (WAIHA) is caused by autoantibody reacting at 37C. Cold AIHA is caused by an IgM antibody whose optimal reactivity is at colder temperatures, down to 4C, but is capable of reacting at temperatures greater than 30C. Cold agglutinin syndrome (CAS) is the most common form of cold AIHA. It usually is caused by an autoantibody with I or i specificity. The rarest form

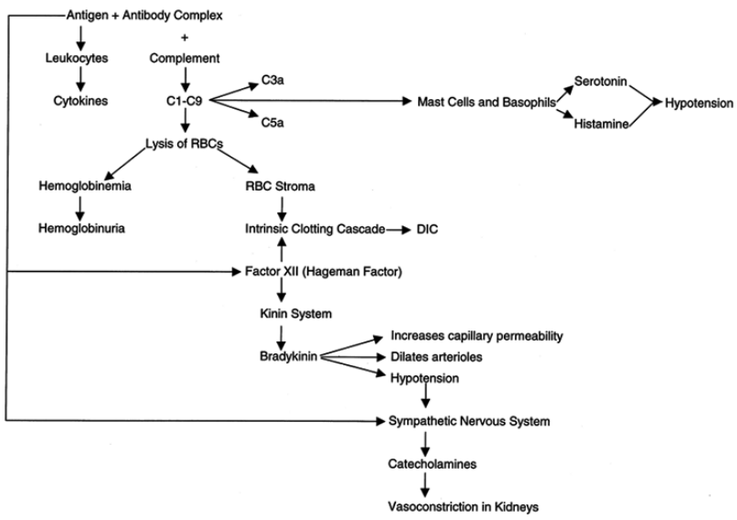


FIGURE 75.1. Pathophysiology of intravascular immune hemolysis. *Modified from Firestone DT (2).

is paroxysmal cold hemoglobinuria (PCH) caused by the biphasic hemolysin, autoanti-P. AIHA may occur rarely as a mixture of both warm and cold-reactive autoantibodies.

The type of AIHA occurring in a patient is established by reviewing the patient's clinical presentation and history, laboratory findings suggesting hemolysis, and results of serologic testing (See Table 75.1). Treatment, including transfusion management, differs for each type of AIHA.

TABLE 75.1. SEROLOGIC FINDINGS IN AUTOIMMUNE HEMOLYTIC ANEMIAS

	WAIHA	CAS	PCH	DIHA
Direct antiglobulin test	3–4+	2–4+	W–2+	W–4+
Anti-IgG	3–4+	0	0–W	W–4+
Anti-C3	0–3+	2–4+	W–2+	W–4+
Eluate	Positive w/all RBCs	Not indicated	Not indicated	Negative
Antibody detection test	0–4+	W–4+	W–2+	O–3+
Reactivity	IAT	Immediate spin, room temperature, (reactive at 37°C &/or 30°C)	Immediate spin, room temperature biphasic hemolysin	Variable, depending on drug and course of therapy
Antibody specificity	Broad	I, i	P	
	Rh			

Warm Autoimmune Hemolytic Anemia

WAIHA is the most common form of AIHA, occurring in approximately 70% of patients with AIHA (3). WAIHA can occur with no underlying or associated disorder. This is referred to as idiopathic or primary WAIHA. More commonly it is a secondary phenomenon associated with another disease. WAIHA may occur secondary to the following disorders: lymphoma, chronic lymphocytic leukemia, system lupus erythematosus, other autoimmune disorders, ovarian tumors, and chronic inflammatory disorders. WAIHA also may follow infection, especially in children.

Clinical Features

Many patients with WAIHA present with symptoms resulting from anemia such as those previously described. Hemoglobinemia and hemoglobinuria may be present in a patient experiencing rapid hemolysis. The extent of anemia is dependent on the degree of hemolysis and on the ability of the patient's bone marrow to respond to red-cell destruction.

Serologic Tests

Direct Antiglobulin Test

The DAT usually is strongly reactive (3-4+) with polyspecific antihuman globulin (AHG) reagent. Tests with AHG specific for IgG and C3 show that IgG alone is found in 67% of patients and IgG plus C3 is present in 24% of cases (4). One percent of cases will react weakly with polyspecific AHG and be negative with anti-IgG and anti-C3b,C3d reagents. Reactivity may be from contaminating anti-IgA and anti-IgM in a rabbit source AHG reagent. Complement alone is found in 7% of cases.

A negative DAT is seen in approximately 1% of patients even though they present with signs and symptoms of immune hemolysis. A well-performed DAT detects approximately 100 to 500 molecules of IgG per red cell. RBCs sensitized with a level of IgG molecules below the level of detection of the standard DAT is one reason for DAT-negative WAIHA. DAT-negative WAIHA also can be due to low-affinity IgG autoantibodies that elute off RBCs during routine washing of the RBCs prior to testing with AHG. Lastly, IgA or IgM autoantibodies are not detected by a DAT performed using monoclonal AHG. More sensitive DATs can be performed in specialized laboratories to try to confirm the diagnosis of autoimmune hemolytic anemia and to characterize the type of DAT-negative WAIHA a patient may be experiencing.

Serum

Approximately 50% of patients with WAIHA have detectable autoantibody in their serum when using standard, no additive, antibody detection tests (screening methods) (3). Autoantibody spills into the plasma *in vivo* when all antigen sites on the RBCs are occupied and there are no antigen sites left for autoantibody binding. Warm autoantibody usually is IgG and reacts optimally at 37C. Autoantibody usually is detected in the indirect antiglobulin test (IAT) and reacts with all reagent red cells tested including the patient's own cells.

It is rarely necessary to attempt to identify autoantibody specificity. Most autoantibodies are directed against antigens in the Rh blood group. Some have apparent, some relative, and some broad specificity within the Rh blood group. An autoantibody with apparent Rh specificity would be directed against normal Rh antigens such as D, C, or more commonly c or e. Antibody in the serum would appear to be anti-D, -C, -c, or -e but the patient's red cells would possess the corresponding antigen and the DAT would be positive.

An autoantibody with relative specificity would react stronger with red cells possessing the corresponding antigen than if the cells lacked the antigen. For example, a relative autoanti-e may be 3+ with e+ RBCs but 1+ with e- red cells.

Most autoantibodies have broad specificity. The only way to confirm that autoantibodies have Rh specificity would be to test them against rare Rh_{null} and/or D-- red cells. This testing is performed rarely because of limited availability of these cells.

Warm autoantibodies may have specificity to antigens in other blood groups. Some of the more common specificities seen include LW, U, Kell system, En^a and Wr^b. Kidd, Duffy, Xg, Vel, AnWj, Sc, R_x, ABO, and I specificities are reported but rarely seen.

One of the biggest challenges in evaluating serum of patients with WAIHA is determining if there are any underlying alloantibodies that would interfere should the patient require transfusion. If the patient's serum reacts with all panel cells, it is difficult to determine if underlying alloantibodies also are present. It is reported that 23% of patients with autoantibodies also have alloantibodies (5).

To determine if alloantibodies are present the autoantibody reactivity must be removed from the patient's serum. This is accomplished by performing adsorptions on the patient's serum using the patient's or other donors' RBCs. There are two types of adsorptions routinely performed. The first is an autologous adsorption performed when the patient has not been transfused recently. Patient's own RBCs are treated to remove autoantibody opening antigen sites and then patient serum is mixed with these treated cells. Autoantibody adsorbs onto the cells and any alloantibody remains in the adsorbed serum. More than one adsorption often is required to fully remove autoantibody reactivity.

When a patient recently has been transfused, allogeneic adsorptions are necessary because multiple populations of RBCs are present in the patient sample. If an autologous adsorption is performed and if the patient is making alloantibody to antigens present on transfused RBCs, the alloantibody could be adsorbed onto transfused RBCs. In an allogeneic adsorption, patient serum is placed on aliquots of RBCs of known phenotype, usually R_1R_1(c- E-), R_2R_2 (e- C-) and rr (D- C- E-), one of which is Jk(a-), or Jk(b-) and/or s- and incubated at 37C. Autoantibody will be adsorbed onto all aliquots regardless of phenotype because of their broad specificity. Alloantibody also will be removed if corresponding antigen is present on the adsorbing cells. Alloantibody will be left in the serum if the adsorbing cells lack the antigen. Selecting adsorbing cells of various phenotypes avoids missing an alloantibody. As with autologous adsorptions, more than one adsorption may be required to remove autoantibody reactivity. As can be imagined, these types of procedures are very time consuming. It normally requires a minimum 4 to 6 hours to determine if a patient has underlying alloantibodies. The extent of the workup, and the resultant time required for testing, needs to be considered when a request for blood is obtained for these patients.

Eluate

An elution procedure is performed to remove antibody from sensitized RBCs. It generally is done in an attempt to identify specificity of antibody coating patient cells. In most cases, it confirms antibody specificity found in serum testing. In patients with WAIHA, an elution may be performed the first time a patient is evaluated to confirm autoantibody and to rule-out the presence of drug-dependent antibody. An eluate typically will be positive with all cells tested showing broad specificity. A negative eluate will be indicative of drug-dependent antibody because the implicated drug must be present in the test for the antibody to react.

Transfusion Management

Many experts believe that transfusion should be avoided in patients with WAIHA. There are several risks associated with transfusing these patients. The first is that no blood selected for transfusion truly will be compatible with the patient's autoantibody. Compatible blood generally is not available for these patients because even if a specificity was determined, the type of blood needed is rare and would not be available for transfusion. Transfused blood will be as incompatible as the patient's own blood and the transfused blood will not enjoy normal survival. Transfusion also may stimulate alloantibody formation. Some fear that a patient's long-term prognosis is worsened by transfusion early in the disease (6).

There are patients in whom transfusion is necessary. If the patient is experiencing severe anemia and is at risk because of underlying heart disease or is experiencing neurological signs, transfusion is necessary. If the patient's clinical condition warrants transfusion before the time-consuming workup is completed, it is better to transfuse than to withhold blood.

Treatment of these patients generally includes corticosteroid immunosuppressive therapy. Transfusion can be avoided if the patient is put on bed rest and oxygen therapy until corticosteroid therapy is effective. Other forms of treatment include splenectomy, cyclosporine A, and intravenous gamma globulin.

COLD AUTOIMMUNE HEMOLYTIC ANEMIA

Part of "75 - Immune Hemolysis"

Cold Agglutinin Syndrome

Cold agglutinin syndrome (CAS), also know as cold hemagglutinin disease is caused by cold-reactive IgM autoantibodies that are capable of reacting at temperatures greater than 30C. Agglutination often occurs in extremities of the body. CAS accounts for approximately 30% of all cases of immune hemolysis (4).

It occurs as an acute or transient form and in a chronic form. The acute form generally is seen in adolescents or young adults following mycoplasma pneumonia or infectious mononucleosis. Chronic CAS often affects older adults occurring as part of another disease process. Lymphoproliferative disorders, including chronic lymphocytic leukemia, lymphomas, and Waldenström's macroglobulinemia are associated with CAS. Chronic CAS also can be idiopathic with no associated disease.

Clinical Features

Individuals often present with symptoms as a result of chronic anemia. Patients exposed to cold may experience acute hemolysis and subsequently show hemoglobinemia and hemoglobinuria. Kidney failure also has been seen in cases of acute hemolysis. Raynaud's phenomenon may be experienced by some patients.

Patients with acute, transient CAS associated with mycoplasma pneumonia or infectious mononucleosis may experience severe hemolysis resulting in a sudden drop in hemoglobin and hematocrit in addition to hemoglobinuria and hemoglobinemia.

Red blood cells in samples drawn from patients with CAS often will spontaneously clump or agglutinate as the sample temperature drops. This clumping makes it difficult to obtain accurate laboratory testing.

Serologic Tests

Direct Antiglobulin Test

Complement is the only protein detected on patient RBCs in most cases. Cold-reactive autoantibodies usually are IgM and bind to red cells at lower body temperatures in the circulation. As red cells move to warmer areas, IgM elutes off the cells and complement remains attached. C3d and sometimes C3b are detected by AHG reagent. Complement component C3 is activated and C3b attaches to the RBCs. Regulatory proteins convert most bound C3b to C3d.

Serum

IgM cold-reactive autoantibodies cause direct agglutination of RBCs. They react optimally at immediate spin and when incubated at colder temperatures. They can agglutinate red cells after the 37C incubation and occasionally positive reactions are seen in the IAT. IgM binding to reagent cells at immediate spin may cause complement activation. If polyspecific AHG is used, positive reactions in the IAT may be from anti-C3d binding to C3d coated red cells.

Most cold autoantibodies detected do not cause CAS. The difference between benign versus pathologic cold autoantibodies is thermal reactivity. Pathologic autoantibodies can attach to red cells *in vivo* at temperatures greater than 30C. The true thermal reactivity of an autoantibody must be determined on a sample that has been drawn and allowed to clot at 37C. The tube must be spun immediately and serum removed quickly. This prevents cold autoantibody from adsorbing onto the RBCs *in vitro*.

Most IgM cold-reactive autoantibodies have I or i specificity. Autoanti-I is seen most often in chronic CAS as well as following mycoplasma pneumonia. Autoanti-i is most often associated with infectious mononucleosis. Other specificities have been reported but are rare.

Eluate

An eluate is performed rarely because in almost all cases the antibody in question is IgM and elution methods work best for eluting IgG antibodies.

Transfusion Management

Transfusion rarely is necessary in CAS. Red-cell destruction often can be avoided if the patient stays in a very warm room. If autoantibody has a wider thermal range and acute hemolysis is occurring, the patient may require transfusion. As with WAIHA, it is essential prior to transfusion to determine if the patient has any underlying alloantibodies. In almost all instances, cold autoantibody reactivity can be avoided by performing a prewarm indirect antiglobulin test or by only reading for agglutination after incubating at 37C and performing the IAT. Once blood is available for transfusion, it may be transfused through a blood warmer. Others feel that this may not be necessary because blood generally is transfused over a 2 to 4 hour period during which time the blood warms to room temperature.

Paroxysmal Cold Hemoglobinuria

Paroxysmal cold hemoglobinuria (PCH) is the rarest form of AIHA. It most often presents as a sudden onset of hemolysis following viral infection in children. In the past, it often was associated with syphilis. PCH is caused by an autoantibody that is IgG in nature, which is capable of binding to RBCs at cold temperatures. When warmed, the antibody causes complement activation and subsequent hemolysis of patient RBCs.

Direct Antiglobulin Test

A routine DAT is positive because of C3 only. Even though PCH is caused by an IgG antibody, it does not remain bound to RBCs at 37C, normal body temperature. More sensitive DATs may detect IgG on red cells.

Serum

A routine antibody detection test may detect positive reactions at immediate spin, sometimes appearing to have anti-P_1 specificity at first glance. If reagent red cells are incubated at colder temperatures positive reactions will be detected with all cells. If the patient's serum is tested against rare P– red cells, negative reactions may be seen since many of these antibodies have P specificity.

The diagnostic test for this disease is the Donath-Landsteiner Test. The patient's serum is mixed with normal red cells and incubated at 4C followed by a 37C incubation. Hemolysis occurs if the Donath-Landsteiner antibody is present.

Eluate

An eluate usually is not indicated because only complement is detected on patient red cells. If an eluate is performed, it generally is negative.

Transfusion Management

Transfusion rarely is necessary in PCH because it usually is self-limiting and recovery occurs within several weeks. If transfusion is necessary, it usually occurs in young children and may be needed as a life-saving measure. Some believe that P-negative red cells survive better than P-positive blood in patients with autoanti-P (6). The incidence of P-negative blood is approximately 1 in 200,000 people. Therefore, if transfusion is urgently needed there often is not time to provide P-negative blood. Moreover, survival of transfused P-positive blood will be about the same as the patient's own blood.

DRUG-INDUCED IMMUNE HEMOLYTIC ANEMIA

Part of "75 - Immune Hemolysis"

Drug-induced immune hemolytic anemia is unusual but when it occurs it may present in a dramatic fashion. The patient often experiences a sudden fall in hemoglobin following ingestion or administration of the implicated drug. These drugs may be varied, being prescription or over the counter, oral or intravenous. Anemia is caused by drug-inducing production of antibody to the drug itself or against new red-cell antigens somehow presented by the drug. The exact mechanism by which this occurs is unknown.

Theories of Immune Response to Drugs

Drug-induced immune hemolysis can be caused by at least three mechanisms, designated Type I, Type II, and Type III (Table 75.2) (8). Mechanisms of causation have been proposed through the years, including drug adsorption, immune-complex formation, nonspecific protein adsorption, and autoantibody production. The latter theories were useful, but there was insufficient data, and current evidence suggests that these proposed mechanisms were inaccurate.

TABLE 75.2. MECHANISMS OF DRUG-INDUCED IMMUNE HEMOLYSIS*

Type	Characteristics	Detection of Drug-Dependent Antibody	Prototype Drugs
I (Hapten-dependent antibody)	Drug binds covalently to membrane proteins and stimulates hapten-dependent antibodies	Antibody reacts with normal RBCs pretreated with drug and then washed.	Penicillin/penicillin derivatives
II (Autoantibody)	Through an unknown mechanism, drug induces autoantibodies specific for red blood cell (RBC) membrane proteins.	Antibody reacts with normal RBCs in the absence of drug.	Alpha methyldopa, procainamide
III (Drug-dependent antibody)	Through an unknown mechanism, drug induces antibodies that bind to RBC only when drug is present in soluble form.	Antibody reacts with normal RBCs when soluble drug is present.	Quinidine, quinine, NSAIDs

NSAIDs, nonsteroidal anti-inflammatory drugs

*Reprinted with permission from *Transfusion 2000*; 40(6):663–668.

Type I drug-related immune hemolysis is likely caused by hapten-dependent antibody. Certain drugs, such as penicillin, penicillin derivatives, and cephalosporins covalently link *in vivo* to red-cell membrane glycoproteins and act as a hapten, triggering antibody production. The patient usually is receiving high doses of the drug (usually intravenous), which leads to the patient's own RBCs becoming coated with drug. Hemolysis usually occurs extravascularly in Type I. The IgG-coated red cells are removed from the circulation by the RE system.

Massive doses of intravenous penicillin causes approximately 3% of patients to develop a positive DAT, but only a small percentage of these go on to develop hemolytic anemia. A more “common” group of drugs, based on current usage, causing hemolytic anemia that likely operates by this mechanism are cephalosporins. The most common of this group are the newer second- and third-generation cephalosporins.

A second type of immune hemolysis (Type II) is associated with sensitivity to drugs such as alpha methyldopa and procainamide. Autoantibody is produced that does not require the drug to be present in order for the antibody to bind to RBCs *in vitro*. The patient serologically presents like a patient with WAIHA.

Alpha methyldopa was a common cause of this phenomenon, but that drug has been replaced by new therapies. In patients continuing to take alpha methyldopa, approximately 15% develop a positive DAT, but only 0.5% to 1.0% develop hemolytic anemia (9). The DAT becomes positive within 3 to 6 months of therapy and the incidence of the positive DAT is dose-dependent. Following cessation of the drug, hemolysis subsides within 2 weeks, but the DAT may continue to be positive for up to 2 years. Hemolysis in Type II DIHA usually is extravascular and is associated with mild to severe anemia.

The third type of drug-induced hemolysis (Type III) occurs when, through an unknown mechanism, drug induces production of antibodies that bind to RBCs only when drug is present in soluble form. Drugs in this group do not bind to RBCs with high affinity. Type III drug-dependent antibodies may recognize compound epitopes (part drug, part protein), or the antibodies may be specific for secondary conformational changes induced by drug on the red cells. Quinine and quinidine are the prototype drugs of Type III antibodies. However, the nonsteroidal antiinflammatory drugs (NSAIDs) currently appear to be a more common cause of immune hemolysis.

Type III antibodies cause acute, intravascular hemolysis with hemoglobinuria, hemoglobinemia, and other signs and symptoms previously described associated with severe hemolysis. Once antibody has been produced by a patient, severe hemolytic episodes may recur after reexposure to even very small quantities of the drug (one tablet in some cases). Numerous fatalities have been reported with this type of hemolysis (9, 10).

In most cases, only one mechanism causes hemolysis. However, some patients appear to have a combination of Types I and III (9, 12) or Types II and III (13). Table 75.3 lists drugs associated with immune heolysis and their reported mechanism for causing heolysis.

Serologic Tests

Type I

The DAT is strongly positive due to IgG coating of the RBCs. Complement sensitization also is occasionally seen but usually is weaker in reactivity.

Routine serum testing is negative. Drug-dependent antibody is demonstrated in the serum by testing against drug-coated RBCs because antibody is directed against the drug (hapten) itself. Caution must be taken when evaluating drug testing as many individuals have naturally occurring drug-dependent antibodies from prior exposure to the medication. Drug-dependent antibody-causing immune hemolysis is high titered, while naturally occurring antibody in an individual with no hemolysis is low titered.

Patients with Type I and III drug-induced hemolytic anemia characteristically have a negative eluate. Test RBCs must be coated with drug in order for drug-dependent antibody in the eluate to bind and cause agglutination.

Type II

Type II drug-induced hemolytic anemia appears serologically identical to WAIHA. The DAT is strongly positive, usually only because of IgG binding. Serum testing shows an IgG, warm autoantibody reacting at 37C and detectable in the IAT. An eluate is positive with all reagent RBCs tested. A separate drug-dependent antibody is not demonstrable.

Type III

The DAT usually is strongly positive because of IgG and complement sensitization of patient RBCs. There are reports of the DAT being positive only because of complement coating the patient's RBCs or being weak from IgG and complement coating.

Standard antibody detection tests will be negative. Drug must be present in the test system in order for the drug-dependent antibody to bind to reagent red cells. When soluble drug in a 1-mg/mL concentration is added to the serum and reagent RBC mixture, positive reactions are obtained. Antibody detected can be either IgG or IgM. Occasionally, the patient may have an antibody that is dependent on a metabolite of the drug. Testing for drug-dependent antibody then requires obtaining a source of drug metabolites. This source is urine and/or plasma from an individual taking the offending drug. Urine and/or plasma, containing metabolites of the drug, is mixed with patient serum and reagent RBCs. The presence of a drug metabolite-dependent antibody results in a positive reaction.

Positive reactions may be obtained in a routine antibody screen if a patient has recently taken the implicated drug and it continues to be present in the patient's circulation.

An eluate will be negative when tested using routine methods. When drug or drug metabolite is mixed with eluate and reagent RBCs a positive reaction will often be seen.

ALLOIMMUNE HEMOLYTIC ANEMIA

Part of "75 - Immune Hemolysis"

Alloimmune hemolysis results from foreign RBCs being introduced into a patient with corresponding alloantibodies. There

TABLE 75.3. DRUGS ASSOCIATED WITH IMMUNE HEMOLYSIS

Drug	Possible Mechanism	Drug	Possible Mechanism
Acetaminophen	III	Interferon	II
Aminopyrine	III	Isoniazid	I, III
Amphotericin B	III	Levodopa	II
Ampicillin	III	Mefenamic acid	II
Antazoline	III	Mefloquine	III
Apazone (azapropazone)	II, I	Melphalan	III
Buthiazide (butizide)	III	6-Mercaptopurine	I
Carbenicillin	I	Methicillin	I
Carbimazole	III	Methotrexate	III
Carboplatin	I, III	Methyldopa	II
Carbromal	I	Moxalactam (latamoxef)	III, II
Catergen	II	Nafcillin	I
Cephalosporins		Nomifensine	II, III
First generation	I	p-Aminosalicylic acid	III
Second generation	III, I, II	Penicillin G	I
Third generation	III, I, II	Phenacetin	II, III
Chaparral	II	Piperacillin	I, III
Chorpropamide	III	Probenecid	III
Chlorpromazine	II, III	Procainamide	II
Cladribine (chlorodeoxyadenosine)	II	Propyphenazone	III
		Pyramidon	III
Cyanidanol	II, I, III	Quinidine	I, III
Cyclofenil	II	Quinine	III
Cyclosporine	II	Rifampin (rifampicin)	III
Diclofenac	II, III	Sodium Pentothal	III
Diethylstilbestrol	III	Stibophen	III
Dipyrone	III, I	Streptomycin	II, I, III
Erythromycin	I	Sulfonamides	III
Etodolac	III	Sulfonylurea derivatives	III
Fenoprofen	II, III	Sulindac	I, II
Fludarabine	II	Suprofen	III, II
Fluorescein	I, III	Temafloxacin	III
Fluorouracil	III	Teniposide	II, III
Glafenine	II, III	Tetracycline	I, III
Hydralazine	III	Thiopental	III
Hydrochlorothiazide	III	Tolbutamide	I
Elliptinium acetate	III	Tolmetin	II, III
Ibuprofen	III	Triamterene	III
Insulin	I, III	Zomepirac	I, III, II

are two forms of alloimmune hemolytic anemia: hemolytic transfusion reactions (HTR) and hemolytic disease of the newborn (HDN). In HTR, the patient with alloantibody is transfused with blood from an allogeneic donor and in HDN, the mother's antibody is exposed to baby's RBCs.

HEMOLYTIC TRANSFUSION REACTIONS

Part of "75 - Immune Hemolysis"

HTR can be classified into two types, immediate and delayed, based on the clinical presentation.

Immediate Transfusion Reactions

Immediate or acute transfusion reactions occur when a patient has preformed antibodies and is transfused with RBCs that possess antigens corresponding to the antibody. Transfusion of ABO-incompatible RBCs accounts for most severe cases of acute, intravascular hemolysis (14).

Symptoms of an acute HTR include any or all of the following:

- Fever with or without chills, defined as a 1C increase in body temperature
- Shaking chills
- Pain at infusion site or in chest, abdomen, or flanks
- Blood pressure changes, either hypertension or hypotension
- Respiratory distress, dyspnea, tachypnea or hypoxemia
- Skin changes: flushing, itching, urticaria, or localized edema
- Nausea
- Circulatory shock
- Hemoglobinemia followed by hemoglobinuria

It is more difficult to identify a HTR in an anesthetized patient. Diffuse bleeding at the surgical site, hypotension, or hemoglobinuria may be the only recognized symptoms.

Evaluation of a Suspected Hemolytic Transfusion Reaction

Personnel transfusing blood usually are the first to suspect a transfusion reaction. When a reaction is suspected, the transfusion must be stopped; however, the intravenous line should be maintained with normal saline or other suitable fluid. Individuals administering blood should perform a clerical check of all labels, forms, and patient identification and the patient's physician should be notified immediately. The patient's physician evaluates the patient and determines if there is probable cause of an immune mediated HTR. A blood sample must be drawn, taking care to avoid mechanical hemolysis and the transfusion service must be notified.

When a transfusion reaction workup is received in the lab, there are three steps that must be taken.

1. Clerical check: A clerical check or review of pretransfusion sample identification, transfused blood product labels, and review of pretransfusion testing is performed, both in the laboratory and by personnel on the patient unit. If an error is found, all transfusions must be stopped until the cause for the error can be determined. When this scenario occurs, it is not uncommon to have two patient samples mixed, thereby exposing two patients to adverse reactions.
2. Visual check for hemolysis: Postreaction serum or plasma must be visually inspected for hemolysis and compared to a prereaction specimen when available. Visible hemoglobinemia occurs when incompatible red cells are lysed and hemoglobin is released into the plasma.

   A postreaction urine also may be examined for evidence of hemoglobinuria. It is important to distinguish hemoglobinuria (free hemoglobin) from hematuria (intact red cells free in the urine).

3. DAT: A DAT must be performed on a postreaction specimen to determine if patient antibody was able to coat transfused red cells. If the postreaction DAT is positive, a DAT should be performed on a pretransfusion sample to validate that it became positive because of transfusion. The DAT will be positive when antibody in the patient coats transfused RBCs and these cells have not been cleared from the circulation. It is usually weakly positive because there are relatively few transfused cells in the individual's total blood volume. Characteristic mixed field reaction as a result of two populations of cells being present often is noted. If all transfused cells are destroyed, or if the sample is drawn, several hours posttransfusion, the DAT may be negative.

Confirmatory Testing

If positive indication of hemolysis is identified in any of the above three checks, additional testing must be performed to confirm the cause of an apparent transfusion reaction.

Eluate

An eluate must be performed on posttransfusion RBCs to determine specificity of the antibody. If the prereaction sample also had a positive DAT, an eluate also should be made from this sample so that a comparison can be made.

ABO, Rh Testing

ABO, Rh typing should be repeated on pre- and postreaction samples. In addition, the ABO and Rh type should be repeated on any units of blood crossmatched for the patient. If there is a discrepancy between the pre- and postsample there has been a mix-up in patient or sample identification or in testing. If there is a discrepancy between patient type and unit type there may be a problem in unit labeling or in selection of blood chosen for the patient.

Antibody Detection Tests

Antibody detection tests must be repeated on pre- and postreaction samples. If a positive reaction is obtained in the pretransfusion sample that was not previously detected, the antibody must be identified to determine if it could have caused the transfusion reaction. If a positive reaction is detected in the post reaction sample, it could be because of an amnestic response following exposure to foreign red cells or it could be because of passive antibody from the transfused unit of blood.

Crossmatch

Serologic crossmatch tests should be repeated using the pre- and postreaction samples. This testing should include an IAT even if it was not performed on the initial crossmatches. Occasionally, an antibody to a low-frequency antigen may be detected in an IAT crossmatch that would not be detected in a standard antibody detection test.

Laboratory Testing

Posttransfusion hemoglobin and hematocrits should be performed to determine if the patient responded appropriately to the transfusion or if there is rapid or slow decline after transfusion. Haptoglobin level testing may assist in determining if hemolysis occurred. A posttransfusion bilirubin approximately 5 to 7 hours after transfusion also may indicate if hemolysis occurred.

Examine Unit Transfused

The unit of blood implicated in the hemologic reaction should be returned to the transfusion service. It should be examined for evidence of hemolysis in the administration tubing or bag itself. If hemolysis is present it could be because of hypotonic solution being added to the bag or a faulty infusion device being used.

The unit also should be inspected for clots and brownish, opaque, muddy, or purple discoloration. If any of these are present bacterial contamination of the unit should be suspected. A Gram's stain and aerobic and anaerobic bacterial culture of the contents of the bag should be performed. Even if the unit looks normal, but the patient experienced a reaction that is unexplained serologically, this testing should be performed.

HLA/Neutrophil Testing

Rarely, the patient may have had signs and symptoms of transfusion-related acute lung injury (TRALI). If this is the case, the patient's pre- and postreaction sample, as well as plasma from the transfused unit should be tested for antibodies to HLA and neutrophil antigens.

Delayed Transfusion Reaction

Antibodies previously formed by an individual with time can fall below the level of detection of antibody screening methods. If red blood cells are transfused that have the corresponding antigen an anamnestic response occurs within hours to days of transfusion. IgG antibodies are produced that are able to bind to transfused antigen positive red cells. Kidd system antibodies (anti-Jk^a and anti-Jk^b) are notorious for causing delayed transfusion reactions. These types of reactions classically occur 7 to 14 days after transfusion.

Many times, the anamnestic antibody production does not cause hemolysis of tranfused red cells. This has been termed a delayed serologic reaction. If hemolysis occurs, the most common signs and symptoms are fever, dropping hemoglobin and hematocrit and jaundice. These patients often present as having an inadequate rise in hemoglobin following transfusion, and additional units of blood are ordered to be transfused. Subsequent pretransfusion testing detects an antibody that was not present in the sample previously tested, and serologic testing will detect a positive DAT. Weaker reactivity (weak-2+) normally is seen because the antibody is binding to transfused red cells, the smaller portion of the patient's total blood volume. An eluate prepared from the patient's red cells usually will show the same antibody specificity as that now found in the serum.

If a new antibody is detected in a posttransfusion sample, it is important that the patient's physician be made aware of this finding so that a possible delayed hemolytic transfusion reaction can be investigated.

HEMOLYTIC DISEASE OF THE NEWBORN

Part of "75 - Immune Hemolysis"

Hemolytic disease of the newborn occurs when IgG alloantibody from the mother crosses the placenta *in utero* and coats fetal red cells. Alloantibody present in mother can be stimulated by previous pregnancies, current pregnancy, or previous transfusions. For antibody to bind to fetal red cells, the corresponding antigen must be fairly well developed. Clinical severity can range from intrauterine death to an infant born with only a positive DAT and no clinical symptoms of red cell destruction.

Antibodies Causing HDN

HDN can be classified into three categories based on the specificity of the antibody involved.

1. Rh Antibodies

   Anti-D is the most common antibody causing HDN and causes the most severe red cell destruction. If left untreated intrauterine death can occur. Prior to the use of Rh immune globulin (RHIG) given to D- mothers bearing D+ infants, anti-D was the most common antibody involved. Most antibodies in the Rh system have been reported to cause HDN including but not limited to anti-C, -E, -c, -e and -G. In many cases the severity of HDN is less than that found with anti-D. If any of these antibodies are detected in a pregnant woman the antibody should be monitored throughout pregnancy.

2. ABO Antibodies

   Anti-A, anti-B and anti-A,B are often a mixture of IgM and IgG.

   Group O individuals are more likely to produce higher levels of IgG ABO antibodies. Therefore, it is not uncommon for group A or B babies born from group O mothers to have a positive DAT. The great majority of cases are subclinical, although some babies may be jaundice and slightly anemic, and there are rare reports of death of the fetus caused by ABO antibodies (15).

3. Other Antibodies

Antibodies to antigens in almost every blood group (Table 75.4) have been described to cause subclinical, mild to severe HDN. The most common is anti-K. In addition, anti-K seems not only to cause red-cell destruction, it also appears to suppress fetal erythropoiesis (16). Incidence of other antibodies is much lower. Many other uncommon or rare antibodies to high and low frequency antigens have been reported to cause red-cell destruction

in infants. Many blood group antigens were first described after causing HDN.

TABLE 75.4. COMMON ANTIBODIES CAUSING IMMUNE HEMOLYSIS*

Antibody	AIHA	HTR	HDN
Anti-A	Rare	No to severe	No to moderate
Anti-B	Rare	No to severe	No to moderate
Anti-A_1	Rare	No to mild/delayed	No
Anti-D	Yes, warm	Mild to severe/immediate or delayed	Mild to severe
Anti-C	Yes, warm	Mild to severe/immediate or delayed	Mild
Anti-E	Yes, warm	Mild to moderate/immediate or delayed	Mild
Anti-c	Yes, warm	Mild to severe/immediate or delayed	Mild to severe
Anti-e	Common warm	Mild to moderate or delayed	Rare, usually mild
Anti-M	Rare, cold	No	Extremely rare
Anti-N	Rare, cold	No	No
Anti-S	Rare	No to moderate (rare)	No to severe (rare)
Anti-s		No to mild (rare)	No to severe (rare)
Anti-U	Yes, warm	Mild to severe	Mild to severe
Anti-Lu^a	No	No	No to mild (rare)
Anti-Lu^b	No	Mild to moderate	Mild
Anti-K	No	Mild to severe/delayed/hemolytic	Mild to severe (rare)
Anti-k	No	Mild to moderate/delayed	Mild to severe (rare)
Anti-Kp^a	No	Mild to moderate/delayed	Mild to moderate
Anti-Kp^b	Yes	No to moderate/delayed	Mild to moderate
Anti-Js^a	No	No to moderate/delayed	Mild to moderate
Anti-Js^b	No	Mild to moderate/delayed	Mild to moderate
Anti-Le^a	No	Hemolytic (rare)	No
Anti-Le^b	No	No	No
Anti-P_1	No	No to moderate/delayed (rare)	No
Anti-Fy^a	Rare, warm	Mild to severe/immediate/delayed	Mild to severe (rare)
Anti-Fy^b	Rare, warm	Mild to severe/immediate/rare; delayed	Mild (rare)
Anti-Jk^a	Yes, warm	No to severe/immediate or delayed/hemolytic	Mild to moderate (rare)
Anti-Jk^b	Yes, warm	No to severe/immediate or delayed/hemolytic	No to mild/rare

*Modified from Reid and Lomas-Francis[27]

There are three commonly occurring antibodies that do not cause HDN: anti-Le^a, anti-Le^b, and anti-P_1. It is much easier to remember these three and beware of all others. There are two reasons that these antibodies do not cause problems. Most are IgM and cannot cross the placenta and secondly these antigens are not well developed on fetal cells.

RECOGNITION OF “AT RISK” PREGNANCIES

Part of "75 - Immune Hemolysis"

The majority of above-described antibodies can be detected in routine antibody detection tests. Serologic studies, including an ABO, Rh type, and antibody detection test, should be performed on all pregnant women as early in the pregnancy as possible (17). If an antibody is detected it must be identified so that its clinical significance can be assessed. The likelihood that the fetus will be at risk often can be predicted by serologically typing the father's red cells to determine if he carries the corresponding antigen. If he possesses the antigen, the fetal red cells also could have the antigen. Fetal genotyping also can be performed by isolating DNA from amniocytes in amniotic fluid, chorionic villus sampling, or fetal blood obtained by cordocentesis. DNA is amplified by polymerase chain reaction (PCR) and typed for the presence of the suspected gene. Genotyping is available for *ABO, Rh (D, C, E, c, e), Jka/Jkb, K1/K2, Fya/Fyb, M* and *S/s*.

Monitoring for HDN

Antibody Titration

If it is determined that the fetus is at risk, or risk is unclear, antibody titrations can aid in monitoring antibody level. This is a semiquantitative method. A baseline titer should be determined in the first trimester and the sample frozen for future comparison studies. Invasive techniques will not be used until 16 to 18 weeks' gestation so repeat titrations are not recommended until that time. Repeat titrations should be performed using the same method and test cells with the same antigen expression as that used previously. The present sample should be titrated with the most current previous sample so that a change can be seen readily. A rising antibody titer does not always predict clinical significance but some institutions have established a critical titer that must be reached before amniocentesis occurs (18).

Amniotic Fluid Analysis

The amount of bilirubin in amniotic fluid is an indicator of the severity of fetal red-cell destruction. Amniotic fluid collected, generally after 27 weeks gestation, is analyzed spectrophotometrically

at wavelength levels of 350 to 700 nm. Peak absorbance of bilirubin is 450 nm. In 1969, Liley described a measurement system in which deviation from linearity at 450 nm (OD) indicates a rise in bilirubin level and increasing severity of red-cell destruction (19). Readings falling in Zone 3 or high in Zone 2 indicate severe disease (the fetus may have hydrops or will develop it within 7 to 10 days) (20). Readings in this zone may require intrauterine transfusion or, if sufficiently late in the pregnancy, immediate delivery (See Fig. 75.2). Readings in Zone 2 indicate moderate disease unless it approaches the higher end of the zone. Repeat analysis usually is required to establish a trend. Zone 1 readings indicate mild or no significant disease or anemia.

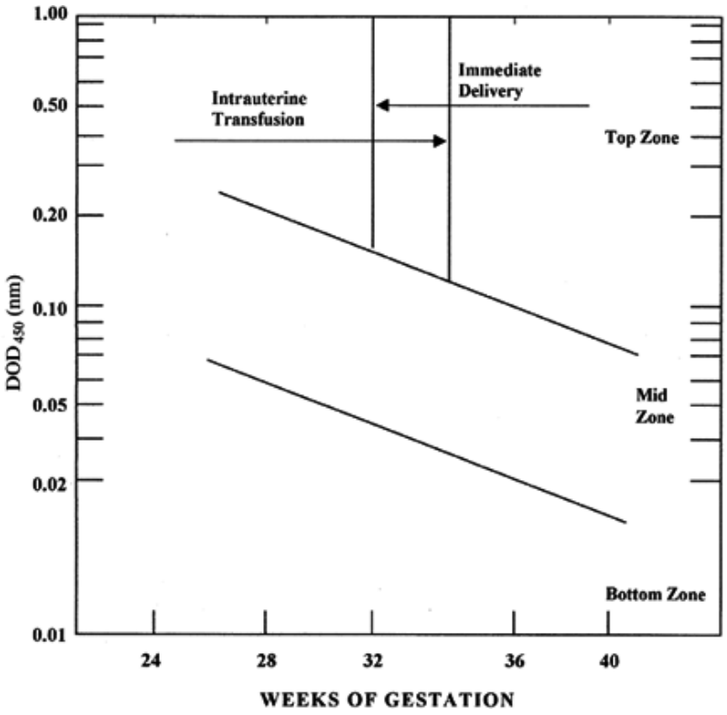


FIGURE 75.2. Liley graph for analyzing amniotic fluid. (Modified from Liley AW. Liquor amnii analysis in the management of the pregnancy complicated by Rhesus sensitization. *Am J Obstet Gynecol* 1961; 82:1359–1370, with permission.)

Percutaneous Umbilical Blood Sampling

Percutaneous umbilical blood sampling (PUBS) became possible in the mid-1980s with the availability of sophisticated ultrasound equipment and skilled perinatologists (21). A needle is directed into the umbilical vessel (usually the vein) that inserts into the placenta. Fetal blood is obtained and laboratory testing to measure immune hemolysis can be performed. PUBS is the most accurate way to determine severity of hemolysis because testing is performed on an actual sample from the fetus. In addition, the technique allowed more direct intrauterine transfusion than infusion into the infant's peritoneal cavity.

Testing at Delivery

When the baby is delivered, a cord-blood sample normally is collected from the infant. Samples usually are stored in the blood-bank/transfusion service for a minimum of 7 days. Serologic testing is performed in the following circumstances: mother is Rh(D) negative and the newborn is Rh(d) positive, mother is known to have an antibody(ies) capable of causing HDN and baby develops signs and symptoms of hemolysis.

Infants of Rh-Negative Women

Cord blood from infants of Rh-negative women should be typed for ABO and Rh(d) at delivery. If the baby is Rh(d) positive, the mother's blood is tested for fetal-maternal hemorrhage (see section on Rh immune globulin prophylaxis).

Infants of Alloimmunized Women

The DAT usually is performed on cord blood from infants born to women with known antibodies. In the case of Rh or “other” HDN, the DAT usually is strongly positive. HDN resulting from ABO antibodies usually presents with a weak or even negative DAT. An eluate may be performed to confirm the specificity of the antibody coating the infant's red cells but is not crucial as long as the causative antibody is detectable in the mother's serum.

If the DAT is positive and the mother's antibody screen is negative, ABO HDN or an antibody to a low-frequency antigen should be suspected. The easiest way to resolve this is to review the ABO types of mother and baby (see previous section on ABO HDN). An eluate can be performed to confirm that ABO antibodies are present or the infant's serum can be tested for the presence of anti-A or anti-B using an indirect antiglobulin test to detect IgG.

If ABO HDN is ruled out, an antibody to a low frequency antigen should be sought. Low frequency antigens are defined as those present in less than 1% of the population. The antibody screen performed on the mother's serum will be negative because the antibody is directed against an antigen not present on reagent red cells. To detect this antibody, a sample is obtained from the father because his red cells should possess the implicated antigen of low frequency. An eluate performed from the infant's red cells and tested against the father's red cells should be positive. In addition, the mother's serum, if ABO compatible with the father, is tested against the father's red cells. A positive result indicates that the mother has an antibody directed against an antigen present on the father's red cells.

Rh Immune Globulin Prophylaxis

Rh immune globulin (RhIG) consists of human IgG anti-D that is administered intramuscularly to Rh(d)-negative women who have had Rh(d)-positive babies. It prevents many of them from becoming sensitized to the D antigen and thus prevents formation of anti-D. Use of RhIG has decreased the rate of alloimmunization. A standard dose of RhIG consists of approximately 300 μg of anti-D. This dose is able to cover about 15 mL of D-positive red cells or approximately 30 mL of fetal whole blood.

Postpartum Administration

Rh(d)-negative women who have Rh(d)-positive babies, who have not already produced anti-D and who have not had an identified excessive fetal-maternal hemorrhage should receive one dose of RhIG within 72 hours of delivery. Occasionally, anti-D is detected in the mother and it is because of RhIG the mother received antenatally. This can be confirmed by checking

the history of the mother to determine if she received RhIG. Evaluating the serologic reactivity of the antibody also can give you clues. More important, if the antibody reactivity is weak in a saline IAT, it is likely to be passively acquired. If there is any doubt, RhIG is given.

If a large FMH occurs, one dose of RhIG will not be enough to cover the bleed. The rosette test is an excellent screening test to determine if excessive bleeding has occurred *in utero*. Red cells from a sample obtained from the mother postdelivery are incubated with anti-D. If fetal D+ red cells are present, the anti-D will bind to these cells. Indicator D+ red cells then are added and they will bind to the exposed anti-D bound to the fetal D+ cells forming rosettes of cells. These rosettes are observed by looking at the mixture microscopically. The rosette test detects about a 10-mL bleed (22). This only is a screening test so an additional test must be performed to quantify the hemorrhage.

The classic test for quantifying FMH is the Kleihauer-Betke acid elution. When a sample of blood from the mother is exposed to acid, adult hemoglobin is denatured and fetal hemoglobin is not. Fetal cells observed microscopically in a field of adult red cells are counted and the volume of bleed is estimated based on the ratio of fetal cells to adult cells observed. This method is known to be imprecise, but it is the one most widely used.

If the bleed is estimated to be less than 30 mL, the standard one vial dose of RhIG will suffice. However, if the hemorrhage is estimated to be greater, additional vials of RhIg will need to be administered. For example:

Kleihauer-Betke is reported 1.5% bleed.

(1.5/100) × 5,000 (mother's assigned blood volume) = 75 mL fetal bleed

75/30 (covered bleed) = 2.5 vials

Because this test is known to have variability, it is not unusual to give additional vials over the calculated dose and one approach follows. If the number to the right of the decimal point is less than five, round down and add one dose of RhIG (23). If the number to the left of the decimal point is 0.5 or greater, round up to the next number and add one dose of RhIG. In the above example, if the calculated dose is 2.5, it would be rounded up to three vials and one additional vial would be added for safety to bring the total to four vials.

Other methods for determining FMH include enzyme-linked antiglobulin test (ELAT) and flow cytometry, but they are not as widely used.

Routine administration of RhIG after delivery has reduced production of anti-D in pregnant women from about 13% to 1% to 2% (24).

Antepartum Administration

To further reduce the risk of alloimmunization, the American College of Obstetricians and Gynecologists (ACOG) recommends giving RhIG at 28 weeks of gestation to all Rh(d)-negative women who do not already have anti-D (25). This is based on the report that approximately 92% of women who produce anti-D during pregnancy, do so on or after 28 weeks (24). The risk of producing anti-D after receiving both antepartum and postpartum RhIG is reduced to 0.1% (24).

Transfusion Support of the Fetus or Newborn

Intrauterine Transfusion - Fetus

Intrauterine transfusion may be necessary if severe hemolysis is occurring in the fetus.

Selection of Blood

The blood selected for transfusion should be O Rh negative because in most cases, the ABO, Rh type of the fetus is not known. In addition, the red blood cells should lack the antigen(s) corresponding to any antibody the mother possesses. The blood should be as fresh as possible to ensure longer survival of transfused red blood cells. It should be irradiated (26) because the fetus does not have normal immunity. Blood provided also should be cytomegalovirus (CMV) seronegative.

The blood can be given using two different techniques, intraperitoneally, or by the umbilical vein. Intraperitoneal transfusion occurs by inserting a needle through the mother's abdominal wall into the fetus' abdominal cavity. The blood is absorbed through lymphatic channels. Direct intravascular fetal transfusion is performed by inserting a needle with the help of ultrasound into the umbilical vein. This method is more effective for the severely affected fetus with hydrops because it does not rely on absorption for the transfused cells to enter the fetal circulation.

Exchange Transfusion — Newborn

In cases of severe HDN, exchange transfusion can be performed. Exchange transfusion consists of replacing the baby's blood that contains mother's antibody, baby's coated red cells and elevated levels of bilirubin with fresh red blood cells. Exchange transfusion removes/reduces all three.

Selection of Blood

Blood selected for transfusion should be compatible with the mother's ABO antibodies as well as with any other antibody the mother may have. Maternal serum or plasma or babies serum or plasma may be used for crossmatching because both will contain the causative antibodies. Again, blood should be irradiated, usually should be CMV-negative, and as fresh as possible.

REFERENCES

1. Davenport RD, Kunkel SL. Cytokine roles in hemolytic and non-hemolytic transfusion reactions. *Transfus Med Rev* 1994;8:57–68, 85–97.

2. Firestone DT. Adverse effects of blood transfusion. In: Rudmann SV, ed. *Textbook of Blood Banking and Transfusion Medicine*, 1st ed. Philadelphia: WB Saunders Company, 1995:411.

3. Petz LD, Garratty G. *Acquired Immune Hemolytic Anemia*. New York: Churchill Livingstone, 1980:29.

4. Sokol RJ, Hewitt S, Stamps BK. Autoimmune haemolysis: An 18 year study of 865 cases referred to a regional transfusion centre. *Br Med J* 1981;282:2023–2027.

5. Issitt PD, Combs MR, Bumgarner DJ, et al. *Transfusion* 1996;36:481.

6. Issitt PD, Anstee DJ. Applied Blood Group Serology. 1998:967–968, 1010–1011.

7. Mollison PL, Engelfriet CP, Contreras M. *Blood Transfusion in Clinical Medicine* 10th ed. Oxford, England: Blackwell Scientific Productions, 1997.

8. D'Cunha P, Lord RS, Johnson ST, et al. Immune hemolytic anemia caused by sensitivity to a metabolite of the nonsteroidal anti-inflammatory drug, etodolac. *Transfusion* 2000;40(6):663–668.

9. Garratty G, ed. Chapter Drug-induced immune hemolytic anemia. In: *Immunobiology of Transfusion Medicine*, 1st ed. New York: Marcel Dekker, Inc., 1994:523–551.

10. Petz LD. Drug-induced autoimmune hemolytic anemia. *Transfus Med Rev* 1993;7:242.

11. Packman CH, Leddy JP. Drug-induced immune hemolytic anemia. In: Beutler E, Lichtman M, Collar BS, et al., eds. *Hematology*, 5th ed. New York: McGraw-Hill, 1995:691.

12. Ehmann WC. Cephalosporin-induced hemolysis: a case report and review of the literature. *Am J Hematol* 1992;40:121–125.

13. Salama A, Kroll H, Wittmann G, et al. Diclofenac-induced immune hemolytic anemias: Similarities occurrence of red blood cell autoantibodies and drug-dependent antibodies. *Br J Haematol* 1996;95:640–644.

14. Simon T. Proficiency testing program. CAP Survey 1991 J-C Northfield, IL: College of American Pathologists, 1991.

15. Miller OF, Petries SJ. Fatal erythroblastosis fetalis secondary to ABO incompatibility: report of a case. *Obstet Gynecol* 1963;22:773.

16. Vaughan JI, Warwick RM, Letsky EA, et al. *Am J Obstet Gynecol* 1994;171:247.

17. Judd WJ, Luban NLC, Ness PM, et al. Prenatal and perinatal immunohematology: Recommendations for serologic management of the fetus, newborn infant and obstetric patient. *Transfusion* 1990;30:175–183.

18. Management of isoimmunization in pregnancy. ACOG Educational Bulletin Number 227. Washington, DC: American College of Obstetricians and Gynecologists, 1996.

19. Liley AW. Liquor amnii analysis in the management of the pregnancy complicated by Rhesus sensitization. *Am J Obstet Gynecol* 1961;82:1359–1370.

20. Bowman JM. Hemolytic disease of the newborn. In: Garratty G, ed. *Immunobiology of Transfusion Medicine*, 1st ed., New York: Marcel Dekker, 1994:570.

21. Daffos F, Capella-Pavlovsky M, Forester F. Fetal blood sampling during pregnancy with use of a needle guided by ultrasound: a study of 606 consecutive cases. *Am J Obstet Gynecol* 1985;153:655.

22. Sebring ES, Poleski HF. Detection of fetal maternal hemorrhage in Rh immune globulin candidates. *Transfusion* 1982;22:468–471.

23. Vergelen-Tyler, ed. *Technical Manual*, 13th ed. Bethesda, Maryland: American Association of Blood Banks, 1999:507–508.

24. Bowman JM. The prevention of Rh immunization. *Transf Med Ref* 1988;2:129–150.

25. Prevention of D isoimmunization. ACOG Technical Bulletin Number 147. Washington, DC: American College of Obstetricians and Gynecologists, 1990.

26. Menitove JE, ed. *Standards for Blood Banks and Transfusion Services*, 19th ed. Bethesda, Maryland: American Association of Blood Banks, 1999.

27. Reid ME, Lomas-Francis C. *Blood Group Antigen Facts Book*, San Diego, California: Academic Press, 1997.

SUBJECT INDEX

Page numbers followed by f indicate figures; page numbers followed by t indicate tables.

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

Q

R

S

T

U

V

W

X

Y

Z